

Een analyse van het leesgedrag van Nederlandse basisschoolleerlingen  
Een onderzoek naar leesfouten en leesstrategieën tijdens hardop voorlezen



Sanne Limonard  
S4646010  
Masterscriptie Taal- en Spraakpathologie  
Radboud Universiteit Nijmegen  
Augustus 2019

Dr. Helmer Strik (scriptiebegeleider)  
Dr. Catia Cucchiarini (tweede lezer)

## Voorwoord

Voor u ligt de masterscriptie ‘Een analyse van het leesgedrag van Nederlandse basisschoolleerlingen. Een onderzoek naar leesfouten en leesstrategieën tijdens hardop voorlezen’. Deze masterscriptie is geschreven in het kader van mijn afstuderen aan de masteropleiding Taal- en Spraakpathologie aan de Radboud Universiteit te Nijmegen.

Het onderwerp van deze scriptie werd aangeleverd door dr. Helmer Strik en dr. Catia Cucchiarini van het Centre for Language and Speech Technology (CLST) van de Radboud Universiteit te Nijmegen, aanvragers van het Dutch ASR-based Reading Tutor (DART)-project dat in oktober 2018 van start ging. Bij dit project zijn ook het Kohnstamm Instituut van de Universiteit van Amsterdam, Uitgeverij Zwijsen, NovoLanguage en scholen betrokken. Het DART-project is een driejarig onderzoeksproject waarbinnen leessoftware met automatisch spraakherkenning wordt ontwikkeld voor basisschoolleerlingen, zodat automatische feedback op hardop voorlezen geboden kan worden. Om dit te kunnen ontwikkelen, moet duidelijk zijn wat voor soort leesfouten leerlingen maken. In Nederland is nog maar weinig onderzoek gedaan naar leesfouten en leesstrategieën tijdens het hardop voorlezen van een tekst. Na het annoteren van spraakopnames en een kwantitatief onderzoek, heb ik hier meer informatie over kunnen geven.

Bij dezen wil ik graag dr. Helmer Strik, dr. Catia Cucchiarini, dr. Marjoke Bakker, prof. dr. Roeland van Hout en dr. Esther Janse bedanken voor hun begeleiding en feedback tijdens het schrijven van deze scriptie. Ook wil ik Emma de Reus bedanken voor de samenwerking tijdens het opstellen van het annotatieprotocol. Daarnaast wil ik mijn vriendinnen Ellen van Sambeek en Mirjam Otters bedanken voor hun wijze raad en feedback. Tot slot wil ik mijn ouders bedanken die het voor mij mogelijk hebben gemaakt deze masteropleiding te volgen.

Ik wens u veel leesplezier toe.

Sanne Limonard

Nijmegen, juli 2019

## Inhoudsopgave

Voorwoord .....	I
Inhoudsopgave.....	II
Samenvatting .....	IV
Definities .....	V
1. Inleiding .....	1
2. Theoretische achtergrond .....	5
2.1 Fasen van geletterdheid .....	5
2.1.1 Ontluikende geletterdheid .....	5
2.1.2 Beginnende geletterdheid .....	5
2.1.3 Gevorderde geletterdheid .....	6
2.2 Informatieniveaus .....	6
2.3 Leesmodellen.....	6
2.3.1 Duale-route model.....	7
2.3.2 Parallele gedistribueerde verwerkingsmodel.....	7
2.3.3 Fonologisch coherentiemodel.....	7
2.4 Leesstrategieën .....	8
2.4.1 Elementaire leeshandeling.....	8
2.4.2 Lezen met behulp van clusters en spellingpatronen .....	9
2.4.3 Lezen met behulp van de visuele woordvorm .....	9
2.4.4 Lezen met behulp van morfologische analyse.....	9
2.4.5 Lezen met behulp van de context .....	9
2.5 Deelvaardigheden van het lezen .....	9
2.6. Leesproblemen .....	10
2.6.1 Oorzaken van leesproblemen .....	11
2.6.2 Soorten leesproblemen .....	11
2.7 Evalueren van de leesontwikkeling .....	13
2.7.1 Toetsen van de taalvaardigheid .....	14
2.7.2 Toetsen van deelvaardigheden .....	14
2.7.3 Toetsen van de leesvaardigheid.....	14
2.7.4 Analyse van toetsingsgegevens .....	15
2.7.5 Eerder onderzoek naar leesfouten en leesstrategieën .....	16
2.8 Hypotheses .....	18
3. Methode.....	19
3.1 Onderzoeksontwerp.....	19
3.2 Materiaal.....	19

3.2.1	Spraakmateriaal JASMIN-CGN Corpus .....	19
3.2.2	Annotatieprotocol .....	19
3.3	Proefpersonen .....	20
3.4	Procedure .....	21
3.5	Data-analyse .....	21
3.5.1	Verband tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau .....	22
3.5.2	Frequenties van leesfouten .....	22
3.5.3	Verband tussen leesfouten en leeftijd, geslacht en AVI-niveau .....	22
3.5.4	Frequenties van leesstrategieën .....	23
3.5.5	Verband tussen leesstrategieën en leeftijd, geslacht en AVI-niveau .....	23
4.	Resultaten .....	24
4.1	Verband tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau .....	24
4.2	Frequenties van leesfouten .....	25
4.2.1	Leesfouten in alle lees pogingen .....	25
4.2.2	Leesfouten in laatste lees pogingen .....	26
4.3	Verband tussen leesfouten en leeftijd, geslacht en AVI-niveau .....	26
4.3.1	Leesfouten in alle lees pogingen .....	26
4.3.2	Leesfouten in laatste lees pogingen .....	27
4.4	Frequenties van leesstrategieën .....	27
4.4.1	Leesstrategieën in alle lees pogingen .....	27
4.4.2	Leesstrategieën in laatste lees pogingen .....	28
4.5	Verband tussen leesstrategieën en leeftijd, geslacht en AVI-niveau .....	28
4.5.1	Leesstrategieën in alle lees pogingen .....	28
4.5.2	Leesstrategieën in laatste lees pogingen .....	29
5.	Discussie .....	30
5.1	Verband tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau .....	30
5.2	Frequenties van leesfouten .....	31
5.3	Verband tussen leesfouten en leeftijd, geslacht en AVI-niveau .....	34
5.4	Frequenties van leesstrategieën .....	36
5.5	Verband tussen leesstrategieën en leeftijd, geslacht en AVI-niveau .....	37
6.	Conclusie .....	39
	Bronnen .....	40
	Bijlage 1: Protocol annotatie leesstrategieën en leesfouten JASMIN-CGN Corpus .....	44
	Bijlage 2: PRAAT-script .....	50

## Samenvatting

In Nederland ervaart een groot aantal leerlingen problemen met het leren lezen (Inspectie van het Onderwijs, 2018). Aan het eind van groep 3 blijkt zo'n 15% van de leerlingen niet goed te kunnen lezen (Struiksma, in Huizenga, 2016, p. 191). Onder allochtone leerlingen is de leesuitval nog groter (Verhoeven, 2002, p. 23). Steeds minder leerlingen halen aan het eind van de basisschool het streefniveau lezen (Inspectie van het Onderwijs, 2018). In 2017 behaalde slechts 65 procent van de leerlingen in het basisonderwijs het streefniveau. Dit is bijna 10 procent lager dan in 2016 (Inspectie van het Onderwijs, 2018).

Een belangrijke oefenvorm voor de ontwikkeling van leesvaardigheid is begeleid hardop voorlezen. Deze oefenvorm kan op scholen echter beperkt ingezet worden, omdat het veel individuele aandacht vergt. Leessoftware met automatische spraakherkenning biedt hier mogelijk een oplossing voor (Strik, 2019). Om dergelijke software te kunnen ontwikkelen, moet duidelijk zijn wat voor soort leesfouten leerlingen maken. In Nederland is nog maar weinig onderzoek gedaan naar leesfouten en leesstrategieën tijdens het hardop voorlezen van een tekst.

Het doel van het huidige onderzoek was te achterhalen welke leesstrategieën Nederlandse basisschoolleerlingen inzetten tijdens het hardop voorlezen van een tekst en welke leesfouten zij daarbij maken. De volgende onderzoeksvragen werden opgesteld:

1. *Is er een verband tussen leesaccuratesse en de variabelen leeftijd, geslacht en AVI-niveau?*
2. *Wat zijn de meest frequente leesfouten die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar maken tijdens het hardop voorlezen van een Nederlandse tekst?*
3. *Is er een verband tussen leesfouten en de variabelen leeftijd, geslacht en AVI-niveau?*
4. *Wat zijn de meest frequente leesstrategieën die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar toepassen tijdens het hardop voorlezen van een Nederlandse tekst?*
5. *Is er een verband tussen leesstrategieën en de variabelen leeftijd, geslacht en AVI-niveau?*

Om antwoord te kunnen geven op de onderzoeksvragen werd een kwantitatief cross-sectioneel onderzoek met bestaande data uitgevoerd. Voor de analyses werd gebruikgemaakt van het JASMIN-CGN Corpus (Jongeren, Anderstaligen en Senioren in Mens-machine Interactie – Corpus Gesproken Nederlands), dat spraakopnames bevat van Nederlandse basisschoolleerlingen die hardop een tekst op beheersingsniveau voorlezen. Allereerst werd een annotatieprotocol opgesteld om leesfouten en leesstrategieën te kunnen analyseren. In het spraakmateriaal van 71 Nederlandse basisschoolleerlingen uit groep 4 tot en met groep 8 werden leesfouten en leesstrategieën geannoteerd. Vervolgens werden de annotaties geanalyseerd. Zoals verwacht, bleek dat leesaccuratesse tijdens het hardop voorlezen van een Nederlandse tekst op beheersingsniveau niet in verband gebracht kon worden met leeftijd, geslacht en AVI-niveau. Leerlingen in Nederland lezen teksten op hun eigen AVI-niveau, waardoor slechts 5% of minder van de woorden uit de tekst fout wordt gelezen. De leesfouten die de leerlingen maakten, bestonden voornamelijk uit vervangingen door orthografisch gelijke woorden, semantisch gelijke of semantisch passende woorden, pseudoworden of woorden die zowel orthografisch als semantisch niet gelijk waren aan het doelwoord. De leesfouten die sublexicale strategieën suggereren, namen proportioneel af naarmate de leerlingen ouder werden en een hoger AVI-niveau beheersten, terwijl leesfouten die lexicale strategieën suggereren proportioneel toenamen. Meisjes vervingen ten opzichte van jongens proportioneel minder vaak woorden door pseudoworden en semantisch gelijke of semantisch passende woorden. Daarnaast bleek dat de Nederlandse basisschoolleerlingen voornamelijk lexicale strategieën gebruikten tijdens het hardop voorlezen van een tekst. De sublexicale strategieën namen proportioneel af naarmate de leerlingen ouder werden en een hoger AVI-niveau beheersten. Meisjes lazen proportioneel gezien minder woorden fout dan jongens.

## Definities

<b>Begrip</b>	<b>Uitleg</b>
<i>Alle lees pogingen</i>	Soms worden er meerdere pogingen gedaan om een woord te lezen. Alle lees pogingen duiden op alle pogingen die gedaan worden om het woord te lezen (bijvoorbeeld wanneer het woord ‘appelmoes’ wordt gelezen als ‘a – appel – appelmoes’, dan werden er drie pogingen gedaan om het woord te lezen).
<i>Analyse van het leesgedrag</i>	Een analyse van het leesgedrag is een kwalitatieve analyse, waarbij gekeken wordt naar de manier waarop wordt gelezen en de verschillende soorten <i>leesfouten</i> die gemaakt worden.
<i>Annotaties</i>	Annotaties zijn aantekeningen bij spraakmateriaal die gemaakt kunnen worden in bijvoorbeeld het computerprogramma PRAAT (Boersma & Weenink, Versie 6.0.52). In het huidige onderzoek werden annotaties gemaakt aan de hand van een annotatieprotocol (zie bijlage 1).
<i>AVI</i>	AVI staat voor Analyse van Individualiseringsvormen. Het AVI-systeem maakt een indeling van teksten op basis van de moeilijkheidsgraad. Daarnaast kan door middel van AVI-toetsen de leesvaardigheid van een leerling bepaald worden.
<i>Corpus</i>	Een corpus is een verzameling van teksten of spraakopnames in een bepaalde taal, die gebruikt wordt voor onderzoek.
<i>Laatste lees poging</i>	Soms worden er meerdere pogingen gedaan om een woord te lezen. De laatste lees poging duidt op de laatste poging die gedaan wordt om het woord te lezen (bijvoorbeeld wanneer het woord ‘appelmoes’ wordt gelezen als ‘a – appel – appelmoes’, dan is appelmoes de laatste poging). Wanneer er slechts één poging gedaan wordt om een woord te lezen, dan is deze poging de laatste lees poging.
<i>Labels</i>	Het begrip labels duidt op de cijfers en letters in de annotaties die gebruikt worden om bepaalde <i>leesfouten</i> en <i>leesstrategieën</i> mee aan te duiden (zie bijlage 1 voor de betekenissen van de labels).
<i>Leesaccuratesse</i>	Leesaccuratesse geeft de verhouding aan tussen het aantal goed gelezen woorden in de laatste lees poging ten opzichte van het totaal aantal woorden in een tekst. Een leesaccuratesse van bijvoorbeeld .95 betekent dat 95% van de woorden in een tekst correct is gelezen.
<i>Leesfout</i>	Een leesfout is een woord dat volledig, maar niet correct wordt voorgelezen (bijvoorbeeld ‘riepen’ in plaats van ‘roepen’). Er is géén sprake van een leesfout wanneer een woord nog niet volledig wordt voorgelezen (bijvoorbeeld wanneer ‘roepen’ wordt gelezen als ‘roep-roepen’, dan wordt ‘roep’ niet als leesfout gerekend).
<i>Leesprobleem</i>	Er is sprake van een leesprobleem wanneer het automatiseren van de leestechniek niet goed op gang komt of achteruitgaat (zie 2.6).

<i>Leesstrategie</i>	Een leesstrategie is een strategie om een tekst te lezen. Er zijn verschillende leesstrategieën, namelijk de elementaire leeshandeling (bijvoorbeeld <i>b...a...l, bal</i> ), lezen met behulp van clusters en spellingpatronen (bijvoorbeeld <i>m...aan, maan</i> ) en lezen met behulp van de visuele woordvorm, waarbij woorden in een keer worden herkend (bijvoorbeeld <i>Emma</i> ). Daarnaast kan gelezen worden met behulp van een morfologische analyse (bijvoorbeeld <i>be-hulp-zaam</i> ) en met behulp van de context (bijvoorbeeld <i>ze bibberden van ...</i> ) (zie 2.4). De leesstrategieën kunnen onderverdeeld worden in <i>lexicale leesstrategieën</i> en <i>sublexicale leesstrategieën</i> (zie 2.3).
<i>Lexicale leesstrategie</i>	De lexicale leesstrategie is een leesstrategie waarbij woorden vrijwel direct worden herkend op basis van het unieke patroon van grafemen dat het woord van alle andere woorden onderscheidt. De veronderstelling is dat de lexicale leesstrategie gebruikt wordt bij het lezen van bekende woorden (zie 2.3).
<i>Orthografisch gelijk</i>	Orthografisch gelijke woorden zijn woorden waarbij slechts één grafeem of het achtervoegsel -en van het doelwoord wordt weggelaten, vervangen of toegevoegd (bijvoorbeeld stad/stap).
<i>Pseudowoord</i>	Een pseudowoord is een niet bestaand woord, waarbij de volgorde van letters wel mogelijk is (bijvoorbeeld het pseudowoord 'flof' dat geen betekenis heeft, maar waarbij de lettervolgorde in het Nederlands wel mogelijk is).
<i>Semantisch gelijk</i>	Semantisch gelijke woorden zijn woorden met dezelfde betekenis (zoals 'kat' en 'poes').
<i>Sublexicale leesstrategie</i>	De sublexicale leesstrategie is een leesstrategie waarbij woorden worden gelezen door grafeem-naar-foneemregels toe te passen, gevolgd door het samenvoegen van losse fonemen of klankgroepen tot een woord. De veronderstelling is dat de sublexicale leesstrategie gebruikt wordt bij het lezen van onbekende of laagfrequente woorden en <i>pseudoworden</i> (zie 2.3).

## 1. Inleiding

In Nederland zijn 2,5 miljoen mensen laaggeletterd. Deze mensen hebben moeite met lezen en schrijven en zijn onvoldoende in staat zich in het dagelijks leven te redden met lezen en schrijven (Huizenga, 2016; Stichting Lezen en Schrijven, z.d.). Daarnaast zijn er in Nederland ongeveer 250.000 mensen analfabeet. In tegenstelling tot laaggeletterde mensen kunnen zij helemaal niet lezen en schrijven (Stichting Lezen en Schrijven, z.d.). De vaardigheid om op een effectieve en efficiënte manier te lezen en te schrijven wordt gevorderde geletterdheid genoemd (Aarnoutse, 2004). Een voorwaarde om tot gevorderde geletterdheid te komen, is het kunnen decoderen van woorden en lezen van teksten (Murnane, Sawhill, & Snow, 2012). Gevorderde geletterdheid is voor volwassenen in de eenentwintigste eeuw een voorwaarde voor succes (Murnane, Sawhill, & Snow, 2012). Met gevorderde geletterdheid is men namelijk in staat toegang te krijgen tot kennis, informatie van verschillende bronnen samen te voegen en te leren over nieuwe onderwerpen (Murnane, Sawhill, & Snow, 2012). Wanneer je niet goed kunt lezen of schrijven, kun je in onze geletterde maatschappij niet goed functioneren (Aarnoutse, 2004; Huizenga, 2016). Er gaat veel aan je voorbij als je niet goed kunt lezen en schrijven. Van reclameborden tot straatnamen en van gebruiksaanwijzingen tot de reisinformatie van het openbaarvervoer. Overal om ons heen zijn letters (Huizenga, 2016). Daarnaast kunnen veel beroepen niet uitgeoefend worden als je niet kunt lezen, omdat je bijvoorbeeld e-mails, notulen en bestelformulieren niet kunt begrijpen (Huizenga, 2016). Ook een andere belangrijke functie van geletterdheid loop je mis als je niet kunt lezen, namelijk ontspanning. Geletterdheid is een voorwaarde voor het kunnen lezen van boeken of stripverhalen, maar ook tijdens het kijken van televisie of films wordt er een beroep gedaan op leesvaardigheid (Huizenga, 2016).

Ook voor kinderen is het van belang dat zij goed kunnen lezen (Huizenga, 2016). Leesvaardigheid is bepalend voor het schoolsucces. Bijna alle vakken op school doen namelijk een beroep op de leesvaardigheid (Huizenga, 2016). Niet goed kunnen lezen heeft grote gevolgen voor de gehele schoolperiode en tevens voor het functioneren in onze geletterde maatschappij (Wentink & Verhoeven, 2008). Leerlingen die niet goed kunnen lezen, ondervinden doorgaans problemen met het begrijpend lezen en tijdens andere vakken waarbij teksten een rol spelen (Struiksmā, van der Leij, & Vieijra, 2009; Verhoeven, 2002). Leerlingen goed leren lezen en schrijven is daarom één van de belangrijkste taken van het basisonderwijs (Huizenga, 2016). Het Nederlandse leesonderwijs is zo georganiseerd dat elke leerling leert lezen op zijn eigen niveau. Door middel van AVI-toetsen wordt de leesvaardigheid van leerlingen vastgesteld waarna ze boeken krijgen die passend zijn bij hun AVI-niveau. Op deze manier wordt voorkomen dat leerlingen gefrustreerd raken door te moeilijke teksten. Echter, niet alle leerlingen halen een AVI-niveau dat passend is bij hun leeftijd. Een groot aantal leerlingen ervaart problemen met het leren lezen (Inspectie van het Onderwijs, 2018). Aan het eind van groep 3 blijkt zo'n 15% van de leerlingen niet goed te kunnen lezen (Struiksmā, in Huizenga, 2016, p. 191). In groep 4 blijkt 10% van de achtjarige leerlingen niet goed te kunnen lezen (Huizenga, 2016). Onder allochtone leerlingen is de leesuitval nog groter (Verhoeven, 2002, p. 23). Het aantal leerlingen dat eind groep 8 niet goed kan lezen, is de afgelopen jaren gestegen. Steeds minder leerlingen halen aan het eind van de basisschool het streefniveau lezen (Inspectie van het Onderwijs, 2018). Het streefniveau voor eind groep 3 is AVI-2, voor eind groep 4 AVI-5 en voor eind groep 5/begin groep 6 AVI-8/9 (Inspectie van het Onderwijs, 2006). In 2017 behaalde slechts 65 procent van de leerlingen in het basisonderwijs het streefniveau. Dit is bijna 10 procent lager dan in 2016 (Inspectie van het Onderwijs, 2018). Bij jongens daalden de leesprestaties meer dan bij meisjes (Inspectie van het Onderwijs, 2018).



De verschillen in leesvaardigheid tussen Nederlandse basisscholen zijn groot, zelfs bij vergelijkbare leerlingenpopulaties. Bij scholen zonder leerlingen met een migratie-achtergrond varieert het percentage leerlingen dat het streefniveau haalt tussen de 47 en 86 procent. Bij scholen waar alle leerlingen een migratie-achtergrond hebben, varieert het percentage leerlingen dat het streefniveau haalt tussen de 18 en 69 procent (Inspectie van het Onderwijs, 2018). Daarnaast halen steeds minder leerlingen het fundamentele niveau lezen, waardoor zij de basisschool laaggeletterd verlaten (Inspectie van het Onderwijs, 2018). Het percentage leerlingen dat de basisschool laaggeletterd verliet, steeg van 1.4 procent in 2015 tot 2.2 procent in 2017, wat overeenkomt met drieënhalf duizend leerlingen (Inspectie van het Onderwijs, 2018). In tegenstelling tot Nederland verbeterde de leesvaardigheid in de meeste landen. Dit heeft ertoe geleid dat Nederland niet meer in de internationale top staat (Inspectie van het Onderwijs, 2018).

Dat steeds minder leerlingen in het basisonderwijs het streefniveau lezen behalen, wordt mogelijk veroorzaakt doordat leerkrachten van basisscholen er onvoldoende in slagen om te gaan met verschillen tussen leerlingen (Inspectie van het Onderwijs, 2008). In de praktijk blijkt dat een groot deel van de leerkrachten de leesproblemen van de leerlingen kent, maar onvoldoende of geen tijd heeft om er iets aan te doen (Stokking, in Ruissenaars & Ghesquière, 2002, p. 25). Daarnaast geven veel leerkrachten aan dat ze leerlingen met leesproblemen niet adequaat kunnen begeleiden, omdat ze niet over de relevante kennis beschikken. Het is voor hen onduidelijk hoe leesproblemen gesignaleerd kunnen worden en hoe zij gerichte interventies kunnen inzetten (Wentink & Verhoeven, 2008).

Basisscholen zijn wettelijk verplicht om de vorderingen van hun leerlingen in de gaten te houden. Ze spelen daarom een belangrijke rol bij het signaleren van leesproblemen. Wanneer een leerling te langzaam vooruitgaat, stilstaat of achteruitgaat in zijn leesontwikkeling, moet achterhaald worden wat hiervan de oorzaak is (Braams, 2007). Leesproblemen zijn vaak al vroeg in de ontwikkeling van leerlingen zichtbaar (Verhoeven, 2002). Zo kan 85% van de mogelijke risicolezers al zeer vroeg opgespoord worden (Vernooy, in Blonk et al., 2006, p. 52). Wanneer leesproblemen in een zo vroeg mogelijk stadium van het leesonderwijs onderkend en aangepakt worden, kunnen ernstige leesachterstanden voorkomen of teruggedrongen worden (Wentink & Verhoeven, 2008). Deze vroegtijdige signalering en behandeling van leesproblemen is effectiever dan het laat signaleren en repareren van leesproblemen (Vernooy, in Blonk et al., 2006, p. 52). Leerlingen die problemen ondervinden met lezen moeten voor hun negende jaar geholpen worden, omdat de leesproblemen daarna veel lastiger te behandelen zijn (Blonk et al., 2006). In de kleuterperiode kunnen problemen met de verwerking van spraakklanken, ook wel fonologische verwerking genoemd, al gesignaleerd worden. Dit kan op basis van specifieke taken, zoals woorddiscriminatie of rijmen (Snow, Burns, & Griffin, in Ruissenaars & Ghesquière, 2002, p. 24). Met een gerichte interventie op het gebied van fonologisch bewustzijn kan geprobeerd worden het risico op leesproblemen te verkleinen (Borström & Elbro, in Ruissenaars & Ghesquière, 2002, p. 24). Vanaf groep 3 kan worden vastgesteld of er sprake is van leesproblemen door middel van (methodegebonden) toetsen en observaties (Verhoeven, 2002).

Op basis van toetsen en observaties worden leerlingen globaal onderverdeeld in spellende en radende lezers of lezers die een combinatie van zowel spellen als raden gebruiken (Paternotte & Oostewechel, 2015; Wentink & Verhoeven, 2008). Daarnaast kunnen leerlingen onderverdeeld worden in drie subgroepen, namelijk zwakke, gemiddelde en gevorderde lezers (Benoit et al., z.d.). Bij de onderverdeling van leerlingen op basis van hun leesprestaties is met name aandacht voor accuratesse en leessnelheid. Er is echter weinig aandacht voor de typen leesfouten die de leerlingen maken, terwijl de leesproblemen bepalend zijn voor de invulling van leesinterventies (Wentink & Verhoeven, 2008).

Op basis van observaties en kwalitatieve analyses van toetsingsgegevens kan iets gezegd worden over de aard van de leesproblemen (Huizenga, 2016). Dit kan gedaan worden door een analyse van het leesgedrag te maken. Zo'n analyse geeft een gedetailleerd inzicht in de leesfouten die een leerling maakt en de leesstrategieën die gebruikt worden. Het leesgedrag wordt geanalyseerd door dat wat de leerling hardop voorleest te vergelijken met de tekst die wordt gelezen (Wentink & Verhoeven, 2008). In de klas kunnen zulke analyse gemaakt worden om te bepalen welke (deel)vaardigheden de leerling nog moet oefenen. Daarnaast kunnen er gedurende een bepaalde periode meerdere analyses worden gemaakt, om zichtbaar te maken of de leerling vooruitgaat en of bepaalde problemen blijven bestaan (Wentink & Verhoeven, 2008). Ook voor onderzoekers kunnen leesfoutenanalyses interessante inzichten opleveren. Wanneer bij meerdere leerlingen analyses van het leesgedrag tijdens hardop voorlezen gemaakt worden, geeft dit inzicht in welke leesfouten leerlingen maken en welke leesstrategieën ze gebruiken (Wentink & Verhoeven, 2008). Dit vergroot de kennis over het verwerken van schriftelijke informatie (Leu, 1982) en het geeft informatie over hoe leerlingen leren lezen (Goikoetxea, 2006). Daarnaast geven de analyses informatie over welke onderdelen binnen het leesonderwijs verbeterd kunnen worden (Au, 1977; Chinn, 1933). Op basis van analyses van het leesgedrag is het mogelijk interventies te ontwikkelen die het leesproces van leerlingen verbeteren (Laing, 2002). Tevens kan onderzocht worden wat de overeenkomsten en verschillen zijn in leesgedrag tussen zwakke en sterke lezers (Cleuren, 2009; Weber, 1970), tussen verschillende tijdstippen in het leesonderwijs (Cleuren, 2009; Schale, 1966; Weber, 1970) en tussen jongens en meisjes (Cleuren, 2009). Daarnaast kunnen foutenanalyses gebruikt worden om leessoftware met automatische spraakherkenning te verbeteren (Cleuren, 2009).

In het huidige onderzoek werd een protocol opgesteld voor het analyseren van leesfouten en leesstrategieën. Dergelijke analyses werden in het huidige onderzoek gemaakt aan de hand van het JASMIN-CGN Corpus (Jongeren, Anderstaligen en Senioren in Mens-machine Interactie – Corpus Gesproken Nederlands). Dit corpus bevat onder andere spraakopnames van Nederlandstalige basisschoolleerlingen die hardop een tekst lezen op beheersingsniveau (Cucchiarini et al., 2008). De spraakopnames van Nederlandse basisschoolleerlingen werden beluisterd en leesfouten en leesstrategieën werden geannoteerd. Een deel van de resultaten kan gebruikt worden voor het Dutch ASR-based Reading Tutor (DART)-project, dat in oktober 2018 van start ging. Dit is een driejarig onderzoeksproject waarbinnen leessoftware met automatische spraakherkenning wordt ontwikkeld, zodat automatische feedback op hardop voorlezen geboden kan worden (Strik, 2019). Leerlingen kunnen met deze leessoftware oefenen met hardop voorlezen op hun eigen AVI-niveau. Hardop voorlezen onder begeleiding is een belangrijke oefenvorm voor de ontwikkeling van technische leesvaardigheid. Deze oefenvorm kan echter op scholen beperkt ingezet worden, omdat het veel individuele aandacht vergt. Leessoftware met automatische spraakherkenning biedt hier mogelijk een oplossing voor (Strik, 2019). Op basis van foutenanalyses kunnen veelvoorkomende leesfouten gekarakteriseerd en gecategoriseerd worden, die vervolgens in de leessoftware geïmplementeerd kunnen worden. Op basis hiervan kan de leessoftware leesfouten voorspellen, detecteren en herstellen (Cleuren, 2009; Mostow et al., 2002).

In Nederland is nog maar weinig onderzoek gedaan naar leesfouten en leesstrategieën tijdens het hardop voorlezen van een tekst. Het huidige onderzoek had doel te achterhalen welke leesstrategieën Nederlandse basisschoolleerlingen inzetten tijdens het hardop voorlezen van een tekst en welke leesfouten zij daarbij maken. Er werd verwacht dat er geen effect gevonden zou worden van leeftijd, geslacht en AVI-niveau op leesaccuratesse, omdat leerlingen teksten lezen op beheersingsniveau. Om dit te kunnen controleren, werd de volgende onderzoeksvraag opgesteld:

1. *Is er een verband tussen leesaccuratesse en de variabelen leeftijd, geslacht en AVI-niveau?*

Vervolgens werd onderzoek gedaan naar leesfouten en leesstrategieën om inzicht te krijgen in het leesgedrag van Nederlandse basisschoolleerlingen. Deze informatie is onder andere bruikbaar voor het ontwikkelen van leessoftware met automatische spraakherkenning, zoals in het DART-project. De volgende onderzoeksvragen werden opgesteld:

2. *Wat zijn de meest frequente leesfouten die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar maken tijdens het hardop voorlezen van een Nederlandse tekst?*
3. *Is er een verband tussen leesfouten en de variabelen leeftijd, geslacht en AVI-niveau?*
4. *Wat zijn de meest frequente leesstrategieën die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar toepassen tijdens het hardop voorlezen van een Nederlandse tekst?*
5. *Is er een verband tussen leesstrategieën en de variabelen leeftijd, geslacht en AVI-niveau?*

## 2. Theoretische achtergrond

### 2.1 Fasen van geletterdheid

Leesvaardigheid ontwikkelt zich bij de meeste kinderen niet vanzelf. Leren lezen is namelijk geen natuurlijk proces dat spontaan op gang komt (Huizenga, 2016; Inspectie van het Onderwijs, 2006). Daarom hebben bijna alle kinderen gericht leesonderwijs nodig, waarin verschillende fasen worden doorlopen (Huizenga, 2016; Inspectie van het Onderwijs, 2006). Basisscholen besteden gemiddeld per week zo'n 110 minuten aan leesonderwijs en in groep 3 zelfs nog meer (Huizenga, 2016). De ontwikkeling van geletterdheid wordt opgedeeld in drie fasen: ontluikende geletterdheid, beginnende geletterdheid en gevorderde geletterdheid (Verhoeven & Aarnoutse, 1999).

#### 2.1.1 Ontluikende geletterdheid

De ontluikende geletterdheid heeft betrekking op de voorschoolse periode, waarin kinderen worden voorbereid op het lezen. Ze maken in deze fase kennis met geschreven taal door onder andere prentenboeken, voorleesverhalen en spelletjes met letters. Kinderen krijgen in deze fase belangstelling voor het geschreven woord (Huizenga, 2016; Verhoeven & Aarnoutse 1999). In deze fase wordt de basis gelegd voor de ontwikkeling van leesvaardigheid (Verhoeven & Aarnoutse, in Verhoeven & van Leeuwe, 2003, p. 259).

#### 2.1.2 Beginnende geletterdheid

De beginnende geletterdheid is de fase van groep 1 tot en met groep 3. Bij de meeste leerlingen start de fase van beginnende geletterdheid echter al voordat ze naar de basisschool gaan (Verhoeven, 2002). In deze fase oriënteren leerlingen zich verder op de functies van geschreven taal (Verhoeven, 2002; Verhoeven & van Leeuwe, 2003). Al in de kleuterperiode kan de ontwikkeling van geletterdheid gestimuleerd worden, door aandacht te besteden aan geletterdheid, mondelinge taalvaardigheid en woordenschat. Dit kan door interactief voorlezen, het bekijken en lezen van prentenboeken, thematafels, taalspelletjes en kringgesprekken. Dit stimuleert het verhaalbegrip en de interesse in lezen (Inspectie van het Onderwijs, 2006). Daarnaast wordt de ontwikkeling van leesvaardigheid bevorderd door gerichte aandacht voor fonologische vaardigheden. Dat zijn vaardigheden die leerlingen nodig hebben om fonemen, ook wel spraakklanken genoemd, te herkennen en van elkaar te onderscheiden (Inspectie van het Onderwijs, 2006). Leerlingen gaan leesgedrag kopiëren, leren hun eigen naam en andere bekende woorden herkennen en een aantal leerlingen leert min of meer spontaan nieuwe woorden lezen door letters samen te voegen (Verhoeven & van Leeuwe, 2003).

Kleuters ontdekken door middel van taal- en (voor)leesactiviteiten dat er een relatie bestaat tussen geschreven en gesproken taal. Daarnaast ontdekken ze dat naast gesproken taal ook via geschreven taal gecommuniceerd kan worden (Verhoeven, 2002). In de fase van beginnende geletterdheid leren leerlingen dat woorden zijn opgebouwd uit klanken en dat deze klanken worden weergegeven door letters uit het alfabet (Verhoeven, 2002; Verhoeven & Aarnoutse, 1999). Ze leren het principe van het alfabetisch schrift. Bij sommige leerlingen ontwikkelt dit spontaan in groep 2 (Verhoeven, 2002). Aan het eind van groep 2 kennen leerlingen in Nederland gemiddeld 10 letters (Inspectie van het Onderwijs, 2006). In de praktijk blijkt een groot aantal leerlingen moeite te hebben met het alfabetisch principe. Deze leerlingen leren het principe van het alfabetisch schrift in groep 3, op het moment dat de formele leesinstructie begint (Verhoeven, 2002). Op een gestructureerde manier wordt het alfabetisch principe geleerd. Wanneer leerlingen fonologische vaardigheden hebben ontwikkeld en een relatie kunnen leggen tussen grafemen en fonemen, leren zij vanuit daar woorden te lezen. De fase van beginnende geletterdheid beperkt zich tot klankzuivere woorden met een gemakkelijke MKM-structuur (medeklinker-klinker-medeklinker). In

klankzuivere woorden geeft elk grafeem een apart foneem weer (Verhoeven & van Leeuwe, 2003).

### 2.1.3 Gevorderde geletterdheid

De gevorderde geletterdheid heeft betrekking op de periode van groep 4 tot en met groep 8 (Verhoeven & Aarnoutse, 1999). Het lezen van gemakkelijke MKM-woorden wordt in deze fase uitgebreid naar woorden met medeklinkerclusters en meerlettergrepige woorden (Verhoeven & van Leeuwe, 2003). In de fase van gevorderde geletterdheid is de leesinstructie gericht op het sneller herkennen van woorden, waardoor de leessnelheid verbetert (Huizenga, 2016; Inspectie van het Onderwijs, 2006; Verhoeven, 2002; Verhoeven & Aarnoutse, 1999). Leerlingen leren sneller woorden te herkennen, zodat ze steeds moeilijkere teksten met begrip kunnen lezen. Dit wordt geoefend op woordniveau, maar ook op zins- en tekstniveau. Leerlingen moeten veel gelegenheid krijgen om te lezen, zodat de leessnelheid kan blijven ontwikkelen (Inspectie van het Onderwijs, 2006). Daarnaast leren ze verschillende leesstrategieën die ze kunnen gebruiken om moeilijke of nieuwe woorden te lezen. Leerlingen kunnen in deze fase steeds beter hun eigen leesgedrag controleren, waardoor ze leesfouten vaker en sneller waarnemen (Verhoeven, 2002).

Naast aandacht voor technisch lezen, is het ook belangrijk aandacht te besteden aan begrijpend lezen. Het gaat er uiteindelijk om dat leerlingen begrijpen wat ze hebben gelezen (Inspectie van het Onderwijs, 2006). Leerlingen leren dat er verschillende soorten teksten bestaan en hoe deze teksten zijn opgebouwd (Verhoeven, 2002). Ze leren de betekenis of de bedoeling van de verschillende soorten teksten te begrijpen (Verhoeven, 2002; Verhoeven & Aarnoutse, 1999). Vaak wordt begrijpend lezen geoefend in aparte lessen, maar ook tijdens de lessen voor technisch lezen kan de inhoud van de teksten besproken worden. Het tekstbegrip is voor een groot deel afhankelijk van de woordenschat. Om deze reden moet er ook aandacht zijn voor het vergroten van de woordenschat (Inspectie van het Onderwijs, 2006). Aan het eind van de basisschool kunnen leerlingen ongeveer evenveel geschreven woorden lezen en begrijpen als ze in gesproken taal kennen (Verhoeven & van Leeuwe, 2003).

## 2.2 Informatieniveaus

Tijdens het lezen wordt gebruikgemaakt van verschillende soorten informatie. Het lezen van woorden kan door herkenning van letters, maar ook door herkenning van lettergrepen of het hele woord. Het verwerken van informatie kan via verschillende niveaus verlopen, namelijk via het visuele niveau, het morfologische niveau, het semantische niveau en het syntactische niveau (Aarnoutse, 2004; Verhoeven & van Leeuwe, 2003). Op het visuele niveau spelen kenmerken van letters en woorden een rol. Het gaat hierbij met name om de vorm, positie en volgorde van losse letters. In de fase van de beginnende geletterdheid maken leerlingen met name gebruik van het visuele niveau. In de fase van de gevorderde geletterdheid spelen ook de andere niveaus van informatieverwerking een rol. Zo wordt op het morfologische niveau gekeken naar de opbouw van woorden. Op het semantische niveau speelt de betekenis van woorden een rol en op het syntactische niveau gaat het om relaties tussen woorden, woordgroepen en zinnen (Huizenga, 2016). Door gebruik te maken van zinsstructuren en de context, worden woorden en woordgroepen sneller herkend (Verhoeven & van Leeuwe, 2003). De manier waarop de niveaus van informatieverwerking een rol spelen tijdens het lezen, is uitgewerkt in verschillende modellen (zie 2.3; Huizenga, 2016).

## 2.3 Leesmodellen

In de literatuur zijn verschillende modellen voor het leesproces ontwikkeld (Verhoeven & van Leeuwe, 2003). In de jaren zeventig waren top-downmodellen populair. Deze modellen gaan ervan uit dat lezen begint op hogere informatieniveaus, waarbij gebruik wordt gemaakt van

algemene kennis en context. De informatie verloopt van hogere informatieniveaus naar lagere informatieniveaus (Eling & Bosman, 1997, p. 139). Vervolgens werden bottom-upmodellen ontwikkeld. Deze modellen suggereren dat informatie van lagere informatieniveaus naar hogere informatieniveaus verloopt, waarbij de hogere niveaus geen invloed hebben op de lagere niveaus (Eling & Bosman, 1997, p. 139). De laatste jaren lijkt er meer aandacht te komen voor interactieve modellen (Eling & Bosman, 1997, p. 139). Hieronder worden de volgende veel geciteerde modellen besproken: het duale-routemodel, het parallelle gedistribueerde verwerkingsmodel en het fonologische coherentiemodel (Verhoeven & van Leeuwe, 2003).

### 2.3.1 Duale-route model

Het duale-route model (Coltheart, in Verhoeven & van Leeuwe, 2003, p. 258) gaat ervan uit dat tijdens het lezen twee verschillende onafhankelijke routes mogelijk zijn, namelijk een indirecte route en een directe route. De directe route, ook wel lexicale of top-down route genoemd, veronderstelt dat woorden vrijwel direct worden herkend op basis van het unieke patroon van grafemen dat het woord van alle andere woorden onderscheidt (Reitsma, in Struiksma et al., 2009, p. 44; Verhoeven & van Leeuwe, 2003). De directe route gaat ervan uit dat het lezen begint op hogere informatieniveaus. Hoe afzonderlijke letters en woorden worden waargenomen, wordt sterk beïnvloed door de verwachtingen die de lezer heeft. De lezer kijkt globaal of het klopt wat hij verwachtte en maakt daarbij weinig gebruik van het visuele informatieniveau. Het visuele informatieniveau wordt pas gebruikt wanneer dat wat gelezen wordt niet klopt met de verwachtingen (Huizenga, 2016).

De indirecte route, ook wel sublexicale of bottom-up route genoemd, veronderstelt dat de betekenis van woorden wordt verkregen door grafeem-naar-foneemregels toe te passen, gevolgd door het samenvoegen van losse fonemen of klankgroepen tot een woord en betekenis-toekenning (Struiksma et al., 2009; Verhoeven & van Leeuwe, 2003). De herkenning van woorden via het omzetten van grafemen in fonemen is in het begin van de leesontwikkeling een traag verlopend proces (Aarnoutse, 2004). De stappen verlopen sneller naarmate de leeservaring toeneemt. De stappen worden altijd allemaal doorlopen en er wordt geen invloed uitgeoefend vanuit andere informatieniveaus (Huizenga, 2016). De veronderstelling is dat de directe route gebruikt wordt bij het lezen van bekende woorden en de indirecte route bij het lezen van onbekende of laagfrequente woorden en pseudoworden (Verhoeven & van Leeuwe, 2003).

### 2.3.2 Parallelle gedistribueerde verwerkingsmodel

Volgens het duale-routemodel zijn de directe en de indirecte route twee onafhankelijke routes. Het parallelle gedistribueerde verwerkingsmodel (Seidenberg & McClelland, 1989) gaat ervan uit dat de directe en de indirecte route parallel en in interactie met elkaar lopen (Seidenberg & McClelland, 1989; Verhoeven & van Leeuwe, 2003). Tijdens het lezen wordt afgewisseld tussen voorspellend en woord voor woord lezen. Bij een bekend onderwerp, wordt met name voorspellend gelezen. Maar wanneer de voorspelling niet klopt of wanneer onbekende woorden gelezen worden, wordt gebruik gemaakt van het visuele informatieniveau (Huizenga, 2016).

### 2.3.3 Fonologisch coherentiemodel

Het fonologisch coherentiemodel (Van Orden, Pennington, & Stone, 1990) gaat uit van een neurale netwerk (Bosman & Van Orden, 2003; Verhoeven & van Leeuwe, 2003). Het netwerk wordt gevormd door zenuwcellen waarin verschillende informatie ligt opgeslagen (Huizenga, 2016). In het netwerk worden drie typen knopen verondersteld die onderling en wederkerig met elkaar verbonden zijn: foneemknopen, grafeemknopen en semantische knopen (Bosman

& Van Orden 2003; Eling & Bosman, 1997, p. 146; Verhoeven & van Leeuwe, 2003). Kennis van spraakklanken ligt opgeslagen in de foneemknopen, kennis van grafemen in de grafeemknopen en betekenissen van woorden liggen opgeslagen in de semantische knopen (Huizenga, 2016).

Wanneer een woord wordt gelezen, worden eerst de grafeemknopen geactiveerd. Vervolgens worden de foneemknopen en de semantische knopen geactiveerd. Als er in plaats van een geschreven woord een gesproken woord wordt aangeboden, worden eerst de foneemknopen geactiveerd. Daarna worden de grafeemknopen en de semantische knopen geactiveerd. Wanneer aan een woord gedacht wordt, worden allereerst de semantische knopen geactiveerd. Vervolgens worden de grafeemknopen en de foneemknopen geactiveerd (Bosman & Van Orden, 2003).

De verbindingen tussen de knopen zijn verschillend in sterkte. Hoe sterker de verbindingen, des te gemakkelijker verloopt het leesproces. De verbindingen tussen foneem- en grafeemknopen zijn het sterkst, omdat in het Nederlands de relaties tussen grafemen en fonemen zeer consistent zijn (Bosman & Van Orden, 2003; Eling & Bosman, 1997, p. 147). De verbindingen tussen foneemknopen en semantische knopen en tussen grafeemknopen en semantische knopen zijn veel minder sterk. Wanneer een woord bijvoorbeeld begint met de letter 'p' of het foneem /p/ geeft dit weinig informatie over de betekenis van het woord (Bosman & Van Orden, 2003). De verbinding tussen grafeemknopen en semantische knopen is minder sterk dan tussen foneemknopen en semantische knopen. Dit kan verklaard worden vanuit het feit dat men eerder leert spreken dan lezen. Foneemknopen en semantische knopen worden al samen geactiveerd vanaf het begin van de taalontwikkeling, terwijl grafeemknopen en semantische knopen pas samen geactiveerd worden vanaf het begin van de leesinstructie (Bosman & Van Orden, 2003; Eling & Bosman, 1997, p. 147).

## 2.4 Leesstrategieën

Tijdens het lezen wordt gebruik gemaakt van verschillende leesstrategieën om een tekst te lezen. Daarbij wordt gebruik gemaakt van de verschillende informatieniveaus (zie 2.2). Per informatieniveau kunnen verschillende leesstrategieën toegepast worden. Het visuele niveau kent de volgende leesstrategieën: de elementaire leeshandeling (bijvoorbeeld *b...a...l, bal*), lezen met behulp van clusters en spellingpatronen (bijvoorbeeld *m...aan, maan*) en lezen met behulp van de visuele woordvorm, waarbij woorden in een keer worden herkend (bijvoorbeeld *Emma*). Op het morfologische niveau kan gelezen worden met behulp van een morfologische analyse (bijvoorbeeld *be-hulp-zaam*) en op het semantische en syntactische niveau kan gelezen worden met behulp van de context (bijvoorbeeld *ze bibberden van ...*) (Huizenga, 2016).

### 2.4.1 Elementaire leeshandeling

De elementaire leeshandeling betreft het spellend lezen, waarbij grafeem voor grafeem geanalyseerd wordt (TULE, 2008). Het is de eerste leesstrategie die leerlingen leren, waarbij leerlingen leren grafemen om te zetten in fonemen (Kleijnen, 1996). Tijdens de elementaire leeshandeling worden vier stappen doorlopen. Allereerst wordt de leesrichting bepaald. Vervolgens worden fonemen gekoppeld aan de grafemen en de volgorde van de fonemen moet onthouden worden. Daarna worden de fonemen samengevoegd tot één woord en als laatste wordt betekenis gegeven aan het gelezen woord. Deze leesstrategie is echter niet ideaal. Door het spellend lezen blijft het leestempo laag en het maakt het lastig om lange woorden te lezen, omdat het moeilijk is om alle letters in de goede volgorde te onthouden (Huizenga, 2016).

#### 2.4.2 Lezen met behulp van clusters en spellingpatronen

De elementaire leeshandeling wordt verkort tot lezen met behulp van clusters en spellingpatronen, waarbij in plaats van losse letters, zoals bij de elementaire leeshandeling, lettercombinaties en spellingpatronen herkend worden (Huizenga, 2016; TULE, 2008). Lettercombinaties zijn medeklinkerclusters (bijvoorbeeld *str-aat*) en vaste lettercombinaties (bijvoorbeeld *aai, ooi, oei*). Spellingpatronen betreffen combinaties van klinkers en medeklinkers (bijvoorbeeld *open, aan*). Een voordeel van deze leesstrategie is dat het geschikt is voor het lezen van niet klankzuivere woorden. Dit zijn woorden die anders geschreven worden dan ze klinken (Huizenga, 2016).

#### 2.4.3 Lezen met behulp van de visuele woordvorm

Tijdens het lezen met behulp van de visuele woordvorm worden woorden onmiddellijk herkend op basis van opvallende grafische elementen (Kleijnen, 1996). De woorden worden in hun geheel in een keer herkend als een soort plaatje. Deze leesstrategie is snel en efficiënt (Huizenga, 2016).

#### 2.4.4 Lezen met behulp van morfologische analyse

Tijdens het lezen met behulp van een morfologische analyse worden woorden gelezen door gebruik te maken van kennis over hoe woorden zijn opgebouwd, waardoor betekenisvolle delen in het woord (morfemen) herkend worden. Het gaat hierbij om betekenisvolle delen in samengestelde woorden (bijvoorbeeld *voet-bal*), voor- en achtervoegsels (bijvoorbeeld *meisje, ge-verfd*) en de woordstam bij vervoegingen (bijvoorbeeld *gefiets*) en meervoudsvormen (bijvoorbeeld *vissen, tafels*) (Huizenga, 2016; Verhoeven & van Leeuwe, 2003).

#### 2.4.5 Lezen met behulp van de context

Wanneer gelezen wordt met behulp van de context, wordt gebruik gemaakt van syntactische en semantische informatie uit de tekst (TULE, 2008). Hierbij wordt kennis over de opbouw van zinnen en betekenissen van woorden en zinnen gebruikt. Wanneer een zin begint met *de kip legt een ...*, dan wordt verwacht op basis van kennis dat de zin eindigt met *...ei* (Huizenga, 2016).

### 2.5 Deelvaardigheden van het lezen

Voor het leren lezen zijn, naast de leesstrategieën, een aantal deelvaardigheden belangrijk. Dit betreffen zowel auditieve als visuele vaardigheden. Daarnaast moeten leerlingen beschikken over voldoende taalvaardigheden om de leesinstructie te kunnen begrijpen. De leerlingen moeten kennis hebben van zowel instructiebegrippen, zoals klank, letter, woord, zin en regel, als relatiewoorden, zoals vooraan, achteraan, middelste, links, rechts, langer en korter (Huizenga, 2016).

De auditieve vaardigheden die belangrijk zijn, zijn auditieve objectivatie, auditieve discriminatie, auditieve analyse, auditieve synthese, temporeel ordenen en klankpositie bepalen. De visuele vaardigheden die belangrijk zijn, zijn visuele discriminatie, visuele analyse, visuele synthese, spatieel ordenen en letterpositie bepalen (Huizenga, 2016).

Auditieve objectivatie is het kunnen reflecteren op hoe een woord klinkt. Er kan op verschillende manieren gelet worden op hoe een woord klinkt. Als vastgesteld wordt wat de overeenkomsten en verschillen zijn tussen klanken en woorden, dan is er sprake van auditieve discriminatie (Huizenga, 2016). Tijdens auditieve discriminatie wordt elke klank van een woord onderscheiden van alle andere klanken in dat woord (Struiksma et al., 2009). Bij visuele discriminatie gaat het om het zien van overeenkomsten en verschillen tussen afbeeldingen, grafemen of woorden (Huizenga, 2016). Tijdens het lezen van een woord wordt door middel van visuele discriminatie elk grafeme onderscheiden van de andere grafemen



(Struiksmā et al., 2009). Bij het analyseren van grafemen is de richting-vormconstantie belangrijk. Dat betekent dat bij grafemen de richting en positie van belang zijn, bijvoorbeeld bij het verschil tussen de ‘p’, ‘d’ en ‘b’ (Huizenga, 2016).

Auditieve analyse is het onderscheiden van verschillende fonemen in een woord. Onder auditieve analyse valt ook het kunnen onderscheiden van afzonderlijke woorden in een zin of tekst. Visuele analyse is het onderscheiden van delen binnen een geheel. Met betrekking tot lezen betekent dit dat afzonderlijke grafemen binnen een woord onderscheiden kunnen worden (Huizenga, 2016; Struiksmā et al., 2009).

Tegenover auditieve analyse staat auditieve synthese. Dit is de vaardigheid om losse fonemen of klankgroepen samen te voegen tot een woord. Tegenover visuele analyse staat visuele synthese. Dit is het samenvoegen van delen tot een geheel. Met betrekking tot lezen gaat het om het samenvoegen van grafemen. Het hardop voorlezen van de afzonderlijke grafemen kan hulp bieden bij de visuele synthese. Visuele synthese vindt plaats bij de elementaire leeshandeling en bij het lezen van clusters en spellingpatronen. Wanneer een leerling een woord in één keer herkent, vindt er geen visuele synthese plaats (Huizenga, 2016).

Om goed auditief te kunnen synthetiseren, moet een leerling fonemen of woorden in een bepaalde volgorde onthouden. Dit valt onder temporeel ordenen. Hierbij wordt een beroep gedaan op het auditief geheugen. Spatieel ordenen is het ruimtelijk ordenen van grafemen tijdens het lezen. In het Nederlands lezen we van links naar rechts. Een leerling moet dus leren om volgens deze leesrichting te lezen. De woorden en letters moeten spatieel geordend worden van links naar rechts. Wanneer een leerling letters omdraait in een woord (bijvoorbeeld ‘lam’ uitsprekt als ‘mal’) is de vaardigheid spatieel ordenen nog onvoldoende ontwikkeld (Huizenga, 2016; Struiksmā et al., 2009).

Bij het bepalen van de klankpositie moet de leerling aangeven wat de plaats is van een klank in een auditief aangeboden woord. Bij het bepalen van de letterpositie moet de leerling in staat zijn om aan te geven op welke plaats een letter in een geschreven woord te vinden is. Het bepalen van de klank- en letterpositie is een samengestelde vaardigheid en bestaat uit een combinatie van analyse, temporeel of spatieel ordenen en kennis van begrippen, zoals vooraan en achteraan (Huizenga, 2016).

## 2.6. Leesproblemen

Leesproblemen variëren sterk en komen voor in elke fase van de leesontwikkeling (Huizenga, 2016). Leesproblemen zijn vaak al vroeg in de leesontwikkeling zichtbaar (Van der Leij, 1998). Er wordt gesproken van leesproblemen wanneer het automatiseren van de leesteknik niet goed op gang komt of achteruitgaat. Aan de hand van toetsings- en observatiegegevens kan vastgesteld worden of een leerling een leesachterstand heeft. Daartoe worden de gegevens van de leerling vergeleken met de gegevens van gemiddelde lezers. Als een leerling aan het eind van groep 3 een halfjaar achterloopt op zijn leeftijdsgenootjes, dan wordt dat als problematisch gezien (Huizenga, 2016).

Leesproblemen kunnen uitgedrukt worden in termen van leerrendement. Een gemiddelde leerling heeft een leerrendement van 100%. Een leerling die aan het eind van groep 3 het niveau beheerst van halverwege groep 4, heeft een leerrendement van 150%. Een leerling die aan het eind van groep 3 het niveau van halverwege groep 3 beheerst, heeft een leerrendement van 50%. Leerlingen met een leerrendement lager dan 50% zijn zeer zwakke lezers en leerlingen met een leerrendement tussen de 50% en 75% zijn zwakke lezers (Van der Leij, 1998).

### 2.6.1 Oorzaken van leesproblemen

De oorzaken van leesproblemen zijn zeer divers en vaak moeilijk te achterhalen (Van der Leij, 1998). Verschillende factoren kunnen leiden tot leesproblemen. Wanneer de leerling de deelvaardigheden nog onvoldoende ontwikkeld heeft of wanneer leesstrategieën niet efficiënt worden ingezet, hangen de problemen samen met leestechnische factoren (Huizenga, 2016). Ook kunnen leesproblemen samenhangen met individuele factoren, zoals verminderde intelligentie (Van der Leij, 1998), gehoorproblemen (Pronk-Boersma, 1992), een hersenbeschadiging of emotionele verwaarlozing (Huizenga, 2016). Bovendien kunnen factoren vanuit de omgeving en/of school samenhangen met leesproblemen. Te denken valt aan een andere thuistaal en de instructie van de leerkracht. Meestal zijn meerdere factoren tegelijk de oorzaak van de leesproblemen (Huizenga, 2016).

### 2.6.2 Soorten leesproblemen

Zwakke lezers worden vaak ingedeeld in verschillende categorieën. Om te bepalen in welke categorie een leerling geplaatst wordt, wordt vaak gekeken naar de aard van de leesproblemen. Zo kan een leerling problemen hebben met de elementaire leeshandeling, problemen met de overgang van spellend naar herkenkend lezen en problemen met technisch lezen (Bouwers & Van Goor, 1996).

#### *Problemen met de elementaire leeshandeling*

Aan het begin van de leesontwikkeling leren leerlingen de elementaire leeshandeling. Ze leren fonemen te koppelen aan grafemen. De volgende stappen worden hierbij doorlopen: werken volgens de leesrichting; visuele analyse in grafemen; koppelen van foneem aan grafeem; fonemen op volgorde onthouden; auditieve synthese en betekenis geven. Voor deze stappen zijn de deelvaardigheden nodig (zie 2.5). De meeste leerlingen hebben geen moeite met de elementaire leeshandeling. Wanneer de elementaire leeshandeling te lang wordt gebruikt, is dit wel een probleem (Huizenga, 2016). Ook zijn er leerlingen die wel moeite hebben met de elementaire leeshandeling. Zij laten tijdens het lezen bijvoorbeeld grafemen weg, vervangen grafemen door andere grafemen of veranderen de volgorde van grafemen. Dit kan veroorzaakt worden door het onvoldoende beheersen van een bepaalde deelvaardigheid of doordat de leerling moeite heeft met het snel achter elkaar toepassen van de deelvaardigheden (Bouwers & Van Goor, 1996; Huizenga, 2016).

Problemen met het werken volgens de leesrichting komen bij Nederlandstalige leerlingen meestal niet voor. Leerlingen die uit bepaalde landen komen waar ze gewend zijn om van rechts naar links te lezen, kunnen hier wel moeite mee hebben (Huizenga, 2016).

Wanneer leerlingen problemen hebben met de visuele analyse, kunnen zij moeite hebben met grafemen die alleen in richting (p-d-b, n-u, m-w) of volgorde (ei-ie) verschillen. Het kan ook zo zijn dat ze moeite hebben met het bepalen van de grafeemvolgorde. Dit wordt meestal veroorzaakt door problemen met de visuele waarneming (Huizenga, 2016).

Leesproblemen in het begin van groep 3 hangen vaak samen met een zwakke letterkennis. Omdat grafemen abstracte symbolen zijn, is het voor leerlingen lastig om hier fonemen aan te koppelen. Daarnaast zijn er combinaties van grafemen die samen een ander foneem vormen, waardoor de koppeling tussen foneem en grafeem bemoeilijkt wordt (Huizenga, 2016).

Problemen met het onthouden van de foneemvolgorde hangt meestal samen met een niet goed functionerend auditief geheugen. Ook problemen met de auditieve synthese hangen samen met een niet goed functionerend auditief geheugen of problemen met auditieve discriminatie (Huizenga, 2016).

Betekenis geven aan woorden wordt bemoeilijkt wanneer de leerling een woord niet kent. Vaak denken leerlingen dan dat ze het woord niet correct hebben gelezen en lezen het woord opnieuw of noemen een woord die ze wel kennen (Huizenga, 2016).

### *Problemen met de overgang van spellend naar herkendend lezen*

Wanneer leerlingen een woord een paar keer gezien hebben, kunnen ze het vaak direct herkennen via de visuele woordvorm. Het spellend lezen gaat geleidelijk over in herkendend lezen. Tijdens het spellend lezen worden alle grafemen van de te lezen woorden gespeld. Dit gaat over via fluisterend naar verinnerlijkt spellen. Vervolgens wordt de leesstrategie 'lezen met behulp van clusters en spellingpatronen' gebruikt. Uiteindelijk worden de woorden direct herkend via de visuele woordvorm (Huizenga, 2016).

Er zijn leerlingen die problemen ondervinden met de overgang van spellend lezen naar herkendend lezen. Vaak wordt er een onderscheid gemaakt tussen spellende lezers en radende lezers, maar veel leerlingen vertonen zowel kenmerken van spellend als radend lezen. In de praktijk wordt vaak een onderscheid gemaakt tussen spellende en radende lezers (Huizenga, 2016; Struiksmā et al., 2009).

### *Spellende lezers*

Aan het begin van de leesontwikkeling leren alle leerlingen spellend lezen. Dit is van belang om de koppeling tussen grafemen en fonemen eigen te kunnen maken. Wanneer leerlingen na acht maanden leesonderwijs nog steeds spellend lezen, is er sprake van een probleem. Een spellende lezer leest meer dan een derde van de woorden door hardop of innerlijk gebruik te maken van de elementaire leeshandeling of van de leesstrategie 'lezen met behulp van clusters of spellingpatronen' (Huizenga, 2016). Spellende lezers lezen langzaam, omdat ze woorden grafeem-voor-grafeem verklanken (Wentink & Verhoeven, 2008). Ze maken meestal weinig fouten. Er zijn echter ook leerlingen die spellend lezen en nog veel fouten maken. Dat komt doordat ze de deelvaardigheden van de elementaire leeshandeling nog niet goed beheersen (Huizenga, 2016).

### *Radende lezers*

Radende lezers proberen op basis van een bepaald visueel kenmerk het woord te herkennen. Hierdoor lezen zij veel woorden direct fout. De leesstrategie 'lezen met behulp van de visuele woordvorm' is bij radende lezers nog onvoldoende ontwikkeld. Daarnaast hebben sommige radende lezers moeite met de leesstrategie 'lezen met behulp van de context'. Zij merken hun fouten namelijk niet op tijdens het lezen (Huizenga, 2016). Er zijn echter ook leerlingen die beter gebruikmaken van de context, de tekst die zij eerder gelezen hebben of de illustratie. Zij lezen dan incorrecte, maar toch semantisch of syntactisch passende woorden (Wentink & Verhoeven, 2008). Bovendien slaan radende lezers vaak woorden over of verwisselen frequente woorden, zoals lidwoorden. Radende lezers hebben een normaal leestempo, maar maken veel fouten (Huizenga, 2016).

### *Problemen met technisch lezen*

Meestal ontstaan leesproblemen in de fase van beginnende geletterdheid. In de fase van gevorderde geletterdheid vallen de leesproblemen steeds meer op, omdat het verschil met leeftijdsgenoten steeds groter wordt. De leesproblemen die voorkomen in de fase van gevorderde geletterdheid zijn spellend lezen, radend lezen, langzaam lezen en eentonig lezen. Spellend lezen en radend lezen uiten zich in deze fase net zoals in de fase van beginnende geletterdheid. Er wordt gesproken van langzaam lezen wanneer leerlingen weinig fouten maken, maar het leestempo achterblijft op dat van leeftijdsgenoten. Zij worden ook wel

vertraagde lezers genoemd (Sas & Wierenga, 1998). Er wordt gesproken van eentonig lezen wanneer leerlingen moeite hebben met de voordrachtsaspecten van lezen. Deze leerlingen hebben moeite met het uiten van expressie tijdens het lezen. Vaak gaat dit samen met langzaam lezen of spellend en radend lezen (Huizenga, 2016).

Een aantal factoren versterken de leesproblemen in de fase van gevorderde geletterdheid. Zo moeten leerlingen steeds langere meerlettergrepige woorden lezen. Herkenning van meerlettergrepige woorden verloopt snel wanneer de leesstrategie 'lezen met behulp van morfologische analyse' goed ontwikkeld is. Leerlingen die al moeite hebben met het lezen van eenlettergrepige woorden, komen bij meerlettergrepige woorden nog meer in de problemen. Daarnaast kan de moeilijkheidsgraad van teksten een belemmerende factor zijn. De moeilijkheidsgraad wordt bepaald door bijvoorbeeld de grootte van de grafemen, de lengte van de tekst en zinnen, de structuur van de tekst, onbekende woorden en figuurlijk taalgebruik (Huizenga, 2016). Het komt vaak voor dat de motivatie om te lezen afneemt tijdens de fase van gevorderde geletterdheid, doordat leerlingen negatieve ervaringen opdoen tijdens het lezen (Vernooy, 1993).

## 2.7 Evalueren van de leesontwikkeling

De ene leerling kan al vroeg goed lezen, terwijl bij de andere leerling het lezen maar niet op gang komt. Het onthouden van letters lukt niet, het leestempo blijft laag of de leerling blijft hangen in de elementaire leeshandeling. Dit wordt meestal veroorzaakt doordat bepaalde deelvaardigheden nog niet verworven zijn. Het vroegtijdig en regelmatig evalueren van de leesontwikkeling van leerlingen geeft inzicht in eventuele problemen (Huizenga, 2016). Vroegtijdige onderkenning en behandeling van problemen die tot leesproblemen kunnen leiden, is effectiever dan laat signaleren en repareren (Vernooy, in Blonk et al., 2006, p. 52). Voor risicolezers kan het onderwijs aangepast worden of vroegtijdig hulp geboden worden (Huizenga, 2016). Daarnaast kan geprobeerd worden iets te doen aan de omstandigheden die de ontwikkeling van de leerling belemmeren (Braams, 2007). Op deze manier kunnen eventuele leesproblemen in de toekomst worden voorkomen (Huizenga, 2016).

Door groepen leerlingen te splitsen op basis van bepaalde kenmerken, kan aan leerlingen met leesproblemen meer hulp geboden worden en leerlingen die al goed kunnen lezen, kunnen meer uitgedaagd worden (Huizenga, 2016). Dit wordt differentiatie genoemd. Differentiëren kan door zowel externe als interne differentiatie. Externe differentiatie houdt in dat leerlingen met verschillende vaardigheden onderverdeeld worden over verschillende klassen, scholen en niveaus (Van Casteren et al., 2017). Bij interne differentiatie wordt rekening gehouden met verschillen tussen leerlingen om op die manier een maximaal leerresultaat te kunnen behalen. Er wordt dan rekening gehouden met verschillen in niveau, tempo en interesse (Couberts et al., in Van Casteren et al., 2017, p. 15). Ook kan gekeken worden naar verschillende kenmerken, zoals leeftijd, geslacht, sociaaleconomische status, intelligentie, persoonlijkheid, leerstijl, motivatie of prestaties (Bosker, in Van Casteren et al., 2017, p. 18). Voordat gedifferentieerd kan worden, moet er informatie over de leerlingen verzameld worden. Deze informatie wordt verkregen door toetsing en observatie. De verzamelde informatie wordt vervolgens gebruikt om de lessen aan te passen aan verschillen tussen leerlingen in dezelfde klas. Dit wordt adaptief onderwijs of onderwijs op maat genoemd. Onderwijs op maat kan de leeropbrengsten van leerlingen verbeteren (Van Casteren et al., 2017).

Leerkrachten spelen een belangrijke rol bij het signaleren van leesproblemen (Braams, 2007). Dit doen zij door de leesontwikkeling op verschillende manieren te evalueren. Tijdens evaluaties wordt informatie over de leesontwikkeling verzameld en geïnterpreteerd, zodat op basis hiervan beslissingen genomen kunnen worden (Huizenga, 2016). Naast observaties die leerkrachten doen tijdens het leesproces, nemen zij toetsen af om na te gaan hoe de

leesontwikkeling van hun leerlingen verloopt (Braams, 2007). Tijdens een observatie worden gedragingen of vaardigheden van een leerling geobserveerd die niet goed meetbaar zijn, door te kijken en te luisteren. Voor een observatie zijn geen duidelijke normen opgesteld. Bij toetsing wordt er een specifieke taak afgenomen, bijvoorbeeld het lezen van een rij woorden of een tekst (Huizenga, 2016). Het leesniveau van de leerling wordt bepaald door de prestaties te vergelijken met de prestaties van een normgroep of met een bepaald criterium (Struiksmā et al., 2009). Bij een gestandaardiseerde toets zijn de normen bepaald op basis van de resultaten van een groep leerlingen die representatief is voor alle Nederlandse leerlingen. Met een gestandaardiseerde toets wordt een leerling vergeleken met leeftijdsgenootjes en op basis hiervan wordt bepaald of de leerling een achterstand heeft (Huizenga, 2016). Toetsen die afgenomen worden om het leesproces te evalueren, hebben betrekking op taalvaardigheid, deelvaardigheden en leesvaardigheid (Huizenga, 2016).

### 2.7.1 Toetsen van de taalvaardigheid

Een toets die gebruikt kan worden om de taalvaardigheid bij kleuters te meten is ‘Cito Taal voor kleuters’ (Lansink, 2009). Bij anderstalige leerlingen kan de ‘Taaltoets Alle Kinderen’ (Verhoeven & Vermeer, 2001) afgenomen worden. Vanaf groep 3 kunnen toetsen van het Cito leerlingvolgsysteem afgenomen worden, zoals de toets Woordenschat (Cito, z.d.). Bovendien is het nuttig de taalvaardigheid te observeren tijdens de les. Daarbij kan gelet worden op het begrijpen van taal, het verwoorden van ervaringen en gedachten, deelname aan gesprekken en belangstelling voor geschreven taal (Huizenga, 2016).

### 2.7.2 Toetsen van deelvaardigheden

Deelvaardigheden van het lezen (zie 2.5) kunnen getoetst worden met het deelvaardighedenonderzoek van Sas en Wieringa (Sas & Wieringa, 1998). Ook met het leerlingvolgsysteem van Cito kunnen sommige deelvaardigheden getoetst worden (Huizenga, 2016). In het boek Diagnostiek van technisch lezen en aanvankelijk spellen (Struiksmā et al., 2009) staat voor de verschillende deelvaardigheden beschreven welke toets afgenomen kan worden (Struiksmā et al., 2009).

### 2.7.3 Toetsen van de leesvaardigheid

Voor het toetsen van de leesvaardigheid zijn twee toetsen essentieel (Struiksmā et al., 2009): de Drie-Minuten-Toets (DMT; Jongen & Krom, 2009b) en de AVI-toetskaarten (Analyse Van Individualiseringsvormen; Jongen & Krom, 2009a). Bij de DMT wordt de leesvaardigheid getoetst via het lezen van losse woorden. De DMT bestaat uit drie kaarten die verschillen in moeilijkheidsgraad (Huizenga, 2016; Struiksmā et al., 2009). Aan de hand van deze kaarten wordt nagegaan hoe snel en accuraat leerlingen woorden kunnen lezen. De leerling krijgt per kaart één minuut de tijd om zoveel mogelijk woorden correct hardop te lezen (Struiksmā et al., 2009; Verhoeven & van Leeuwe, 2003). Overgeslagen, verkeerd gelezen of spellend gelezen woorden worden genoteerd (Huizenga, 2016). Op de eerste leeskaart staan 150 eenlettergrepige klankzuivere MKM (medeklinker-klinker-medeklinker) woorden. De tweede leeskaart bestaat uit 150 eenlettergrepige MMKM, MKMM, MMKMM, MMMKM en MKMMM-woorden en de derde leeskaart bestaat uit 120 woorden van twee lettergrepen, drie lettergrepen en vier lettergrepen. Op basis van het aantal correct gelezen woorden, wordt het leesniveau bepaald (Verhoeven & van Leeuwe, 2003).

Toetsing via het lezen van teksten kan met de AVI-toetskaarten. De AVI-toets bestaat uit elf teksten die toenemen in moeilijkheidsgraad. Het is de bedoeling dat de leerling de tekst zo snel mogelijk en met zo min mogelijk fouten leest (Struiksmā et al., 2009). Ook bij de AVI-toetsen worden overgeslagen, verkeerd of spellend gelezen woorden genoteerd, evenals woorden die zijn herhaald, tussengevoegd of voorgezegd door de onderzoeker. Het leesniveau

wordt bepaald op basis van de leestijd en het aantal fouten. Sinds 2009 is de aanduiding van niveaus gerelateerd aan de groepsindeling van de basisschool. Het AVI-M4 niveau betekent het niveau dat een gemiddelde leerling midden groep 4 beheerst. AVI-E5 is het niveau dat een gemiddelde leerling aan het eind van groep 5 beheerst (Huizenga, 2016). Door middel van tabel 1 (Balans, 2008) zijn de oude AVI-niveaus om te rekenen naar de nieuwe AVI-niveaus. De leerling wordt op basis van vooraf opgestelde normen ingedeeld in beheersingsniveau, instructieniveau en frustratieniveau (Struiksma et al., 2009). Op het beheersingsniveau leest de leerling vrijwel foutloos met een behoorlijk tempo. Op dit niveau wordt minder dan  $\pm 5\%$  van de woorden fout gelezen. Het leestempo is afhankelijk van het niveau. Op niveau AVI-E3 worden ongeveer 40 woorden per minuut gelezen en op niveau AVI-E6 worden meer dan 100 woorden per minuut gelezen. Het frustratieniveau is het niveau waarop de leerling een laag leestempo heeft en veel fouten maakt. Globaal leest een leerling op het frustratieniveau wanneer meer dan  $\pm 5\%$  van de woorden fout wordt gelezen. Op het instructieniveau kost het lezen nog moeite, maar door voldoende oefening en begeleiding kan de leerling dit niveau toch aan (Struiksma et al., 2009).

Tabel 1. Omreken tabel van oude AVI-niveaus naar nieuwe AVI-niveaus

Oud	Nieuw											
1	<b>Start</b>	<b>M3</b>										
2		<b>M3</b>	<b>E3</b>									
3			<b>E3</b>	M4								
4				<b>M4</b>	E4							
5				M4	<b>E4</b>	M5						
6					E4	<b>M5</b>	E5					
7							<b>E5</b>	M6				
8							E5	<b>M6</b>	<b>E6</b>	M7		
9								M6	<b>E6</b>	<b>M7</b>		
9+										M7	<b>E7</b>	<b>Plus</b>

Opmerking. Herdrukt van ‘‘Omreken tabel van AVI-oud naar AVI-nieuw’’ door Balans, 2008. Geraadpleegd van <http://www.de-edelsteen.nl/Portals/71/docs/Diversen/AVI%20omreken tabel.pdf>.

#### 2.7.4 Analyse van toetsingsgegevens

Toetsingsgegevens kunnen zowel kwantitatief als kwalitatief geanalyseerd worden. Bij een kwantitatieve analyse wordt gekeken naar ruwe scores, zoals het aantal goed dan wel fout gelezen woorden, terwijl bij een kwalitatieve analyse wordt gekeken naar de manier waarop gelezen wordt. Bij een kwalitatieve analyse kan bijvoorbeeld gelet worden op de leesstrategieën die een leerling gebruikt (Huizenga, 2016). De onderzoeker geeft de leerling de opdracht een tekst hardop voor te lezen en noteert vervolgens de uitingen die afwijken van de tekst (Leu, 1982). Afwijkende uitingen worden letterlijk opgeschreven zoals de leerling het heeft voorgelezen. Vervolgens kunnen de uitingen geanalyseerd worden. Dit kan bijvoorbeeld door te noteren welke woorden spellend of vloeiend werden gelezen, welke woorden werden gecorrigeerd en welke woorden op basis van de context werden gelezen. Op basis van de analyse wordt bepaald welke (deel)vaardigheden extra geoefend moeten worden. Daarnaast wordt, wanneer regelmatig een analyse gemaakt wordt, zicht verkregen op de leesontwikkeling van de leerling, de vooruitgang en de eventuele blijvende problemen (Wentink & Verhoeven, 2008).

Analyses van het leesgedrag staan al lang in de belangstelling (Leu, 1982). Leu (1982) onderscheidt twee perioden die te onderscheiden zijn op basis van de theoretische achtergrond en de categorieën die gebruikt worden om het leesgedrag te beschrijven. De eerste periode eindigde eind jaren zestig en de tweede periode loopt vanaf eind jaren zestig tot heden. Voor 1968 misten de analyses van leesgedrag een duidelijk theoretisch kader. De fouten die gedefinieerd werden, gaven weinig inzicht in de onderliggende cognitieve en linguïstische

processen. Onderzoekers keken naar algemene type leesfouten, zoals vervangingen, weglatingen, toevoegingen, herhalingen en de frequentie van deze fouten. Eind jaren zestig vond er een verandering plaats in het analyseren van leesgedrag. Onderzoekers gingen gebruikmaken van categorieën die in relatie staan met de informatieniveaus (zie 2.2), zoals semantisch of syntactisch acceptabele en niet acceptabele fouten en fouten die grafisch gelijk of niet gelijk zijn aan het doelwoord in de tekst (Leu, 1982).

### 2.7.5 Eerder onderzoek naar leesfouten en leesstrategieën

Cleuren (2009) onderzocht in Vlaanderen de aard en prevalentie van regelmatig voorkomende leesfouten. Dit onderzocht zij bij 400 Nederlandstalige basisschoolleerlingen met en zonder leesproblemen. Zij deed dit op basis van opgenomen en geannoteerde voorgelezen spraak, waarbij leesfouten en leesstrategieën werden gelabeld. Uit de resultaten bleek dat jongens teksten op hun beheersingsniveau significant minder accuraat lazen dan meisjes. Daarnaast was er een significant verband tussen het AVI-niveau en leesaccuratesse bij het lezen van een tekst op beheersingsniveau. De resultaten lieten niet zien dat de leerlingen met een lager AVI-niveau meer leesfouten maakten dan leerlingen met een hoger AVI-niveau. Dit was zoals verwacht, omdat de moeilijkheid van de teksten paste bij het AVI-niveau van de leerling. Het verband tussen klas en leesaccuratesse bij het lezen van een tekst op beheersingsniveau was niet significant.

De tien meest frequente leesfouten die Cleuren (2009) vond, waren:

- leesfouten die volgens haar leesfoutenclassificatie niet geassocieerd konden worden;
- vervanging van een medeklinker door een willekeurige andere medeklinker (bijv. knop/kjop);
- toevoeging van een willekeurige medeklinker (bijv. poes/poest);
- vervanging van korte klinkers door lange klinkers of omgekeerd (bijv. man/maan);
- verandering in de volgorde van aangrenzende letters (bijv. droon/doorn);
- vervanging van klinkers (bijv. bal/bel);
- een bestaand woord werd vervangen door een pseudowoord (bijv. straat/skraat);
- (foute of correcte) gedeeltelijke (bestaande uit (een combinatie van) enkele letters of lettergrepen) decodering die niet werd gevolgd door een poging tot gesynthetiseerd lezen (bijv. v...oe...t...bal);
- een bestaand woord werd vervangen door een ander bestaand woord dat orthografisch gelijk is (bijv. staart/straat);
- een woord werd weggelaten.

De resultaten toonden aan dat de leerlingen met name lexicale strategieën gebruikten om woorden in een tekst te lezen. Wanneer woorden langer werden, werden meer sublexicale strategieën gebruikt. Daarnaast was er een leeftijdsafhankelijke verschuiving in het gebruik van leesstrategieën. Hoe ouder de leerlingen, des te meer gebruik werd gemaakt van lexicale strategieën. Dit bleek uit het resultaat dat oudere leerlingen ten opzichte van jongere leerlingen vaker woorden vervingen door een bestaand woord en minder vaak door een niet bestaand woord. Daarnaast lieten oudere leerlingen minder woorden weg. Ook gold hoe lager het decoderingsniveau (zwak, gemiddeld, sterk), des te meer de leerlingen gebruikmaakten van lexicale strategieën. Leerlingen met een lager decoderingsniveau vervingen ten opzichte van leerlingen met een hoger leesniveau vaker woorden door een bestaand woord en minder vaak door een niet bestaand woord. Daarnaast lieten leerlingen met een lager decoderingsniveau vaker woorden weg dan leerlingen met een hoger decoderingsniveau. Deze resultaten suggereren dat hoe lager het decoderingsniveau tijdens het lezen van de tekst, des te meer de leerling probeert het woord te raden door een lexicale strategie toe te passen in plaats van grafeem-naar-foneem regels te gebruiken bij het lezen van een onbekend woord. Ook onderzocht Cleuren (2009) het verschil tussen jongens en meisjes. Ze vond dat meisjes ten

opzichte van jongens iets minder woorden vervangen door niet bestaande woorden. De proportie woorden die werd vervangen door bestaande woorden was voor meisjes en jongens ongeveer even groot. Dit suggereert dat meisjes ten opzichte van jongens minder vaak gebruikmaken van sublexicale strategieën.

Goikoetxea (2006) deed onderzoek naar leesfouten die gemaakt werden door beginnende lezers. Zij legde de focus van haar onderzoek op zowel contextafhankelijke (c, g, r) als visueel vergelijkbare medeklinkers (b, d, g, p, q). Aan het onderzoek namen negentig Spaanssprekende leerlingen deel, waarvan 45 uit groep 3 en 45 uit groep 4. De leerlingen moesten zowel bestaande als niet bestaande woorden voorlezen. Uit het onderzoek bleek dat er meer fouten werden gemaakt in contextafhankelijke medeklinkers dan in visueel vergelijkbare dan wel de overige medeklinkers. De oudere leerlingen maakten minder fouten in medeklinkers die visueel vergelijkbaar waren dan de jongere leerlingen. Dit impliceert dat bij het leren van de koppeling tussen grafemen en fonemen, het koppelen van fonemen aan contextafhankelijke medeklinkers meer problemen oplevert dan het koppelen van fonemen aan visueel vergelijkbare medeklinkers.

Laing (2002) onderzocht welke leesfouten zich normaal ontwikkelende leerlingen en leerlingen met beneden gemiddelde taal- en leesvaardigheden maakten. Ze onderzocht of deze twee groepen leerlingen vergelijkbare fouten maakten. Daarnaast onderzocht ze of typisch ontwikkelende leerlingen meer fouten maakten door woorden te vervangen door grafemisch en/of fonemisch vergelijkbare woorden en niet bestaande woorden dan leerlingen met beneden gemiddelde taal- en leesvaardigheden. Aan het onderzoek namen 22 Amerikaanse leerlingen met een leeftijd tussen 8;5 en 9;6 jaar deel, waarvan 11 typisch ontwikkelende leerlingen en 11 leerlingen met beneden gemiddelde taal- en leesvaardigheden. De leerlingen lazen een tekst hardop voor, die vervolgens geannoteerd werd door de onderzoeker. De resultaten toonden aan dat typisch ontwikkelende leerlingen ten opzichte van leerlingen met beneden gemiddelde taal- en leesvaardigheden meer fouten maakten waarbij de betekenis van de tekst behouden bleef. Beide groepen maakten evenveel grafemisch en/of fonemisch vergelijkbare fouten of vervangen even vaak woorden door niet bestaande woorden.

Rutter en collega's (2004) analyseerden vier onafhankelijke studies die onderzoek deden naar sekseverschillen bij leesproblemen. In de eerste studie werden in Nieuw-Zeeland de leesprestaties van 989 kinderen gevolgd. De leesprestaties werden onderzocht op de leeftijd van 7, 9 en 11 jaar. De tweede studie onderzocht ook in Nieuw-Zeeland de leesprestaties van 895 kinderen tussen de 8 en 10 jaar. De derde studie was een onderzoek uit Groot-Brittannië. Hier werden bij 5752 kinderen tussen de 9 en 15 jaar de leesprestaties onderzocht. De laatste studie onderzocht bij 2163 zevenjarige tweelingkinderen uit Engeland en Wales de leesprestaties. De resultaten van alle vier de studies lieten zien dat het percentage leesproblemen bij jongens significant hoger was dan bij meisjes. Dit suggereert dat leesproblemen vaker voorkomen bij jongens dan bij meisjes.

Alhoewel al veel onderzoek gedaan is naar leesfouten en leesstrategieën, zijn de resultaten vanuit andere talen niet direct te generaliseren naar het Nederlandse leesproces. Verschillende studies hebben aangetoond dat verschillen in leesstrategieën een gevolg zijn van verschillen in orthografie (Protopapas & Skaloumbakas, 2007). Orthografie is de manier waarop een taal schriftelijk wordt weergegeven. De orthografie van een taal wordt bepaald door de mate waarin de relatie tussen grafemen en fonemen consistent is (Lieberman et al., in Guron & Lundberg, 2004, p. 44). Het Nederlands is een taal met een vrijwel transparante orthografie, waarbij woorden vaak worden geschreven zoals ze klinken en vaak worden uitgesproken zoals ze worden geschreven (Cleuren, 2009).

Daarnaast zijn de resultaten vanuit andere landen niet direct te generaliseren naar het leesproces in Nederland, omdat andere landen andere prestaties behalen op het gebied van lezen. Zo scoort Nederland hoger dan België (Vlaanderen) op het gebied van leesvaardigheid



en heeft Nederland een kleinere spreiding tussen zwakke en sterke lezers dan Vlaanderen. Bovendien is er een verschil in leesonderwijs tussen verschillende landen. Zo krijgt in Nederland slechts 7% van de leerlingen voornamelijk klassikaal leesonderwijs, terwijl dit percentage nog lager ligt in Vlaanderen (Gubbels, Netten & Verhoeven, 2017).

## 2.8 Hypotheses

Op basis van de theoretische achtergrond en eerder onderzoek naar leesfouten en leesstrategieën werden de volgende hypothesen opgesteld:

Onderzoeksvraag	Hypothesen
<i>1. Is er een verband tussen leesaccuratesse en de variabelen leeftijd, geslacht en AVI-niveau?</i>	Leeftijd, geslacht en AVI-niveau hebben geen effect op leesaccuratesse.
<i>2. Wat zijn de meest frequente leesfouten die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar maken tijdens het hardop voorlezen van een Nederlandse tekst?</i>	Vervangingen van medeklinkers en klinkers, vervangingen in de volgorde van aangrenzende letters, vervangingen van een bestaand woord door een pseudowoord of een ander bestaand woord dat orthografisch gelijk is, toevoegingen van een willekeurige medeklinker, decodering die niet werd gevolgd door een poging tot gesynthetiseerd lezen en weglatingen van woorden komen vaker voor dan de andere geannoteerde leesfouten (zie bijlage 1).
<i>3. Is er een verband tussen leesfouten en de variabelen leeftijd, geslacht en AVI-niveau?</i>	Oudere leerlingen en leerlingen met een hoger AVI-niveau vervangen proportioneel gezien minder woorden door orthografisch gelijke woorden of pseudowoorden en meer door semantisch gelijke of semantisch passende woorden ten opzichte van jongere leerlingen en leerlingen met een lager AVI-niveau. Daarnaast vervangen meisjes proportioneel gezien minder vaak woorden door pseudowoorden ten opzichte van jongens.
<i>4. Wat zijn de meest frequente leesstrategieën die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar toepassen tijdens het hardop voorlezen van een Nederlandse tekst?</i>	De meest frequente leesstrategieën zijn lexicale leesstrategieën. Minder frequent zijn sublexicale leesstrategieën.
<i>5. Is er een verband tussen leesstrategieën en de variabelen leeftijd, geslacht en AVI-niveau?</i>	Oudere leerlingen en leerlingen met een hoger AVI-niveau maken proportioneel gezien meer gebruik van lexicale leesstrategieën ten opzichte van jongere leerlingen en leerlingen met een lager AVI-niveau. Daarnaast maken meisjes proportioneel gezien meer gebruik van lexicale leesstrategieën.

### 3. Methode

#### 3.1 Onderzoeksontwerp

Het onderzoek bestond uit een kwantitatief cross-sectioneel onderzoek met bestaande data. De onafhankelijke variabelen waren leeftijd (6 tot en met 13 jaar), geslacht (jongen of meisje) en AVI-niveau (1 tot en met 9). De afhankelijke variabelen waren de frequenties van leesstrategieën, de frequenties van leesfouten, de proporties van leesstrategieën, de proporties van leesfouten en leesaccuratesse.

#### 3.2 Materiaal

##### 3.2.1 Spraakmateriaal JASMIN-CGN Corpus

Voor het onderzoek werd gebruik gemaakt van het JASMIN-CGN (Jongeren, Anderstaligen en Senioren in Mens-machine Interactie – Corpus Gesproken Nederlands) Corpus (Taalunie, 2008). Dit corpus bevat spraak van groepen sprekers die niet zijn vertegenwoordigd in het CGN, zoals kinderen, anderstaligen en senioren. Per persoon bestaat ongeveer 50% van het materiaal uit voorgelezen spraakmateriaal. De andere 50% bestaat uit onvoorbereide spraak, geproduceerd in mens-machine dialogen, waarbij sprekers in gesprek gaan met een computer (Cucchiari et al., 2008). Dit is in het corpus verdeeld over twee componenten: comp-p (dialogen) en comp-q (voorgelezen spraak). De sprekers zijn ingedeeld in verschillende groepen: groep 1, kinderen tussen 7 en 11 jaar; groep 2, kinderen tussen 12 en 16 jaar; groep 3, anderstalige kinderen; groep 4, anderstalige volwassenen en groep 5, senioren ouder dan 65 jaar. Het spraakmateriaal is verdeeld over zes DVD's (Cucchiari et al., 2008). Voor het huidige onderzoek werd het voorgelezen spraakmateriaal (comp-q) van Nederlandstalige basisschoolleerlingen gebruikt (groep 1). Dit is te vinden op DVD01 en DVD02. Toegang tot het JASMIN-CGN Corpus werd verkregen via het CLS (Centre for Language Studies) van de Radboud Universiteit Nijmegen.

Het materiaal bevat voorgelezen spraak, waarbij leerlingen een tekst voorlezen uit boeken van de leesmethode *Veilig Leren Lezen*<sup>1</sup> (Mommers, Verhoeven, & Van der Linden, 1990). De leerlingen lazen een tekst voor passend bij hun beheersingsniveau, variërend van AVI-1 tot en met AVI-9 (Cucchiari et al., 2008). De leerlingen kregen de instructie om de teksten te lezen die op een scherm verschenen (Cucchiari et al., 2008).

De spraakopnames werden gemaakt in stille ruimtes op scholen, alhoewel achtergrondgeluiden en galm niet altijd voorkomen konden worden. De opnameapparatuur bestond uit een Sennheiser 835 cardioïde microfoon, versterker, geluidskaart en opnamesoftware. De opnames hebben een sample frequentie van 16 kHz en een resolutie van 16 bits PCM (pulsocodemodulatie). De orthografische transcripties werden handmatig gemaakt en de fonetische transcripties werden automatisch gegenereerd (Cucchiari et al., 2008). In het corpus zijn per spreker twee bestanden beschikbaar, namelijk een .awd-bestand en een .wav-bestand. Het .awd-bestand bevat de orthografische transcriptie (tier 1), een automatisch gegenereerde woordoplijning met fonemische transcriptie (tier 2) en een automatisch gegenereerde foneemoplijning met fonemische annotaties (tier 3) (Cucchiari et al., 2008). Het .wav-bestand bevat de spraakopname.

##### 3.2.2 Annotatieprotocol

Voor het annoteren van leesfouten en leesstrategieën in het JASMIN-CGN Corpus werd een vooraf opgesteld protocol gebruikt (zie bijlage 1). Dit protocol werd in samenwerking met E.L.M. de Reus (masterstudent General Linguistics) opgesteld op basis van het Annotatieprotocol voor CHOREC (Cleuren, Duchateau, & Sips, 2008). Het Annotatieprotocol

---

<sup>1</sup> De teksten mogen vanwege auteursrechten niet gepubliceerd worden.

voor CHOREC werd vertaald naar het Nederlands en op basis van literatuur en relevantie aangevuld en aangepast.

### 3.3 Proefpersonen

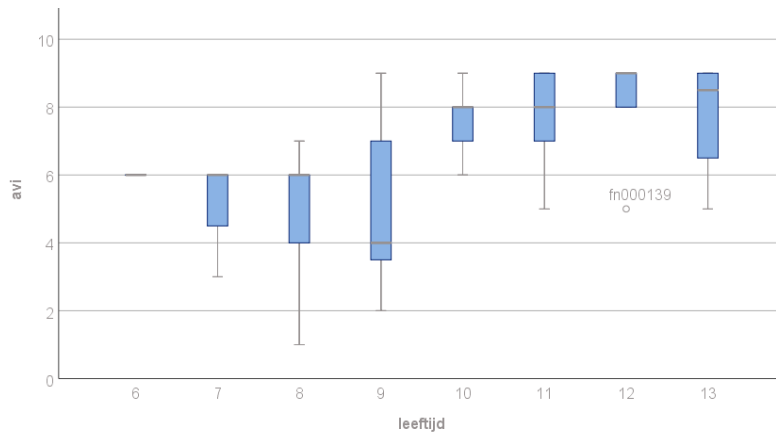
Voor het huidige onderzoek werden bestaande spraakopnames gebruikt van 71 basisschoolleerlingen ( $M = 9.8$  jaar,  $SD = 1.6$  jaar, bereik = 6-13) uit groep 4 tot en met groep 8, waarvan 35 meisjes en 36 jongens. Deze leerlingen hadden Nederlands als moedertaal. Van de 71 leerlingen kwamen er 29 uit regio Rivierenland (inclusief Arnhem en Nijmegen), 20 uit Noord-Brabant en 22 uit Overijssel. De leerlingen werden geworven op basisscholen (Cucchiarini et al., 2008). De AVI-niveaus van de leerlingen varieerden tussen AVI-1 en AVI-9 (*mediaan* = 7, *modus* = 8). In tabel 2 wordt een overzicht gegeven van de AVI-niveaus per leeftijdsgroep. Sommige leerlingen beheersten een lager AVI-niveau dan verwacht werd bij hun leeftijd. Daarnaast waren er leerlingen die een hoger AVI-niveau beheersten dan verwacht werd bij hun leeftijd. De spreiding van AVI-niveaus per leeftijdsgroep was groot (zie figuur 1). Er was een matig positief verband tussen leeftijd en AVI-niveau,  $r = .55$ ,  $p < .001$ .

Van drie leerlingen was de leeftijd niet bekend, namelijk van de leerlingen met proefpersoonnummer 491, 540 en 543. Deze leerlingen werden niet meegenomen in de analyses die betrekking hadden op leeftijd. Ze werden wel meegenomen in de analyses die betrekking hadden op geslacht en AVI-niveau.

Leerlingen met een accuratessescore lager dan .95 werden niet meegenomen in het onderzoek. Deze leerlingen lazen meer dan 5% van de woorden fout, waardoor niet met zekerheid gezegd kan worden dat ze de tekst op beheersingsniveau gelezen hebben. Op het beheersingsniveau wordt namelijk minder dan  $\pm 5\%$  van de woorden fout gelezen (zie 2.7.3). Leerlingen met een beheersingsniveau lager dan het streefniveau, werden wel meegenomen in de analyses. Bij deze leerlingen kan gedacht worden aan leesproblemen zoals dyslexie, maar er is geen bewijs dat leerlingen met dyslexie andere leesfouten maken dan leerlingen zonder dyslexie (Rozeman; Baddeley, Logie, & Ellis, in Braams, 2007).

Tabel 2. Overzicht van de AVI-niveaus per leeftijdsgroep

		Leeftijd									
		6	7	8	9	10	11	12	13	Ontbrekend	Totaal
AVI-niveau	1	0	0	1	0	0	0	0	0	0	1
	2	0	0	1	1	0	0	0	0	0	2
	3	0	1	0	2	0	0	0	0	0	3
	4	0	0	3	3	0	0	0	0	0	6
	5	0	0	1	0	0	2	1	1	0	5
	6	1	2	6	0	3	0	0	0	0	12
	7	0	0	2	3	2	3	0	0	1	11
	8	0	0	0	1	6	6	1	1	1	16
	9	0	0	0	1	3	5	3	2	1	15
Totaal		1	3	14	11	14	16	5	4	3	71



Figuur 1. Boxplot AVI-niveau afgezet tegen leeftijd

### 3.4 Procedure

Allereerst werd het spraakmateriaal vanuit het JASMIN-CGN Corpus gedownload. Vervolgens werd in het computerprogramma PRAAT (Boersma & Weenink, Versie 6.0.52) het PRAAT-script (zie bijlage 2), opgesteld door E.P. Sanders (onderzoeker bij Centre for Language Studies en Centre for Language and Speech Technology van de Radboud Universiteit Nijmegen), geopend en uitgevoerd. Dit PRAAT-script genereerde automatisch drie extra tiers. Tier 4 was een lege tier waar plek was voor de originele tekst. In tier 5 werden de leesstrategieën geannoteerd en in tier 6 de leesfouten. Tier 5 en 6 bevatten dezelfde woordgrenzen als tier 2 en 3, zodat per woord geannoteerd kon worden.

Het spraakmateriaal van de zes-, zeven-, en achtjarige leerlingen werd in samenwerking met E.L.M. de Reus geannoteerd. Het overige spraakmateriaal werd zelfstandig geannoteerd. In tier 4 werd de originele tekst per zin ingevoegd, vanaf de hoofdletter tot aan de punt. De zinnen werden van elkaar onderscheiden door segmentgrenzen te plaatsen (*Boundary > Add on tier 4*). Daarna werd geluisterd en gecontroleerd of dat wat hardop werd gelezen, overeenkwam met de tekst. In tier 5 en 6 werden de labels voor leesstrategieën en leesfouten uit het annotatieprotocol (zie bijlage 1) ingevoegd in overeenstemming met het protocol. Vervolgens werden de annotaties opgeslagen als textgrid-bestand.

Uit alle textgrid-bestanden werd zowel tier 5 als tier 6 gehaald (*Extract one tier...*). Overzichten van de annotaties werden opgevraagd in tabellen en vervolgens opgeslagen (*Tabulate > List... > OK > File > Save as...*). De tabellen werden geopend in Excel en de frequenties van alle labels werden geteld en ingevoerd in Excel. Daarnaast werd er een tweede dataset gemaakt voor de laatste leesopingen. Ook van de laatste leesopingen werden de frequenties van de labels geteld en ingevoerd in een nieuwe dataset in Excel.

### 3.5 Data-analyse

De statistische analyses werden uitgevoerd met het statistische programma Statistical Package for the Social Sciences (IBM SPSS, Versie 21). De variabelen die in SPSS werden ingevoerd, waren: proefpersoonnummer, leeftijd, geslacht, AVI-niveau, leesaccuratesse, frequentie leesstrategieën, frequentie leesfouten, proportie leesstrategieën en proportie leesfouten.

Leesaccuratesse, proportie leesstrategieën en proportie leesfouten werden als volgt berekend:

- Leesaccuratesse = aantal goed gelezen woorden in laatste leesopingen gedeeld door totaal aantal woorden in de tekst (zoals in Cleuren, 2009).
- Proportie leesfouten (alle leesopingen) = aantal keer dat een bepaalde leesfout geannoteerd is, gedeeld door totale frequentie van alle leesfouten.

- Proportie leesfouten (laatste lees pogingen) = aantal keer dat een bepaalde leesfout in de laatste lees poging geannoteerd is, gedeeld door totale frequentie van alle leesfouten in de laatste lees poging.
- Proportie leesstrategieën (alle lees pogingen) = aantal keer dat een bepaalde leesstrategie geannoteerd is, gedeeld door totale frequentie van alle leesstrategieën (exclusief *g*).
- Proportie leesstrategieën (laatste lees pogingen) = aantal keer dat een bepaalde leesstrategie in de laatste lees poging geannoteerd is, gedeeld door totale frequentie van *f*, *0*, *W*, *alg*, *ssg* in de laatste lees poging.

### 3.5.1 Verband tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau

Om antwoord te kunnen geven op de onderzoeksvraag ‘*Is er een verband tussen leesaccuratesse en de variabelen leeftijd, geslacht en AVI-niveau?*’, werden Pearson’s correlatiecoëfficiënten berekend. Pearson’s correlatiecoëfficiënt is een test die gebruikt wordt om te onderzoeken of er een verband bestaat tussen twee variabelen (Field, 2009). De sterkte van de correlaties werd bepaald op basis van tabel 3 (Schober, Boer, & Schwarte, 2018). Vervolgens werd voor jongens en meisjes apart het verband berekend tussen leesaccuratesse en de variabelen leeftijd en AVI-niveau.

Om leesaccuratesse te voorspellen op basis van leeftijd, geslacht en AVI-niveau, werd een meervoudige lineaire regressie berekend met leesaccuratesse als afhankelijke variabele en leeftijd, geslacht en AVI-niveau als onafhankelijke variabelen (Field, 2009). Vervolgens werd voor jongens en meisjes apart een meervoudige lineaire regressie berekend met leesaccuratesse als afhankelijke variabele en leeftijd en AVI-niveau als onafhankelijke variabelen.

Tabel 3. Interpretatie van geobserveerde correlatiecoëfficiënten

Absolute waarde van de geobserveerde correlatiecoëfficiënt	Interpretatie
0.00-0.10	Zeer zwakke correlatie
0.10-0.39	Zwakke correlatie
0.40-0.69	Matige correlatie
0.70-0.89	Sterke correlatie
0.90-1.00	Zeer sterke correlatie

*Opmerking.* Aangepast van ‘‘Correlation Coefficients: Appropriate Use and Interpretation’’ door Schober, P., Boer, C., & Schwarte, L.A., 2018, *Anesthesia & Analgesia*, 126(5), p. 1765

### 3.5.2 Frequenties van leesfouten

Om antwoord te kunnen geven op de onderzoeksvragen ‘*Wat zijn de meest frequente leesfouten die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar maken tijdens het hardop voorlezen van een Nederlandse tekst?*’ werden staafdiagrammen gemaakt waarbij de totale absolute frequenties van leesfouten van alle leerlingen samen werden afgezet tegen de typen leesfouten. Dit werd gedaan voor zowel alle lees pogingen als de laatste lees pogingen.

### 3.5.3 Verband tussen leesfouten en leeftijd, geslacht en AVI-niveau

Om antwoord te kunnen geven op de onderzoeksvraag ‘*Is er een verband tussen leesfouten en de variabelen leeftijd, geslacht en AVI-niveau?*’, werden Pearson’s correlatiecoëfficiënten berekend. Dit werd gedaan voor zowel alle lees pogingen als de laatste lees pogingen. De correlaties werden berekend voor de 10 meest geannoteerde leesfouten. Hierbij werd gerekend met proportionele waarden. Vervolgens werden de correlaties opnieuw berekend voor jongens en meisjes apart.

#### 3.5.4 Frequenties van leesstrategieën

Om antwoord te kunnen geven op de onderzoeksvragen *‘Wat zijn de meest frequente leesstrategieën die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar toepassen tijdens het hardop voorlezen van een Nederlandse tekst?’* werden staafdiagrammen gemaakt waarbij de totale absolute frequenties van leesstrategieën van alle leerlingen samen werden afgezet tegen de typen leesstrategieën. Dit werd gedaan voor zowel alle lees pogingen als de laatste lees pogingen.

#### 3.5.5 Verband tussen leesstrategieën en leeftijd, geslacht en AVI-niveau

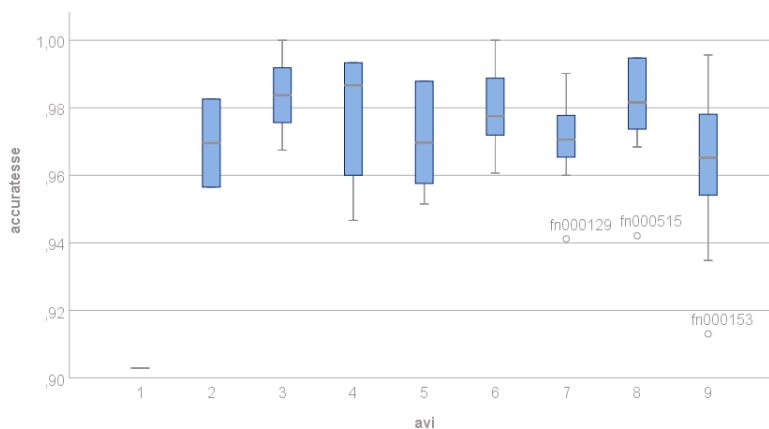
Om antwoord te kunnen geven op de onderzoeksvraag *‘Is er een verband tussen leesstrategieën en de variabelen leeftijd, geslacht en AVI-niveau?’*, werden Pearson’s correlatiecoëfficiënten berekend. Dit werd gedaan voor zowel alle lees pogingen als de laatste lees pogingen. De correlaties werden berekend voor de vijf meest geannoteerde leesstrategieën. Hierbij werd gerekend met proportionele waarden. Vervolgens werden de correlaties opnieuw berekend voor jongens en meisjes apart.

## 4. Resultaten

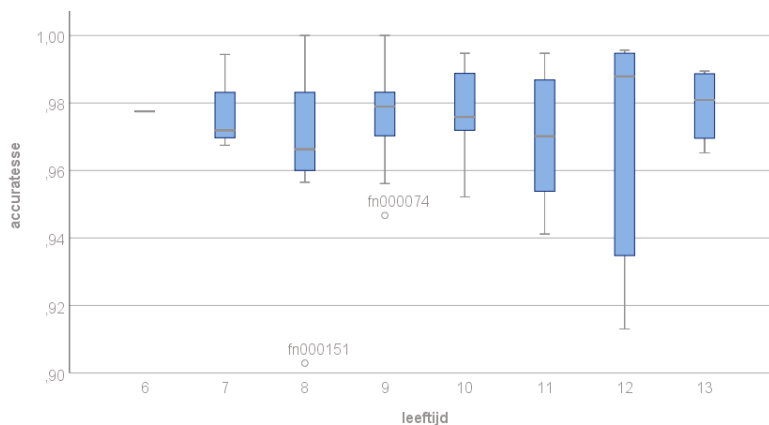
### 4.1 Verband tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau

Leesaccuratesse werd per leerling berekend door het aantal woorden dat goed werd gelezen in de laatste leesponing te delen door het totaal aantal woorden in de gelezen tekst.

Leesaccuratesse werd in een boxplot afgezet tegen zowel AVI-niveau (figuur 2) als leeftijd (figuur 3). Hieruit bleek dat er een aantal uitbijters waren, namelijk de leerlingen met proefpersoonnummer 129, 515, 153, 74 en 151. Behalve de leerling met nummer 74, scoorden deze leerlingen lager dan de minimale accuratessescore van .95. Daarnaast scoorden een elfjarige en twaalfjarige lager dan de minimale accuratessescore. Dit waren de leerlingen met proefpersoonnummer 106 en 108. Deze leerlingen maakten proportioneel gezien meer fouten dan de andere leerlingen, maar lieten verder geen bijzonderheden zien. De leerlingen met een accuratessescore lager dan .95 werden niet meegenomen in de analyses. De gemiddelde leesaccuratesse zonder uitbijters was 0.98 ( $SD = .01$ , bereik = .95-1.00).



Figuur 2. Boxplot AVI-niveau afgezet tegen leesaccuratesse



Figuur 3. Boxplot leeftijd afgezet tegen leesaccuratesse

Om te kijken of er een verband bestond tussen leesaccuratesse en de variabelen leeftijd, geslacht en AVI-niveau, werden Pearson's correlatiecoëfficiënten berekend. Hieruit bleek dat leesaccuratesse niet significant correleerde met leeftijd,  $r = .15$ ,  $p = .23$ , geslacht,  $r = -.14$ ,  $p = .91$  en AVI-niveau,  $r = < .01$ ,  $p = .99$  (zie tabel 4).

Vervolgens werd voor jongens en meisjes apart berekend of er een verband bestond tussen leesaccuratesse en de variabelen leeftijd en AVI-niveau. Bij jongens bleek dat er een matig positief verband was tussen leesaccuratesse en leeftijd,  $r = .41$ ,  $p = .02$ . Er werd geen verband gevonden tussen leesaccuratesse en AVI-niveau,  $r = .16$ ,  $p = .37$ . In tegenstelling tot

de jongens bleek bij de meisjes leesaccuratesse niet significant te correleren met leeftijd,  $r = -.12$ ,  $p = .55$  en tevens niet met AVI-niveau,  $r = -.19$ ,  $p = .33$  (zie tabel 4).

Een meervoudige lineaire regressie werd berekend om leesaccuratesse te voorspellen op basis van leeftijd, AVI-niveau en geslacht. Hieruit bleek dat leeftijd, AVI-niveau en geslacht geen significante voorspellers waren van leesaccuratesse,  $F(3, 58) = .67$ ,  $p = .57$ ,  $R^2 = .03$ .

Omdat bij jongens leesaccuratesse bleek te correleren met leeftijd, werd voor jongens en meisjes apart een meervoudige lineaire regressie berekend om leesaccuratesse te voorspellen op basis van leeftijd en AVI-niveau. Bij jongens bleek dat leeftijd en AVI-niveau geen significante voorspellers waren van leesaccuratesse ( $F(2, 29) = 2.89$ ,  $p = .07$ ,  $R^2 = .17$ ). Ook bij meisjes bleken leeftijd en AVI-niveau geen significante voorspellers te zijn van leesaccuratesse,  $F(2, 27) = .49$ ,  $p = .62$ ,  $R^2 = .04$ .

Tabel 4. Pearson's correlatiecoëfficiënten met afhankelijke variabele accuratesse en onafhankelijke variabelen leeftijd, geslacht en AVI-niveau

		Leeftijd		Geslacht		AVI-niveau	
		<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Accuratesse	Alle leerlingen	.15	.23	-.14	.91	< .01	.99
	Jongens	.41*	.02	.	.	.16	.37
	Meisjes	-.12	.55	.	.	-.19	.33

\*  $p \leq .05$  (tweezijdig)

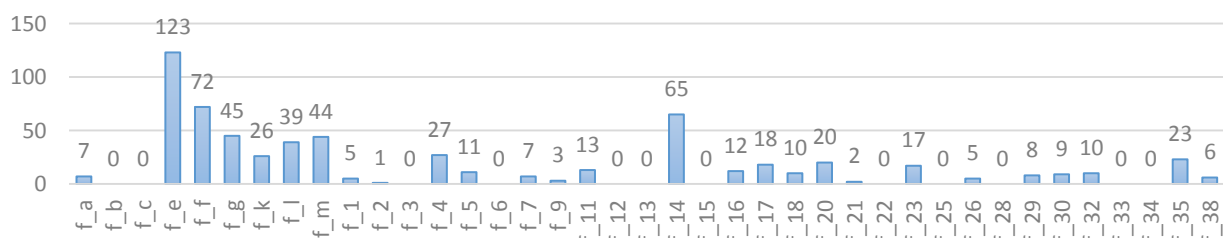
## 4.2 Frequenties van leesfouten

### 4.2.1 Leesfouten in alle leesopgingen

In totaal werden 628 fouten geannoteerd, waarbij sommige leesopgingen meerdere labels kregen (zie figuur 4). De 10 meest frequente leesfouten waren:

- *e*: een woord werd vervangen door een ander bestaand woord dat orthografisch gelijk is;
- *f*: een woord werd vervangen door een ander bestaand woord dat semantisch gelijk is;
- *l4*: vervanging van medeklinker door een andere medeklinker;
- *m*: een woord werd vervangen door een ander bestaand woord dat zowel orthografisch als semantisch niet gelijk is en niet semantisch passend is binnen het stuk tekst dat werd gelezen;
- *l*: invoeging van een extra woord dat niet in de tekst staat en semantisch niet passend is binnen het stuk tekst dat werd gelezen;
- *g*: een woord werd vervangen door een pseudowoord;
- *k*: invoeging van een extra woord dat niet in de tekst staat en semantisch passend is binnen het stuk tekst dat werd gelezen;
- *4*: vervanging van klinker door een andere klinker;
- *35*: invoeging van een willekeurige medeklinker;
- *17*: vervanging van een lettergreep door een andere lettergreep of letter.

Een aantal labels voor leesfouten werd nooit geannoteerd of slechts een aantal keer.



Figuur 4. Frequentie per leesfout (alle leesopgingen) in absolute waarden

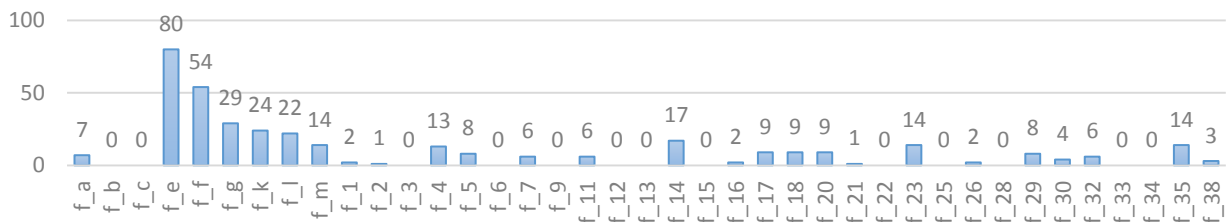


## 4.2.2 Leesfouten in laatste leespogingen

In totaal werden 364 fouten geannoteerd voor de laatste leespogingen (zie figuur 5). De 10 meest frequente leesfouten in de laatste leespoging waren:

- *e*: een woord werd vervangen door een ander bestaand woord dat orthografisch gelijk is;
- *f*: een woord werd vervangen door een ander bestaand woord dat semantisch gelijk is;
- *k*: invoeging van een extra woord dat niet in de tekst staat en semantisch passend is binnen het stuk tekst dat werd gelezen;
- *l*: invoeging van een extra woord dat niet in de tekst staat en semantisch niet passend is binnen het stuk tekst dat werd gelezen;
- *g*: een woord werd vervangen door een pseudowoord;
- *m*: een woord werd vervangen door een ander bestaand woord dat zowel orthografisch als semantisch niet gelijk is en niet semantisch passend is binnen het stuk tekst dat werd gelezen;
- *14*: vervanging van medeklinker door een andere medeklinker;
- *4*: vervanging van klinker door een andere klinker;
- *23*: weglating van een willekeurige medeklinker;
- *35*: invoeging van een willekeurige medeklinker.

Alle labels die geannoteerd werden voor alle leespogingen, kwamen ook voor bij de laatste leespogingen.



Figuur 5. Frequentie per leesfout (laatste leespogingen) in absolute waarden

## 4.3 Verband tussen leesfouten en leeftijd, geslacht en AVI-niveau

### 4.3.1 Leesfouten in alle leespogingen

Om te kijken of er een verband bestond tussen leesfouten en de variabelen leeftijd, geslacht en AVI-niveau, werden Pearson's correlatiecoëfficiënten berekend over de proportionele waarden. Wanneer alle leespogingen werden meegenomen, bleek dat er zwak negatieve correlatie was tussen leeftijd en *e*,  $r = -.30$ ,  $p = .02$  en een zwak positieve correlatie tussen leeftijd en *g*,  $r = .28$ ,  $p = .03$ . Er was een zwak negatieve correlatie tussen geslacht en *g*,  $r = -.25$ ,  $p = .04$ , waarbij meisjes gemiddeld een kleine proportie woorden vervingen door pseudowoorden. En er was zwak positieve correlatie tussen AVI-niveau en *f*,  $r = .24$ ,  $p = .05$  en tussen AVI-niveau en *35*,  $r = .25$ ,  $p = .50$  (zie tabel 5).

De correlaties werden opnieuw berekend voor jongens en meisjes apart. Hieruit bleek dat er bij jongens een matig negatieve correlatie was tussen leeftijd en *e*,  $r = -.53$ ,  $p < .01$  en een zwak positieve correlatie tussen leeftijd en *g*,  $r = .39$ ,  $p = .03$ . Bij meisjes was er een zwak positieve correlatie tussen leeftijd en *k*,  $r = .41$ ,  $p = .03$ .

Tabel 5. Pearson's correlatiecoëfficiënten afhankelijke en onafhankelijke variabelen (alle leespogingen)

	Leeftijd		Geslacht		AVI-niveau	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
<i>e</i>	-.30*	.02	.05	.70	-.01	.95
<i>f</i>	.06	.66	-.14	.28	.24*	.05
<i>g</i>	.28*	.03	-.25*	.04	.12	.36
<i>k</i>	.23	.07	.02	.87	.07	.60
<i>l</i>	-.01	.95	.19	.13	-.09	.46
<i>m</i>	.10	.46	.01	.92	-.11	.40
4	-.02	.86	-.14	.27	.03	.83
14	-.05	.68	.07	.57	.03	.85
17	.20	.13	-.22	.08	.22	.08
35	.06	.67	-.13	.29	.25*	.05

\*  $p \leq .05$  (tweezijdig)

#### 4.3.2 Leesfouten in laatste leespogingen

Wanneer alleen de laatste leespogingen werden meegenomen, bleek dat er een zwak negatieve correlatie was tussen leeftijd en *e*,  $r = -.37$ ,  $p < .01$  en een zwak positieve correlatie tussen leeftijd en *g*,  $r = .28$ ,  $p = .03$ . Er was een zwak negatieve correlatie tussen geslacht en *f*,  $r = -.30$ ,  $p = .02$  en tussen geslacht en *g*,  $r = -.29$ ,  $p = .02$ , waarbij meisjes gemiddeld een kleinere proportie woorden vervingen door semantisch gerelateerde woorden of pseudoworden dan jongens. Er was een zwak positieve correlatie tussen AVI-niveau en *f*,  $r = .26$ ,  $p = .04$  (zie tabel 6).

De correlaties werden opnieuw berekend voor jongens en meisjes apart. Hieruit bleek dat er bij jongens een matig negatieve correlatie was tussen leeftijd en *e*,  $r = -.54$ ,  $p < .01$ . Bij meisjes was er een matig positieve correlatie tussen leeftijd en *f*,  $r = .42$ ,  $p = .02$ .

Tabel 6. Pearson's correlatiecoëfficiënten afhankelijke en onafhankelijke variabelen (laatste leespogingen)

	Leeftijd		Geslacht		AVI-niveau	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
<i>e</i>	-.37**	<.01	-.01	.95	<-.01	.97
<i>f</i>	.16	.23	-.30*	.02	.26*	.04
<i>g</i>	.28*	.03	-.29*	.02	.01	.95
<i>k</i>	.18	.16	-.15	.24	-.05	.70
<i>l</i>	.13	.30	.12	.32	.17	.19
<i>m</i>	.14	.28	<-.01	.97	-.03	.81
4	<-.01	.99	-.23	.06	.09	.46
14	-.22	.09	.19	.12	-.05	.71
23	.13	.31	-.15	.23	.22	.08
35	.06	.62	-.11	.39	.21	.09

\*  $p \leq .05$  (tweezijdig)

\*\*  $p \leq .01$  (tweezijdig)

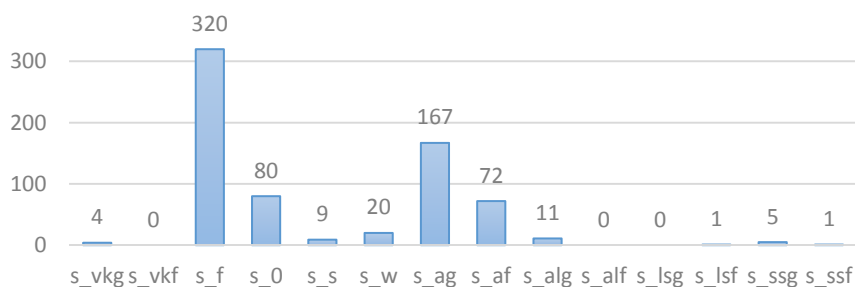
#### 4.4 Frequenties van leesstrategieën

##### 4.4.1 Leesstrategieën in alle leespogingen

In totaal werden er voor alle leespogingen 12860 leesstrategieën geannoteerd, waarvan 12170 keer *g* (woord werd correct gelezen). Om te kijken wat de meest frequente leesstrategieën waren, werden staafdiagrammen gemaakt waarbij de totale frequenties van leesstrategieën van alle leerlingen samen werden afgezet tegen de leesstrategieën (figuur 6). De op *g* na meest frequente leesstrategieën waren:

- *f*: woord werd fout gelezen;
- *ag*: eerste letter of begin van het woord werd correct gelezen;
- *0*: woord werd weggelaten;
- *af*: eerste letter of begin van het woord werd fout gelezen;
- *W*: woord werd gewisseld van plaats met een ander woord.

Verlengingen van klanken (*vkg* en *vkf*) en spelstrategieën (*\*s*, *lsg*, *lsf*, *ssg*, en *ssf*) werden niet of slechts een aantal keer geannoteerd.

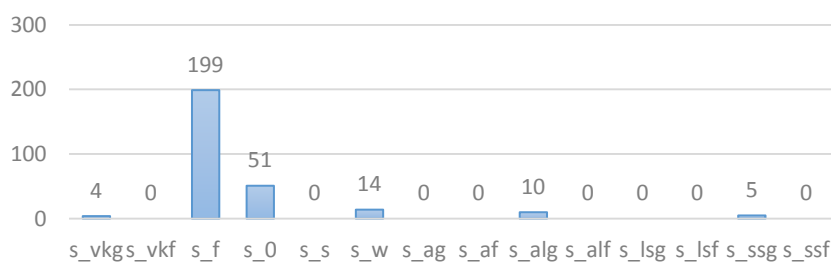


Figuur 6. Frequentie per leesstrategie (alle lees pogingen) in absolute waarden

#### 4.4.2 Leesstrategieën in laatste lees pogingen

In totaal waren 12081 lees pogingen de laatste lees poging (zie figuur 7). Van de totaal 12860 lees pogingen waren 779 lees pogingen niet de laatste lees poging. In totaal werden 11802 woorden correct gelezen (*g* en *vkg*) en 279 woorden werden fout gelezen in de laatste lees poging. De op *g* na meest frequente leesstrategieën waren:

- *f*: woord werd fout gelezen;
- *0*: woord werd weggelaten;
- *W*: woord werd gewisseld van plaats met een ander woord.
- *alg*: laatste letter of laatste deel van het woord werd correct gelezen, maar het begin van het woord werd niet volledig herhaald bij het corrigeren van een vorige lees poging;
- *ssg*: woord werd correct gespeld in lettergrepen, maar werd niet gevolgd door een poging tot gesynthetiseerd lezen.



Figuur 7. Frequentie per leesstrategie (laatste lees pogingen) in absolute waarden

### 4.5 Verband tussen leesstrategieën en leeftijd, geslacht en AVI-niveau

#### 4.5.1 Leesstrategieën in alle lees pogingen

Om te kijken of er een verband bestond tussen leesstrategieën en de variabelen leeftijd, geslacht en AVI-niveau, werden Pearson's correlatiecoëfficiënten berekend over de proportionele waarden. Wanneer alle lees pogingen werden meegenomen, bleek er geen correlatie te zijn tussen leeftijd en leesstrategieën en geslacht en leesstrategieën. Er was een

zwak positieve correlatie tussen AVI-niveau en  $f$ ,  $r = .29$ ,  $p = .02$  en een zwak negatieve correlatie tussen AVI-niveau en  $ag$ ,  $r = -.37$ ,  $p < .01$  (zie tabel 7).

De correlaties werden opnieuw berekend voor jongens en meisjes apart. Hieruit bleek dat er bij jongens geen correlatie was tussen leeftijd en leesstrategieën en AVI-niveau en leesstrategieën. Bij meisjes was er een zwak positieve correlatie tussen leeftijd en  $f$ ,  $r = .40$ ,  $p = .03$ . Daarnaast was er een zwak positieve correlatie tussen AVI-niveau en  $f$ ,  $r = .39$ ,  $p = .03$  en een matig negatieve correlatie tussen AVI-niveau en  $ag$ ,  $r = -.49$ ,  $p = .01$ .

Tabel 7. Pearson's correlatiecoëfficiënten met afhankelijke variabelen leesstrategieën en onafhankelijke variabelen leeftijd, geslacht en AVI-niveau (alle leesposingen)

	Leeftijd		Geslacht		AVI-niveau	
	$r$	$p$	$r$	$p$	$r$	$p$
$f$	.19	.14	-.12	.33	.29*	.02
$0$	.03	.83	.15	.24	.01	.95
$W$	-.11	.41	.15	.24	-.02	.86
$ag$	-.15	.26	.03	.84	-.37*	< .01
$af$	.02	.89	-.22	.09	.03	.84

\*  $p \leq .05$  (tweezijdig)

#### 4.5.2 Leesstrategieën in laatste leesposingen

Wanneer alleen de laatste leesposingen werden meegenomen, bleek er geen correlatie te zijn tussen leeftijd en leesstrategieën en AVI-niveau en leesstrategieën. Er was een matig negatieve correlatie tussen geslacht en  $f$ ,  $r = -.43$ ,  $p < .001$  (zie tabel 8), waarbij meisjes gemiddeld een kleinere proportie woorden fout lezen dan jongens.

De correlaties werden opnieuw berekend voor jongens en meisjes apart. Hieruit bleek dat er bij zowel jongens als meisjes geen correlatie was tussen leeftijd en leesstrategieën en AVI-niveau en leesstrategieën.

Tabel 8. Pearson's correlatiecoëfficiënten met afhankelijke variabelen leesstrategieën en onafhankelijke variabelen leeftijd, geslacht en AVI-niveau (laatste leesposingen)

	Leeftijd		Geslacht		AVI-niveau	
	$r$	$p$	$r$	$p$	$r$	$p$
$f$	.14	.29	-.43**	<.001	.15	.23
$0$	.02	.87	.17	.19	-.11	.38
$W$	-.04	.75	.17	.19	.03	.85
$alg$	-.22	.09	.15	.22	.03	.82
$ssg$	.11	.42	.18	.15	.03	.83

\*  $p \leq .05$  (tweezijdig)

\*\*  $p \leq .01$  (tweezijdig)

## 5. Discussie

Voor dit onderzoek werden spraakopnames beluisterd van 71 basisschoolleerlingen tussen de 6 en 13 jaar die hardop een tekst voorlezen. In het spraakmateriaal werden de leesstrategieën die toegepast werden en de leesfouten die daarbij gemaakt werden geannoteerd. Vervolgens werd onderzocht of er een verband bestond tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau. Daarnaast werd gekeken wat de meest frequente leesfouten en leesstrategieën waren. Bovendien werd onderzocht of er een verband bestond tussen de afhankelijke variabelen proporties leesfouten en proporties leesstrategieën en de onafhankelijke variabelen leeftijd, geslacht en AVI-niveau.

Uit de resultaten bleek dat sommige leerlingen een lager AVI-niveau beheersten dan verwacht werd bij hun leeftijd. Daarnaast bleek uit de resultaten dat niet alle leerlingen het streefniveau lezen hadden behaald aan het eind van de basisschool. Dit is in overeenstemming met de literatuur, omdat is gebleken dat in 2017 slechts 65 procent van de leerlingen het streefniveau lezen behaalde (Inspectie van het Onderwijs, 2018).

Een mogelijke verklaring voor dit resultaat is dat hardop voorlezen onder begeleiding te weinig wordt ingezet, omdat er veel individuele aandacht voor nodig is (Strik, 2019). Bovendien slagen leerkrachten van basisscholen er onvoldoende in om te gaan met verschillen tussen leerlingen (Inspectie van het Onderwijs, 2018). Een andere mogelijke verklaring is dat leesproblemen te laat onderkend en aangepakt worden, waardoor leesachterstanden ontstaan (Wentink & Verhoeven, 2008). Wanneer leerlingen met leesproblemen pas na hun negende jaar hulp krijgen, zijn de leesproblemen lastig te behandelen (Blonk et al., 2006). Daarnaast hebben steeds minder leerlingen plezier in lezen. Deze leerlingen behalen minder vaak het streefniveau lezen dan leerlingen die lezen wel leuk vinden. Dat lezen minder gewaardeerd wordt door leerlingen, zou te maken kunnen hebben met de toename van digitale media (Gubbels, Netten, & Verhoeven, 2017).

Het huidige onderzoek laat daarmee nogmaals zien dat niet alle leerlingen het streefniveau lezen hebben behaald als zij de basisschool verlaten. Op basis van dit onderzoek kan gesuggereerd worden dat het leesonderwijs meer aandacht nodig heeft en dat vooral het onderwijs aan en de ondersteuning van zwakke lezers beter kan (Inspectie van het Onderwijs, 2019). Hierbij moet wel rekening gehouden worden dat dit onderzoek is gebaseerd op een steekproef van 71 leerlingen, verdeeld over acht leeftijdsgroepen. Daarnaast was het niet bekend hoe het beheersingsniveau van de leerlingen werd bepaald. Mogelijk werd dit bepaald op basis van verouderde toetsingsgegevens. Wanneer de leerlingen recent voor het moment van de spraakopnames een AVI-toets zouden maken, zouden de resultaten mogelijk anders zijn. Bovendien was het niet bekend in welke groep de leerlingen zaten. Om deze reden kan geen algemene uitspraak gedaan worden over percentages leerlingen die wel of niet het streefniveau lezen hebben behaald aan het eind van een bepaalde groep of aan het eind van de basisschool. Het advies voor vervolgonderzoek is om recente toetsingsgegevens te gebruiken voor het bepalen van het beheersingsniveau en daarnaast het aantal leerlingen per leeftijdsgroep uit te breiden.

### 5.1 Verband tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau

De onderzoeksvraag luidde: *‘Is er een verband tussen leesaccuratesse en de variabelen leeftijd, geslacht en AVI-niveau?’*. Verwacht werd dat leeftijd, geslacht en AVI-niveau geen effect zouden hebben op leesaccuratesse. Uit de resultaten bleek dat de gemiddelde leesaccuratesse bijna maximaal was. Zoals verwacht, werd er geen verband gevonden tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau. Echter, wanneer voor jongens en meisjes apart het verband werd berekend tussen accuratesse, leeftijd en AVI-niveau, bleek dat er bij jongens een matig positief verband was tussen accuratesse en leeftijd. Uit een meervoudige lineaire regressie bleek dat leeftijd, geslacht en AVI-niveau geen significante voorspellers

waren van accuratesse. Ook niet wanneer de meervoudige lineaire regressie werd berekend voor jongens en meisjes apart.

Hieruit kan worden opgemaakt dat leesaccuratesse tijdens het lezen van een tekst op beheersingsniveau niet in verband staat tot leeftijd, geslacht en AVI-niveau. Dat er geen verband werd gevonden tussen leeftijd en leesaccuratesse en was in overeenstemming met de verwachting. Ook Cleuren (2009) vond geen verband tussen klas en leesaccuratesse. Dat er geen verband werd gevonden tussen geslacht en leesaccuratesse en tussen AVI-niveau en leesaccuratesse was ook in overeenstemming met de verwachting. De leerlingen lazen namelijk teksten op beheersingsniveau, waardoor het logisch is dat leerlingen vergelijkbare accuratessescores behaalden. In tegenstelling tot deze resultaten vond Cleuren (2009) in haar onderzoek dat er tijdens het lezen van een tekst op beheersingsniveau wel een significant verband was tussen AVI-niveau en leesaccuratesse en tussen geslacht en leesaccuratesse. Uit de resultaten bleek echter niet dat leerlingen met een lager AVI-niveau meer leesfouten maakten dan leerlingen met een hoger AVI-niveau. Uit de resultaten bleek wel dat jongens minder accuraat lazen dan meisjes. Daarnaast vonden Rutter en collega's (2004) in vier studies dat de percentages leesproblemen groter waren bij jongens dan bij meisjes.

Een mogelijke verklaring voor deze resultaten is het feit dat leerlingen op hun beheersingsniveau hebben gelezen. Het leesmateriaal werd aangepast aan het leesniveau van de leerlingen. Op het beheersingsniveau wordt minder dan  $\pm 5\%$  van de woorden fout gelezen (Struiksma et al., 2009), waardoor hoge accuratessescores worden behaald. Hierdoor werden mogelijke plafondeffecten bereikt. Deze effecten treden op wanneer scores op een afhankelijke variabele een maximum naderen. De invloed van onafhankelijke variabelen is klein, omdat de waarden van de afhankelijke variabele bijna zo hoog zijn als dat ze kunnen zijn (Cramer & Howitt, 2004). Aan de resultaten ligt mogelijk ook ten grondslag dat leessnelheid niet werd meegenomen. Dit werd niet meegenomen, omdat de leerlingen niet werden geïnstrueerd om zo snel mogelijk te lezen. Hierdoor hadden leerlingen mogelijk meer tijd en aandacht om leesfouten te verbeteren, waardoor de leesaccuratesse hoog was.

De resultaten zijn een aanvulling op de bestaande literatuur over leesaccuratesse, omdat het laat zien dat leesaccuratesse tijdens het lezen van een tekst op beheersingsniveau niet wordt beïnvloed door leeftijd, geslacht en AVI-niveau. Hierbij moet wel rekening worden gehouden dat in dit onderzoek uitsluitend teksten op beheersingsniveau werden gelezen en dat de invloed van leessnelheid niet werd meegenomen. Wanneer alle leerlingen dezelfde tekst zouden lezen of wanneer leessnelheid zou worden meegenomen, zouden de resultaten mogelijk anders zijn. Om deze reden kan geen uitspraak worden gedaan over het verband tussen leesaccuratesse en leeftijd, geslacht en AVI-niveau in het algemeen. Voor vervolgonderzoek naar leesaccuratesse wordt aangeraden alle leerlingen dezelfde tekst aan te bieden en leessnelheid mee te nemen.

## 5.2 Frequenties van leesfouten

De onderzoeksvraag luidde: *'Wat zijn de meest frequente leesfouten die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar maken tijdens het hardop voorlezen van een Nederlandse tekst?'*. Verwacht werd dat vervangingen van medeklinkers en klinkers, vervangingen in de volgorde van aangrenzende letters, vervangingen van een bestaand woord door een pseudowoord of een ander bestaand woord dat orthografisch gelijk is, toevoegingen van een willekeurige medeklinker, decodering die niet werd gevolgd door een poging tot gesynthetiseerd lezen en weglatingen van woorden vaker zouden voorkomen dan de andere geannoteerde leesfouten (zie bijlage 1). Uit de resultaten bleek dat, wanneer alle lees pogingen werden meegenomen, de volgende leesfouten het meest frequent voorkwamen:

- een woord werd vervangen door een ander bestaand woord dat orthografisch gelijk was;

- een woord werd vervangen door een ander bestaand woord dat semantisch gelijk was;
- vervanging van een medeklinker door een andere medeklinker;
- een woord werd vervangen door een ander bestaand woord dat zowel orthografisch als semantisch niet gelijk was en niet semantisch passend was binnen het stuk tekst dat werd gelezen;
- invoeging van een extra woord dat niet in de tekst stond en semantisch niet passend was binnen het stuk tekst dat werd gelezen;
- een woord werd vervangen door een pseudowoord;
- invoeging van een extra woord dat niet in de tekst stond en semantisch passend was binnen het stuk tekst dat werd gelezen;
- vervanging van een klinker door een andere klinker;
- invoeging van een willekeurige medeklinker;
- vervanging van een lettergreep door een andere lettergreep of letter.

Wanneer alleen de laatste lees pogingen werden meegenomen, bleken vrijwel dezelfde leesfouten het meest frequent voor te komen. Het enige verschil was dat bij de laatste lees pogingen willekeurige medeklinkers vaker werden weggelaten dan dat lettergrepen werden vervangen door een andere lettergreep of letter. Een aantal leesfouten werd nooit of slechts een aantal keer geannoteerd.

De meest frequente leesfouten die werden gevonden in het onderzoek zijn deels in overeenstemming met de verwachting op basis van de studie van Cleuren (2009). Uit haar studie bleken fouten door verandering in de volgorde van aangrenzende letters, vervanging van korte klinkers door lange klinkers of omgekeerd en decoding die niet werd gevolgd door een poging tot gesynthetiseerd lezen tot de meest frequente leesfouten te behoren. Een mogelijke verklaring voor dit verschil kan zijn dat het protocol dat werd gebruikt in het onderzoek van Cleuren (2009) verschillend is van het protocol dat werd gebruikt in dit onderzoek. Hierdoor veranderden mogelijk de proporties van leesfouten. Daarnaast is het in het onderzoek van Cleuren (2009) niet duidelijk of de meest frequente leesfouten zijn gebaseerd op fouten tijdens het lezen van een tekst of ook op fouten tijdens het lezen van woordenlijsten met niet gerelateerde (pseudo)woorden. Bovendien nam zij, in tegenstelling tot het huidige onderzoek, groep 3 leerlingen mee in haar onderzoek. Aan het resultaat ligt mogelijk ook ten grondslag dat het onderzoek van Cleuren (2009) werd uitgevoerd in Vlaanderen. Mogelijk maken leerlingen uit Vlaanderen andere leesfouten dan leerlingen uit Nederland.

Uit de resultaten kan worden opgemaakt dat woorden vaker werden vervangen door bestaande woorden dan door pseudowoorden. Dit is in overeenstemming met de resultaten uit het onderzoek van Cleuren (2009), die vond dat Vlaamse basisschoolleerlingen meer woorden vervingen door echte woorden dan door pseudowoorden tijdens het lezen van een Nederlandse tekst. Ook Danielsson (2001) vond dat beginnende lezers meer woorden vervingen door echte woorden dan door pseudowoorden tijdens het lezen van teksten in het Zweeds (dat orthografisch vrijwel transparant is, net zoals het Nederlands). Dit resultaat kan verklaard worden vanuit de leesstrategieën die de leerlingen gebruikten. Leerlingen gebruikten vaker lexicale strategieën dan sublexicale strategieën (nader toegelicht in 5.4). Wanneer lexicale strategieën worden gebruikt tijdens het lezen van een tekst, is het waarschijnlijk dat de poging tot het lezen van het woord resulteert in een echt woord (Biemiller; Weber, in Danielsson, 2001, p. 413), terwijl een poging gebaseerd op sublexicale strategieën kan resulteren in een pseudowoord (Danielsson, 2001).

Daarnaast kan uit de resultaten worden opgemaakt dat woorden vaker werden vervangen door orthografisch gelijke woorden dan door semantisch gelijke of semantisch passende woorden. Dit is niet in overeenstemming met de verwachting, omdat orthografische gelijkheid het gebruik van de minder frequente sublexicale strategieën suggereert, terwijl

orthografisch niet gelijke woorden de frequente lexicale strategieën suggereren (Cleuren, 2009; Laing, 2002). Een mogelijke verklaring voor dit verschil in resultaat is dat semantische gelijkheid in het huidige onderzoek te streng werd beoordeeld, waardoor dit label minder frequent werd geannoteerd. Zo werd bijvoorbeeld de lezing ‘groot’ in plaats van ‘grote’ geannoteerd als orthografisch gelijk, maar niet als semantisch gelijk. Semantische gelijkheid zou daarom in het annotatieprotocol duidelijker gedefinieerd moeten worden.

Andere fouten die voorkwamen, waren fouten door invoegingen van woorden. Dat fouten door invoeging minder frequent voorkwamen dan fouten door vervanging, is in overeenstemming met de studie van Goodman (in D’Angelo & Mahlios, 1983, p. 780). Een mogelijke verklaring voor de invoegingen is dat leerlingen woorden invoegen naarmate ze vaardig worden in lezen (D’Angelo & Mahlios, 1983). Leerlingen leren vanaf groep 4 lezen met behulp van context. Leerlingen proberen op basis van context te voorspellen welke woorden er komen. Mogelijk produceren zij hierdoor woorden die niet in de tekst staan, maar die zij wel hadden verwacht.

Bovendien bleek dat fouten door vervanging, invoeging of weglating van medeklinkers en klinkers vaker betrekking hadden op willekeurige medeklinkers en klinkers dan op specifieke (clusters van) medeklinkers (zoals /sch-/, /ng/ en visueel gelijke medeklinkers) en specifieke (clusters van) klinkers (zoals klinkers in leenwoorden en tweeklanken). Dit resultaat kan mogelijk verklaard worden vanuit de leesinstructie die leerlingen krijgen. Vanaf groep 4 leren leerlingen woorden met medeklinkerclusters en meerlettergrepige woorden lezen (Verhoeven & van Leeuwe, 2003). De elementaire leeshandeling wordt verkort tot lezen met behulp van clusters en spellingspatronen. Leerlingen leren in plaats van losse letters, zoals bij de elementaire leeshandeling, lettercombinaties en spellingpatronen te herkennen (Huizenga, 2016; TULE; 2008). Mogelijk herkennen leerlingen deze specifieke (clusters van) medeklinkers en specifieke (clusters van) klinkers daardoor gemakkelijker en worden deze minder frequent fout gelezen dan andere medeklinkers of klinkers.

Het huidige onderzoek draagt bij aan de bestaande literatuur over leesfouten die gemaakt worden door Nederlandse basisschoolleerlingen tijdens het lezen van een tekst, omdat hier in Nederland nog maar weinig onderzoek naar is gedaan. Het vergroot de kennis over hoe Nederlandse basisschoolleerlingen schriftelijke informatie verwerken (Leu, 1982) en het geeft inzicht in hoe leerlingen leren lezen (Goikoetxea, 2006). Tevens kunnen de resultaten gebruikt worden om veelvoorkomende leesfouten te karakteriseren en te categoriseren, die vervolgens in leessoftware, gebaseerd op automatische spraakherkenning, geïmplementeerd kunnen worden. De spraakherkenner moet getraind worden met een corpus dat voorgelezen spraak en annotaties van het leesgedrag bevat. Op basis hiervan kan de leessoftware leesfouten voorspellen, detecteren en herstellen (Cleuren, 2009; Mostow et al., 2002). Bij de interpretatie van de resultaten moet wel rekening gehouden worden dat de leesfouten van alle leerlingen werden samengenomen. Om deze reden kan geen uitspraak gedaan worden over de verdeling van fouten per leerling. Daarnaast moet er rekening mee worden gehouden dat de leesfouten die niet geannoteerd werden, mogelijk wel voorkomen. De leerlingen lazen teksten op beheersingsniveau en er werden in het onderzoek geen leerlingen uit groep 3 meegenomen. Mogelijk veranderen de frequenties van leesfouten wanneer leerlingen teksten op instructie- of frustratieniveau of woordenlijsten met niet gerelateerde (pseudo)woorden lezen. Bovendien is het mogelijk dat leerlingen uit groep 3 andere leesfouten maken dan de leerlingen die werden meegenomen in dit onderzoek. Om deze reden kan geen algemene uitspraak gedaan worden over welke leesfouten het meest frequent gemaakt worden, omdat dit afhankelijk zal zijn van het leesmateriaal en de groepen die worden meegenomen. Het advies voor vervolgonderzoek is om het materiaal uit te breiden met teksten op instructie- en frustratieniveau en woordenlijsten met niet gerelateerde



(pseudo)woorden en daarnaast het corpus aan te vullen met spraakmateriaal van leerlingen uit groep 3.

### 5.3 Verband tussen leesfouten en leeftijd, geslacht en AVI-niveau

De onderzoeksvraag luidde: *'Is er een verband tussen leesfouten en de variabelen leeftijd, geslacht en AVI-niveau?'*. Verwacht werd dat oudere leerlingen en leerlingen met een hoger AVI-niveau proportioneel gezien minder vaak woorden zouden vervangen door orthografisch gelijke woorden of pseudowoorden en meer door semantisch gelijke of semantisch passende woorden. Daarnaast werd verwacht dat meisjes proportioneel gezien minder vaak woorden zouden vervangen door pseudowoorden ten opzichte van jongens. Uit de resultaten bleek dat, wanneer alle lees pogingen werden meegenomen, er een zwak negatieve correlatie was tussen leeftijd en de proportie woorden die werd vervangen door woorden die orthografisch gelijk waren en een zwak positieve correlatie tussen leeftijd en de proportie woorden die werd vervangen door pseudowoorden. Er was een zwak negatieve correlatie tussen geslacht en de proportie woorden die werd vervangen door pseudowoorden. Daarnaast was er een zwak positieve correlatie tussen AVI-niveau en de proportie woorden die werd vervangen door semantisch gelijke of semantisch passende woorden en tussen AVI-niveau en de proportie woorden waarbij een willekeurige medeklinker werd ingevoegd. De correlaties werden opnieuw berekend voor jongens en meisjes apart. Hieruit bleek dat er bij jongens een matig negatieve correlatie was tussen leeftijd en de proportie woorden die werd vervangen door woorden die orthografisch gelijk waren en een zwak positieve correlatie tussen leeftijd en de proportie woorden die werd vervangen door pseudowoorden. Bij meisjes was er een zwak positieve correlatie tussen leeftijd en de proportie invoegingen van een extra woord dat niet in de tekst stond en semantisch passend was binnen het stuk tekst dat werd gelezen.

Uit de resultaten bleek dat, wanneer alleen de laatste lees pogingen werden meegenomen, er wederom een zwak negatieve correlatie was tussen leeftijd en de proportie woorden die werd vervangen door woorden die orthografisch gelijk waren en een zwak positieve correlatie tussen leeftijd en de proportie woorden die werd vervangen door pseudowoorden. Er was een zwak negatieve correlatie tussen geslacht en de proportie woorden die werd vervangen door pseudowoorden en tussen geslacht en de proportie woorden die werd vervangen door semantisch gelijke of semantisch passende woorden. Bovendien was er een zwak positieve correlatie tussen AVI-niveau en de proportie woorden die werd vervangen door semantisch gelijke of semantisch passende woorden. De correlaties werden opnieuw berekend voor jongens en meisjes apart. Hieruit bleek dat er bij jongens een matig negatieve correlatie was tussen leeftijd en de proportie woorden die werd vervangen door woorden die orthografisch gelijk waren. Bij meisjes was er een matig positieve correlatie tussen leeftijd en de proportie woorden die werd vervangen door semantisch gelijke of semantisch passende woorden.

Hieruit kan worden opgemaakt dat naarmate leerlingen ouder worden of naarmate ze een hoger AVI-niveau bereiken, de proportie woorden die wordt vervangen door orthografisch gelijke woorden afneemt, terwijl de proportie pseudowoorden, semantisch gelijke of semantisch passende woorden en invoegingen van een willekeurige medeklinker (alleen voor alle lees pogingen) toeneemt. Daarnaast vervangen meisjes ten opzichte van jongens minder vaak woorden door pseudowoorden en semantisch gelijke of semantisch passende woorden (alleen voor de laatste lees pogingen). Het gaat voorbij aan het doel van deze studie om de correlaties voor jongens en meisjes apart te bespreken.

Een afname van orthografisch gelijke woorden en een toename van semantisch gelijke of semantisch passende woorden bij oudere leerlingen en leerlingen met een hoger AVI-niveau is in overeenstemming met de verwachting, omdat sublexicale strategieën afnemen naarmate leerlingen ouder worden en een hoger AVI-niveau beheersen. Dit resultaat kan

verklaard worden vanuit de veronderstelling dat orthografische gelijkheid het gebruik van sublexicale strategieën suggereert, terwijl orthografisch niet gelijke woorden lexicale strategieën suggereren (Cleuren, 2009; Laing, 2002). Cleuren (2009) vond een leeftijdsafhankelijke verschuiving in het gebruik van leesstrategieën. Hoe ouder de leerlingen, des te meer gebruik gemaakt werd van lexicale strategieën. Aangezien leeftijd en AVI-niveau samenhangen, kan gesuggereerd worden dat een toename van leeftijd en daarmee een toename van AVI-niveau gepaard gaat met een grotere proportie lexicale strategieën. Deze grotere proportie lexicale strategieën leidt waarschijnlijk tot meer semantische fouten dan tot fouten door orthografische gelijkheid.

Een toename van pseudowoorden bij oudere leerlingen is niet in overeenstemming met de verwachting. Verwacht werd dat de proportie pseudowoorden afneemt naarmate leerlingen ouder worden, omdat het gebruik van pseudowoorden sublexicale strategieën suggereert. Een mogelijke verklaring voor dit resultaat is dat de toegenomen woordlengte in teksten met een hoger AVI-niveau leidt tot een grotere proportie pseudowoorden. Leerlingen maken tijdens het lezen van niet gerelateerde woorden meer gebruik van sublexicale strategieën naarmate de woordlengte toeneemt (Cleuren, 2009).

Een toename van de proportie fouten door invoeging van een willekeurige medeklinker naarmate leerlingen een hoger AVI-niveau bereiken werd niet gevonden in eerder onderzoek. Voorgaand onderzoek naar leesfouten in het Nederlands richtte zich met name op fouten op woordniveau. Om deze reden was er geen verwachting over het verband tussen deze fout en leeftijd, geslacht en AVI-niveau.

Dat meisjes ten opzichte van jongens proportioneel gezien minder vaak woorden vervangen door pseudowoorden werd verwacht op basis van de studie van Cleuren (2009). Zij vond dat meisjes ten opzichte van jongens tijdens het lezen van woordenlijsten met niet gerelateerde woorden kleinere proporties woorden vervangen door niet bestaande woorden. Een mogelijke verklaring voor dit resultaat is dat meisjes ten opzichte van jongens minder geneigd zijn sublexicale strategieën te gebruiken om een woord te lezen. Ze zijn daarentegen meer geneigd een woord te raden of over te slaan (Cleuren, 2009).

Het resultaat dat meisjes ten opzichte van jongens proportioneel gezien minder vaak woorden vervangen door semantisch gerelateerde woorden werd niet verwacht op basis van de studie van Cleuren (2009). Zij vond namelijk dat meisjes en jongens proportioneel gezien evenveel woorden vervangen door bestaande woorden tijdens het lezen van woordenlijsten met niet gerelateerde woorden. Echter maakte Cleuren (2009) hierbij geen onderscheid tussen bestaande woorden die orthografisch dan wel semantisch gelijk waren. Tevens was het leesmateriaal verschillend.

De resultaten vullen de bestaande literatuur over de leesontwikkeling van Nederlandse basisschoolleerlingen aan, omdat het laat zien dat leesfouten veroorzaakt door sublexicale strategieën afnemen terwijl leesfouten door lexicale strategieën toenemen. Hierbij moet wel rekening worden gehouden dat de gevonden relaties zwak tot matig waren. In dit onderzoek werden veel leesfouten meegenomen, waardoor mogelijke bodemeffecten werden bereikt. Deze effecten treden op wanneer de scores op een afhankelijke variabelen zo laag zijn, dat effecten van onafhankelijke variabelen niet worden gevonden (Cramer & Howitt, 2004). Deze bodemeffecten werden mogelijk ook bereikt doordat de leerlingen teksten op beheersingsniveau lazen. Hierdoor maakten zij weinig leesfouten. Uit onderzoek van Goodman (1965) is gebleken dat er minder leesfouten worden gemaakt wanneer woorden in context worden gelezen ten opzichte van wanneer woordlijsten met niet gerelateerde (pseudo)woorden worden gelezen. Daarnaast moet er rekening mee worden gehouden dat dit onderzoek een cross-sectioneel onderzoek betrof. Hierdoor was er slechts een meetmoment, waardoor causale verbanden moeilijk na te gaan zijn. Bij een longitudinaal onderzoek daarentegen wordt op verschillende tijdstippen gemeten en kunnen verbanden gelegd

worden tussen de metingen op de verschillende momenten. Hierdoor kan onderzocht worden of het waarschijnlijk is dat de verandering in leeftijd of AVI-niveau voorafgaat aan de verandering in leesfouten (Baarda et al., 2017). Zodra leerlingen teksten op instructie- of frustratieniveau of woordlijsten met niet gerelateerde (pseudo)woorden zouden lezen, onder tijdsdruk zouden lezen of wanneer er een longitudinaal onderzoek zou plaatsvinden, zouden de resultaten mogelijk anders zijn. Om deze redenen kan geen algemene uitspraak gedaan worden over een causaal verband tussen leesfouten en leeftijd, geslacht en AVI-niveau. Het advies voor vervolgonderzoek is om het verband tussen leesfouten en leeftijd, geslacht en AVI-niveau te onderzoeken in een longitudinaal onderzoek, waarbij leerlingen woordlijsten of teksten boven hun beheersingsniveau lezen onder tijdsdruk.

#### 5.4 Frequenties van leesstrategieën

De onderzoeksvraag luidde: *‘Wat zijn de meest frequente leesstrategieën die Nederlandse basisschoolleerlingen van 6 t/m 13 jaar toepassen tijdens het hardop voorlezen van een Nederlandse tekst?’*. Verwacht werd dat leerlingen vaker gebruik zouden maken van lexicale leesstrategieën dan sublexicale leesstrategieën. Wanneer alle lees pogingen werden meegenomen, bleek dat correcte lezing van het woord het meest frequent voorkwam. De leesstrategieën die daarna het meest frequent voorkwamen, waren: foute lezing van het woord; eerste letter of begin van het woord werd correct gelezen; woord werd weggelaten; eerste letter of begin van het woord werd fout gelezen; en woord werd gewisseld van plaats met een ander woord. Verlengingen van klanken en spelstrategieën kwamen vrijwel niet voor. Wanneer alleen de laatste lees pogingen werden meegenomen, bleek wederom dat correcte lezing van het woord het meest frequent voorkwam. De leesstrategieën die daarna het meest frequent voorkwamen, waren: foute lezing van het woord; woord werd weggelaten; woord werd gewisseld van plaats met een ander woord; laatste letter of laatste deel van het woord werd correct gelezen, maar het begin van het woord werd niet volledig herhaald bij het corrigeren van een vorige poging; en woord werd correct gespeld in lettergrepen, maar werd niet gevolgd door een poging tot gesynthetiseerd lezen.

Hieruit kan worden opgemaakt dat woorden vaker met behulp van de visuele woordvorm of context werden gelezen dan met behulp van de elementaire leeshandeling, clusters en spellingspatronen of morfologische analyse. Dit is in overeenstemming met de verwachting dat lexicale leesstrategieën vaker voorkomen dan sublexicale leesstrategieën, dat ook werd gevonden in de studie van Cleuren (2009).

Een mogelijke verklaring voor dit resultaat is dat woorden in aanwezigheid van context vrijwel succesvol gelezen kunnen worden op basis van slechts een gedeeltelijke visuele analyse van de te lezen woorden (Cleuren, 2009). Daarnaast is de veronderstelling dat lexicale strategieën worden toegepast bij bekende woorden en sublexicale strategieën bij onbekende of laagfrequente woorden (Verhoeven & van Leeuwe, 2003). Mogelijk bevatten de teksten op beheersingsniveau voornamelijk bekende woorden. Een andere verklaring zou kunnen zijn dat de leesinstructie vanaf groep 4 gericht is op het sneller herkennen van woorden (Huizenga, 2016; Inspectie van het Onderwijs, 2006; Verhoeven, 2002; Verhoeven & Aarnoutse, 1999), waardoor sublexicale strategieën waarschijnlijk afnemen. Aan het resultaat ligt mogelijk ook ten grondslag dat leerlingen alleen teksten op hun beheersingsniveau hebben gelezen. Om een bepaald beheersingsniveau te halen, moet binnen een bepaalde tijd gelezen worden. Dat kan alleen als niet alle woorden gespeld worden. Op het beheersingsniveau zullen leerlingen daarom weinig gebruikmaken van sublexicale strategieën.

De resultaten vullen de bestaande literatuur over leesstrategieën die toegepast worden door Nederlandse basisschoolleerlingen tijdens het lezen van een tekst aan, omdat hier in Nederland nog maar weinig onderzoek naar is gedaan. Ook deze resultaten vergroten de

kennis over hoe Nederlandse basisschoolleerlingen schriftelijke informatie verwerken (Leu, 1982) en hoe leerlingen leren lezen (Goikoetxea, 2006). Op basis van dit onderzoek kan gesuggereerd worden dat leerlingen tijdens het lezen van een tekst op beheersingsniveau vaker gebruikmaken van lexicale strategieën, omdat ze gebruik kunnen maken van syntactische en semantische informatie (Archer & Bryant, 2001; Goodman, 1965).

Hierbij moet wel rekening gehouden worden dat de leesstrategieën van alle leerlingen werden samengenomen. Om deze reden kan geen uitspraak gedaan worden over de verdeling van leesstrategieën per leerling. Ook moet rekening gehouden worden dat dit onderzoek alleen teksten op beheersingsniveau heeft meegenomen. Daarnaast werden er geen leerlingen uit groep 3 meegenomen in het onderzoek. De resultaten zouden mogelijk anders zijn wanneer het materiaal uitgebreid zou worden met teksten op instructie- en frustratieniveau en woordenlijsten met niet gerelateerde (pseudo)woorden en wanneer leerlingen uit groep 3 zouden worden meegenomen. Wanneer leerlingen teksten lezen op instructie- of frustratieniveau, zullen zij moeilijkere woorden tegenkomen en waarschijnlijk vaker gebruikmaken van sublexicale strategieën dan wanneer zij een tekst lezen op beheersingsniveau. Daarnaast is uit onderzoek van Archer en Bryant (2001) en onderzoek van Goodman (1995) gebleken dat leerlingen in staat zijn woorden te lezen in context die zij niet kunnen lezen in isolatie. Dit suggereert dat leerlingen tijdens het lezen van woorden in een tekst andere strategieën gebruiken dan wanneer zij woorden lezen in isolatie. Dit wordt bevestigd door het onderzoek van Cleuren (2009). Uit haar onderzoek bleek dat leerlingen voornamelijk sublexicale strategieën gebruikten tijdens het lezen van niet gerelateerde (pseudo)woorden en voornamelijk lexicale strategieën gebruikten tijdens het lezen van woorden in een tekst. Wanneer leerlingen uit groep 3 worden meegenomen in het onderzoek, zullen sublexicale strategieën vaker voorkomen, omdat er in groep 3 voornamelijk aandacht is voor de elementaire leeshandeling. Om deze reden kan geen algemene uitspraak gedaan worden over welke leesstrategieën het meest frequent gebruikt worden, omdat dit afhankelijk zal zijn van het leesmateriaal en de groepen die worden meegenomen. Het advies voor vervolgonderzoek is om het materiaal uit te breiden met teksten op instructie- en frustratieniveau en woordenlijsten met niet gerelateerde (pseudo)woorden. Daarnaast wordt er geadviseerd het corpus aan te vullen met spraakmateriaal van leerlingen uit groep 3.

### 5.5 Verband tussen leesstrategieën en leeftijd, geslacht en AVI-niveau

De onderzoeksvraag luidde: *'Is er een verband tussen leesstrategieën en de variabelen leeftijd, geslacht en AVI-niveau?'*. Verwacht werd dat oudere leerlingen en leerlingen met een hoger AVI-niveau proportioneel gezien meer gebruik zouden maken van lexicale leesstrategieën ten opzichte van jongere leerlingen en leerlingen met een lager AVI-niveau. Daarnaast werd verwacht dat meisjes vaker gebruikmaken van lexicale leesstrategieën ten opzichte van jongens. Uit de resultaten bleek dat, wanneer alle leesprogingen werden meegenomen, er een zwak positieve correlatie was tussen AVI-niveau en de proportie fout gelezen woorden en dat er een zwak negatieve correlatie was tussen AVI-niveau en de proportie woorden waarvan de eerste letter of het begin van het woord correct werd gelezen. De correlaties werden opnieuw berekend voor jongens en meisjes apart. Hieruit bleek dat er bij jongens geen correlatie was tussen leeftijd en leesstrategieën en AVI-niveau en leesstrategieën. Bij meisjes was er een zwak positieve correlatie tussen leeftijd en de proportie fout gelezen woorden. Daarnaast was er een zwak positieve correlatie tussen AVI-niveau en de proportie fout gelezen woorden en een matig negatieve correlatie tussen AVI-niveau en de proportie woorden waarvan de eerste letter of het eerste deel van het woord correct werd gelezen.

Hieruit kan worden opgemaakt dat wanneer AVI-niveau toeneemt, de proportie fouten toeneemt en de proportie woorden waarvan de eerste letter of het begin van het woord correct

werd gelezen afneemt. Bij meisjes nam daarnaast de proportie fouten toe wanneer de leeftijd toenam. Dit is in overeenstemming met de verwachting dat jongere leerlingen en leerlingen met een lager AVI-niveau vaker sublexicale strategieën toepassen, terwijl oudere leerlingen en leerlingen met een hoger AVI-niveau vaker lexicale strategieën gebruiken. Dit werd ook gevonden in de studie van Cleuren (2009). Zij vond een leeftijdsafhankelijke verschuiving in het gebruik van leesstrategieën. Hoe ouder de leerlingen, des te meer gebruik gemaakt werd van lexicale strategieën. Aangezien leeftijd en AVI-niveau samenhangen, kan gesuggereerd worden dat een toename van leeftijd en daarmee een toename van AVI-niveau gepaard gaan met een grotere proportie lexicale strategieën.

Een mogelijke verklaring voor dit resultaat is dat leerlingen allereerst sublexicale strategieën leren toepassen en later pas lexicale strategieën. Naarmate het leesonderwijs vordert, is de leesinstructie meer gericht op het sneller herkennen van woorden (Huizenga, 2016; Inspectie van het Onderwijs, 2006; Verhoeven, 2002; Verhoeven & Aarnoutse, 1999). Aan het resultaat ligt mogelijk ook ten grondslag dat oudere leerlingen en leerlingen met een hoger AVI-niveau meer leeservaring hebben, waardoor zij woorden vaker hebben gelezen en daardoor deze woorden in een keer kunnen herkennen.

Uit de resultaten bleek dat, wanneer alleen de laatste lees pogingen werden meegenomen, er een matig negatieve correlatie was tussen geslacht en de proportie fout gelezen woorden, waarbij meisjes proportioneel gezien minder woorden fout lazen. De correlaties werden opnieuw berekend voor jongens en meisjes apart. Hieruit bleek dat er bij zowel jongens als meisjes geen correlatie was tussen leeftijd en leesstrategieën en tussen AVI-niveau en leesstrategieën.

Hieruit kan worden opgemaakt dat meisjes proportioneel gezien minder woorden fout lazen dan jongens. Dit is in overeenstemming met de verwachting, omdat uit onderzoek van Cleuren (2009) is gebleken dat meisjes tijdens het lezen van een tekst meer woorden weglieten dan jongens. Dit resultaat werd ook in het huidige onderzoek gevonden. Doordat de proportie weggelaten woorden bij meisjes groter is dan bij jongens, is de proportie fout gelezen woorden kleiner bij meisjes dan bij jongens.

Een verklaring voor dit resultaat hangt mogelijk samen met de leessnelheid (Cleuren, 2009). Mogelijk lezen meisjes sneller, waardoor zij eerder woorden over het hoofd zien en daardoor minder woorden fout lezen. Aan het resultaat zou ook ten grondslag kunnen liggen dat meisjes niet graag fouten maken en daarom het woord overslaan als ze het niet kunnen lezen.

Het huidige onderzoek is een aanvulling op de bestaande literatuur over de leesontwikkeling van Nederlandse basisschoolleerlingen, omdat het laat zien dat sublexicale strategieën afnemen naarmate leerlingen ouder worden en een hoger AVI-niveau beheersen. Hierbij moet wel rekening gehouden worden dat gevonden relaties zwak tot matig waren. In dit onderzoek werden veel leesstrategieën, waardoor mogelijke bodemeffecten werden bereikt (zie 5.3). Doordat de leerlingen teksten op beheersingsniveau lazen, lazen zij de meeste woorden direct goed. Andere leesstrategieën werden daardoor minder vaak ingezet. Daarnaast moet er rekening mee worden gehouden dat dit onderzoek een cross-sectioneel onderzoek betrof (zie 5.3). Zodra leerlingen teksten op instructie- of frustratieniveau of woordlijsten met niet gerelateerde (pseudo)woorden zouden lezen, onder tijdsdruk zouden lezen of wanneer er een longitudinaal onderzoek zou plaatsvinden, zouden de resultaten mogelijk anders zijn. Om deze redenen kan geen algemene uitspraak gedaan worden over een causaal verband tussen leesstrategieën en leeftijd, geslacht en AVI-niveau. Het advies voor vervolgonderzoek is om het verband tussen leesstrategieën en leeftijd, geslacht en AVI-niveau te onderzoeken in een longitudinaal onderzoek, waarbij leerlingen woordlijsten of teksten boven hun beheersingsniveau lezen onder tijdsdruk.

## 6. Conclusie

In dit onderzoek werd het leesgedrag van Nederlandse basisschoolleerlingen geanalyseerd door onderzoek te doen naar leesfouten en leesstrategieën tijdens hardop voorlezen van een tekst op beheersingsniveau. Hiertoe werd het verband tussen leesaccuratesse en de variabelen leeftijd, leesniveau en geslacht onderzocht. Vervolgens werd onderzocht wat de meest frequente leesfouten en leesstrategieën waren tijdens het hardop voorlezen van een Nederlandse tekst en als laatste werd het verband onderzocht tussen leesfouten en leesstrategieën en de variabelen leeftijd, geslacht en AVI-niveau. Hiervoor werd een kwantitatief cross-sectioneel onderzoek met bestaande data uitgevoerd.

Uit de resultaten bleek de gemiddelde leesaccuratesse bijna maximaal. Hierdoor werd er geen verband gevonden tussen leesaccuratesse en leeftijd, leesniveau en geslacht tijdens het hardop voorlezen van een tekst op beheersingsniveau. De leesfouten die de leerlingen maakten tijdens het hardop voorlezen van een Nederlandse tekst op beheersingsniveau bestonden voornamelijk uit vervangingen door orthografisch gelijke woorden, semantisch gelijke of semantisch passende woorden, pseudowoorden of woorden die zowel orthografisch als semantisch niet gelijk waren aan het doelwoord. Ook voegden zij woorden toe aan de tekst die al dan niet passend waren binnen de tekst die werd gelezen. Op grafeemniveau werden medeklinkers en klinkers vervangen en er werden medeklinkers ingevoegd. Bovendien werden lettergrepen vervangen door een andere lettergreep of letter. Wanneer alle lees pogingen werden meegenomen, bleek dat oudere leerlingen en leerlingen met een hoger AVI-niveau proportioneel gezien minder vaak woorden vervingen door orthografisch gelijke woorden, terwijl de proportie pseudowoorden, semantisch gelijke of semantisch passende woorden en invoegingen van een willekeurige medeklinker groter was. Wanneer alleen de laatste lees pogingen werden meegenomen, bleek daarnaast dat meisjes ten opzichte van jongens proportioneel gezien minder vaak woorden vervingen door pseudowoorden en semantisch gelijke of semantisch passende woorden.

Leerlingen gebruikten met name lexicale strategieën tijdens het lezen van een tekst. Absoluut gezien werden woorden het meest frequent correct gelezen, fout gelezen of het begin van het woord werd correct gelezen. Verleningen van klanken en spelstrategieën kwamen vrijwel niet voor. Leerlingen met een hoger AVI-niveau maakten proportioneel gezien meer leesfouten, omdat zij proportioneel gezien minder vaak het begin van het woord lazen. Wanneer alleen de laatste lees pogingen werden meegenomen, bleek dat meisjes proportioneel gezien minder woorden fout lazen dan jongens.

Uit dit kwantitatieve onderzoek is gebleken dat leesaccuratesse tijdens het hardop voorlezen van een Nederlandse tekst op beheersingsniveau niet in verband gebracht kon worden met leeftijd, geslacht en AVI-niveau. Daarnaast bleek dat de Nederlandse basisschoolleerlingen absoluut gezien voornamelijk lexicale strategieën gebruikten tijdens het hardop voorlezen van een tekst op beheersingsniveau. De sublexicale strategieën namen proportioneel gezien af naarmate leerlingen ouder werden en een hoger AVI-niveau beheersten. Ook de leesfouten die sublexicale strategieën suggereren namen proportioneel gezien af naarmate leerlingen ouder werden en een hoger AVI-niveau beheersten, terwijl leesfouten die lexicale strategieën suggereren proportioneel gezien toenamen.

Het advies voor vervolgonderzoek is om het JASMIN-CGN Corpus uit te breiden per leeftijdsgroep en daarnaast aan te vullen met leerlingen uit groep 3. Daarnaast wordt aangeraden het materiaal uit te bereiden met teksten op instructie- en frustratieniveau en woordlijsten met niet gerelateerde (pseudo)woorden. Bovendien kan leerlingen tijdsdruk opgelegd worden. Een longitudinaal onderzoek kan uitwijzen of het waarschijnlijk is dat verandering in leeftijd of AVI-niveau voorafgaat aan verandering in leesfouten of leesstrategieën.

## Bronnen

- Aarnoutse, C.A.J. (2004). *Ontwikkeling van beginnende geletterdheid*. Nijmegen: Thieme MediaCenter Nijmegen.
- Archer, N., & Bryant, P. (2001). Investigating the role of context in learning to read: A direct test of Goodman's model. *British Journal of Psychology*, 92(4), 579-591. <http://doi.org/10.1348/000712601162356>
- Au, K.H.-P. (1977). Analyzing Oral Reading Errors to Improve Instruction. *The Reading Teacher*, 31(1), 46-49. Geraadpleegd van <https://www.jstor.org/stable/20194454>
- Baarda, B., Bakker, E., Fischer, T., Julsing, M., Van Vianen, R., & Van der Hulst, M. (2017). *Basisboek methoden en technieken : kwantitatief praktijkgericht onderzoek op wetenschappelijke basis*. Groningen/Houten: Noordhoff Uitgevers.
- Balans. (2008). *Omrekening van AVI-oud naar AVI-nieuw*. Geraadpleegd van <http://www.de-edelsteen.nl/Portals/71/docs/Diversen/AVI%20omrekening.pdf>
- Benoit, A., Geudens, A., Irausquin, R., Koekebacker, E., van der Linden, S., van Loosbroek, ... Warnaar, J. (z.d.). *Gebruikswijzer en verantwoording: kim-versie Veilig leren lezen*. Tilburg: Uitgeverij Zwijsen B.V.
- Blonk, A., Hageman, J., Janssen, M., van den Hoogen, C., Kleijnen, R., Klein, B. ... Visser, J. (2006). *Dyslexie: Zorg van ons allemaal*. Amsterdam: Garant.
- Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer (Versie 6.0.52) [Computerprogramma]. Gedownload van <http://www.praat.org/>
- Bosman, A.M.T., & Van Orden, G.C (2003). Het fonologisch coherentiemodel voor lezen en spellen. *Pedagogische studiën*, 80(5), 391-406. Geraadpleegd van <http://pedagogischestudien.nl/download?type=document&identificer=616620>
- Bouwers, H., & van Goor, H. (1996). *Diagnostiek en behandeling van leesproblemen*. Nijkerk: Intro.
- Braams, T. (2007). *Dyslexie: een complex taalprobleem*. Amsterdam: Uitgeverij Boom.
- Chinn, C. (1933). Situated Actions During Reading Lessons: A Microanalysis of Oral Reading Error Episodes. *American Educational Research Journal*, 30(2), 361-392. <http://doi.org/10.3102/00028312030002361>
- Cito. (z.d.). *Woordenschat*. Geraadpleegd van <https://www.cito.nl/onderwijs/primair-onderwijs/lvs-toetsen/toetsen/woordenschat>
- Cleuren, L. (2009). *Elements of speech technology based reading assessment and intervention* (PhD Thesis). KU Leuven. Geraadpleegd van <https://lirias.kuleuven.be/retrieve/84719>
- Cleuren, L., Duchateau, J., & Sips, A. (2008). *Annotation Protocol for CHOREC*. Beschikbaar bij het instituut voor de Nederlandse taal: <http://hdl.handle.net/10032/tm-a2-j5>
- Cramer, D., & Howitt, D. (2004). *The SAGE Dictionary of Statistics*. <http://doi.org/10.4135/9780857020123>
- Cucchiari, C., Driesen, J., Van hamme, H., & Sanders, E. (Mei 2008). Recording Speech of Children, Non-Natives and Eldery People for HLT Applications: the JASMIN-CGN Corpus. Paper gepresenteerd op de *Language Resources and Evaluation Conference*. Geraadpleegd van <http://lands.let.ru.nl/literature/catia.2008.1.pdf>
- Cucchiari, C., Van Hamme, H., Driesen, J., & Sanders, E. (2008). *THE JASMIN-CGN CORPUS: Design, recording, transcription and structure of the corpus*. Beschikbaar bij het instituut voor de Nederlandse taal: <http://hdl.handle.net/10032/tm-a2-j7>
- D'Angelo, K., & Mahlios, M. (1983). Insertion and Omission Miscues of Good and Poor Readers. *The Reading Teacher*, 36(8), 778-782. Geraadpleegd van <https://www.jstor.org/stable/20198326>

- Danielsson, K. (2001). Beginning readers' sensitivity to different linguistic levels: An error and correction analysis at the lexical, syntactic, and semantic levels. *Reading and Writing, 14*(5), 395-421. <http://doi.org/10.1023/A:1011182004179>
- Eling, P.A.T.M., & Bosman, A.M.T. (1997). Lezen en schrijven. In H.F.M. Peters, R. Bastiaanse, J. Van Borsel, P.H.O. Dejonckere, K. Jansoniuss-Schultheiss, S.J. Van der Meulen, & B.J.E. Mondelaers (Eds.), *Taal. Deel 7 – Handboek Stem-, Spraak- en Taalpathologie* (pp. 131-155). Houten: Bohn Stafleu van Loghum.
- Field, A. (2009). *Discovering Statistics Using SPSS*. Londen: SAGE Publications Ltd.
- Goodman, K. S. (1965). A linguistic study of cues and miscues in reading. *Elementary English Journal, 42*(6), 639-643. Geraadpleegd van <http://www.jstor.org/stable/41387554>
- Goikoetxea, E. (2006). Reading errors in first- and second-grade readers of a shallow orthography: Evidence from Spanish. *British Journal of Educational Psychology, 76*(2), 333-350. <http://doi.org/10.1348/000709905X52490>
- Gubbels, A., Netten, A., & Verhoeven, L. (2017). *Vijftien jaar leesprestaties in Nederland: PIRLS-2016*. Nijmegen: Expertisecentrum Nederlands, Radboud Universiteit, Behavioural Science Institute.
- Guron, L.M., & Lundberg, I. (2004). Error Patterns in Word Reading Among Primary School Children: A Cross-orthographic Study. *Dyslexia, 10*(1), 44-60. <http://doi.org/10.1002/dys.260>
- Huizenga, H. (2016). *Aanvankelijk en technisch lezen*. Groningen/Houten: Noordhoff Uitgevers.
- IBM SPSS Statistics (Versie 21) [Computerprogramma]. Armonk, NY: IBM Corp.
- Inspectie van het Onderwijs. (2006). *Iedereen kan leren lezen*. Geraadpleegd van [http://masterplandyslexie.nl/public/files/documenten/Iedereen\\_kan\\_leren\\_lezen.pdf](http://masterplandyslexie.nl/public/files/documenten/Iedereen_kan_leren_lezen.pdf)
- Inspectie van het Onderwijs. (2018). *Onderwijsverslag 2016/2017*. Geraadpleegd van [https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2018/04/11/rapport-de-staat-van-het-onderwijs/108126\\_IvhO\\_StaatvanhetOnderwijs\\_TG.pdf](https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2018/04/11/rapport-de-staat-van-het-onderwijs/108126_IvhO_StaatvanhetOnderwijs_TG.pdf)
- Inspectie van het Onderwijs. (2019). *De Staat van het Primair Onderwijs 2019*. Geraadpleegd van <https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2019/04/10/deelrapport-primair-onderwijs/De+Staat+van+het+primair+onderwijs+2019.pdf>
- Jongen, I., & Krom, R. (2009a). *AVI-toets Leeskaart A en B*. Arnhem: Cito.
- Jongen, I., & Krom, R. (2009b). *Drie-Minuten-Toets (DMT)*. Arnhem: Cito.
- Kleijnen, R. (1996). *Strategieën van zwakke lezers en spellers in het voortgezet onderwijs*. Leiden: Swets & Zeitlinger Publishers.
- Laing, S.P. (2002). Miscue Analysis in School-Age Children. *American Journal of Speech-Language Pathology, 11*(4), 407-416. [http://doi.org/10.1044/1058-0360\(2002/044\)](http://doi.org/10.1044/1058-0360(2002/044))
- Lansink, N. (2009). *Cito Taal voor kleuters*. Arnhem: Cito.
- Leu, D.J. (1982). Oral Reading Error Analysis: A Critical Review of Research and Application. *Reading Research Quarterly, 17*(3), 420-437. <http://doi.org/10.2307/747528>
- Mommers, M.J.C., Verhoeven, L. and Van der Linden, S. (1990) *Veilig Leren Lezen*. Tilburg: Zwijzen.
- Mostow, J., Beck, J., Winter, S.V., Wang, S., & Tobin, B. (2002). Predicting Oral Reading Miscues. Paper gepresenteerd op de *International Conference on Spoken Language Processing, Colorado*. Geraadpleegd van <https://corescholar.libraries.wright.edu/cgi/viewcontent.cgi?article=1277&context=knoesis>



- Murnane, R., Sawhill, I., & Snow, C. (2012). Literacy Challenges for the Twenty-First Century: Introducing the Issue. *The Future of Children*, 22(2), 3-15. <http://doi.org/10.2307/23317408>
- Paternotte, A., & Oostewechel, N. (2015). *Dyslexie en leesproblemen*. Houten: Uitgeverij LannooCampus.
- Pronk-Boersma, M. (1992). *Logopedie voor onderwijsgeevenden*. Baarn: Uitgeverij H. Nelissen.
- Protopapas, A., & Skaloumbakas, C. (2007). Traditional and computerbased screening and diagnosis of reading disabilities in Greek. *Journal of Learning Disabilities*, 40(1), 15-36. <https://doi.org/10.1177/00222194070400010201>
- Ruissenaars, A.J.J.M., & Ghesquière, P. (2002). *Dyslexie en dyscalculie: ernstige problemen in het leren lezen en rekenen. Recente ontwikkelingen in onderkenning en aanpak*. Leuven: Acco.
- Rutter, M., Caspi, A., Fergusson, D., Horwood, L.J., Goodman, R., Maughan, B., ... Carroll, J. (2004). Sex Differences in Developmental Reading Disability. New Findings From 4 Epidemiological Studies. *JAMA*, 291(16), 2007-2012. <http://doi.org/10.1001/jama.291.16.2007>
- Sas, J. & Wierenga, C. (1998). *Leesmoeilijkheden. Achtergronden, analyse en behandeling*. Groningen: Wolters-Noordhoff.
- Schale, F.C. (1966). Changes in Oral Reading Errors at Elementary and Secondary Levels. *Academic Therapy Quarterly*, 1(4), 225-229. <http://doi.org/10.1177/105345126600100408>
- Schober, P., Boer, C., & Schwarte, L.A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, 126(5), 1763-1768. <http://doi.org/10.1213/ANE.0000000000002864>
- Seidenberg, M.S., & McClelland, J.L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review*, 96(4), 523-568. <http://doi.org/10.1037//0033-295X.96.4.523>
- Stichting Lezen en Schrijven. (z.d.). *Over laaggeletterdheid*. Geraadpleegd van <https://www.lezenenschrijven.nl/over-laaggeletterdheid/>
- Strik, H. (2019). *DART: Dutch ASR-based Reading Tutor*. Geraadpleegd van <http://hstriik.ruhosting.nl/dart/>
- Struiksma, A.J.C., van der Leij, A., & Vieijra, J.P.M. (2009). *Diagnostiek van technisch lezen en aanvankelijk spellen*. Amsterdam: VU Uitgeverij.
- Taalunie. (2008). *JASMIN Spraakcorpus (Version 1.0)* [Data set]. Beschikbaar bij het instituut voor de Nederlandse taal: <http://hdl.handle.net/10032/tm-a2-j7>
- TULE. (2008). *Kerdoel 4 – Technisch lezen*. Geraadpleegd van <http://tule.slo.nl/Nederlands/bestand/P-L04a.pdf>
- Van Casteren, W., Bendig-Jacobs, J., Wartenbergh-Cras, F., van Essen, M., & Kurver, B. (2017). *Differentiëren en differentiatievaardigheden in het primair onderwijs*. Geraadpleegd van <http://www.researchned.nl/wp-content/uploads/2017/12/Differenti%C3%ABren-en-differentiatievaardigheden-in-het-primair-onderwijs-.pdf>
- Van der Leij, A. (1998). *Leesproblemen: Beschrijving, verklaring en aanpak*. Rotterdam: Lemniscaat.
- Van Orden, G.C., Pennington, B.F., & Stone, G.O. (1990). Word Identification in Reading and the Promise of Subsymbolic Psycholinguistics. *Psychological Review*, 97(4), 488-522. <http://doi.org/10.1037/0033-295X.97.4.488>

- Verhoeven, L. (2002). Vroege onderkenning en aanpak van leesproblemen. In A.J.J.M. Ruijsenaars & P. Ghesquière (Eds.), *Dyslexie en dyscalculie: ernstige problemen in het leren lezen en rekenen. Recente ontwikkelingen in onderkenning en aanpak* (pp. 23-43). Leuven: Acco.
- Verhoeven, L., & Aarnoutse, C. (1999). *Tussendoelen beginnende geletterdheid. Een leerlijn voor groep 1 tot en met 3*. Nijmegen: Expertisecentrum Nederlands.
- Verhoeven, L., & van Leeuwe, J. (2003). Ontwikkeling van decodeervaardigheid in het basisonderwijs. *Pedagogische Studiën*, 80(4), 257-271. Geraadpleegd van <http://pedagogischestudien.nl/download?type=document&identificer=616546>
- Verhoeven, L., & Vermeer, A. (2001). *Taaltoets Alle Kinderen. Diagnostische toets voor de mondelinge vaardigheid Nederlands bij kinderen van groep 1 tot en met groep 4*. Arnhem: Cito.
- Vernooy, C.G.Th. (1993). *Leesproblemen voorkomen en effectief mee omgaan*. Hoewelaken: Christelijk Pedagogisch Studiecentrum.
- Weber, R-M. (1970). A Linguistic Analysis of First-Grade Reading Errors. *Reading Research Quarterly*, 5(3), 427-451. <http://doi.org/10.2307/747079>
- Wentink, H., & Verhoeven, L. (2008). *Protocol Leesproblemen en Dyslexie*. Nijmegen: Expertisecentrum Nederlands.

## Protocol annotatie leesstrategieën en leesfouten JASMIN-CGN Corpus

Het JASMIN-CGN (Jongeren, Anderstaligen en Senioren in Mens-machine Interactie – Corpus Gesproken Nederlands) Corpus (Taalunie, 2008) bevat spraak van groepen sprekers die niet vertegenwoordigd zijn in het CGN, zoals kinderen, anderstaligen en senioren. Per persoon bestaat ongeveer 50% van het materiaal uit voorgelezen spraakmateriaal en 50% uit onvoorbereide spraak, geproduceerd in mens-machine dialogen (Cucchiari et al., 2008). Het spraakmateriaal bevat een handmatig gemaakte orthografische transcriptie, een automatisch gegenereerde woordoplijning met fonemische transcriptie en een automatisch gegenereerde foneemoplijning met fonemische annotaties (Cucchiari et al., 2008).

Dit protocol is geschreven om de leesstrategieën en leesfouten te annoteren die voorkomen in het voorgelezen spraakmateriaal uit het JASMIN-CGN Corpus. Er zijn in het verleden vele annotatieprotocollen gemaakt voor voorgelezen spraak. Deze protocollen verschillen van elkaar waardoor het lastig is studies gebaseerd op deze protocollen met elkaar te vergelijken (Donald, 1982). Het protocol van de leesfouten is gebaseerd op eerdere protocollen zoals het protocol van CHOREC (Cleuren, Duchateau, & Sips, 2008) en classificaties in verschillende onderzoeken (Goodman, 1969; Au, 1977; Stemberger, 1985; Laing, 2002; Goikoetxea, 2006; Black et al., 2007; Cleuren, 2009). In appendix 1 staat een gedetailleerd overzicht met de verantwoording voor de foutenclassificatie.

### 1. PRAAT

Het annoteren van leesstrategieën en leesfouten kan met behulp van het programma PRAAT (Boersma & Weenink, Versie 6.0.52). In PRAAT wordt het PRAAT-script (zie bijlage 2) geopend en uitgevoerd. Het PRAAT-script moet op de computer in eenzelfde map staan als het spraakmateriaal. Het PRAAT-script genereert automatisch drie extra tiers. Tier 4 is een lege tier, waar plek is voor de originele tekst. In tier 5 worden de leesstrategieën geannoteerd en in tier 6 de leesfouten.

### 2. Tier 4: originele tekst

Neem in tier 4 de segmentgrenzen aan het begin en aan het eind van de zin over uit tier 1 (*Boundary > Add on tier 4*). Voeg vervolgens per zin de originele tekst toe, vanaf de hoofdletter tot aan de punt. Tier 4 kan aangepast worden door met de linkermuisknop te klikken in het betreffende segment. De inhoud kan door middel van de tekstbalk links boven worden aangepast.

### 3. Tier 5: Leesstrategieën

In tier 5 worden de leesstrategieën geannoteerd. De segmentverdeling is automatisch gegenereerd en identiek aan de verdeling in tier 1. Voeg segmentgrenzen toe indien woorden in tier 1 als samengesteld woord worden gepresenteerd, terwijl het in de oorspronkelijke tekst als twee woorden wordt gepresenteerd (bijvoorbeeld *samenwonen* i.p.v. *samen wonen* of *tenslotte* i.p.v. *ten slotte*). Verwijder segmentgrenzen indien woorden in tier 1 als twee losse woorden worden gepresenteerd, terwijl het in de oorspronkelijke tekst als samengesteld woord wordt gepresenteerd.

Tier 5 kan aangepast worden door met de linkermuisknop te klikken in het betreffende segment. De inhoud kan door middel van de tekstbalk links boven worden aangepast.

Methode:

- Vergelijk tier 1 en tier 4 en annotateer voor elk woord het juiste label in tabel 5. Indien een woord niet correct werd gelezen tijdens de eerste poging, krijgt elke afzonderlijke poging in tier 1 een label in tier 5.

### 3.1 Labels

Eén poging

- Een woord werd correct gelezen tijdens de eerste poging: g
- Een woord werd correct gelezen, maar één of meerdere klanken werden verlengd: vkg
- Een woord werd correct gelezen, maar één of meerdere klanken werden dusdanig verlengd dat het niet langer als correct beschouwd werd: vkf
- Een woord werd fout gelezen tijdens de eerste poging: f
- Een woord werd overgeslagen: 0 (maak hiervoor een nieuw segment waar het woord had moeten komen)
- Een woord werd gewisseld van plaats met een ander woord; het eerste woord krijgt een W. Dit kan in combinatie met een codering voor goed of fout, bijvoorbeeld g(W) of f(W).

Meer dan één poging voor één woord

*Laatste poging:*

- Een woord werd correct gelezen tijdens de laatste poging: g of vkg
- Een woord werd fout gelezen tijdens de laatste poging: vkf, f, 0, (W)

*Vorige pogingen:*

- Een woord werd correct gelezen tijdens de eerste poging: g
- Een woord werd correct gelezen, maar de klanken werden verlengd: vkg
- Een woord werd correct gelezen, maar de klanken werden dusdanig verlengd dat het niet langer als correct beschouwd werd: vkf
- Een woord werd fout gelezen tijdens de eerste poging: f
- Een woord werd overgeslagen: 0
- De spellingstrategie is niet duidelijk hoorbaar: \*s (wanneer het niet hoorbaar is hoe gespeld wordt)
- Een woord werd gewisseld van plaats met een ander woord; het eerste woord krijgt een W. Dit kan in combinatie met een codering voor goed of fout, bijvoorbeeld g(W) of f(W).
- De eerste letter of het eerste deel van het woord is correct gelezen: ag
- De eerste letter of het eerste deel van het woord is fout gelezen: af
- De laatste letter of het laatste deel van het woord is correct gelezen: alg
- De laatste letter of het laatste deel van het woord is fout gelezen: alf
- Een woord is correct gespeld (letters): lsg (bijv. b...a...l voor bal)
- Een woord is fout gespeld (letters): lsf (bijv. t...a...l voor bal)
- Een woord is correct gespeld (lettergrepen): ssg (bijv. voet...bal voor voetbal)
- Een woord is fout gespeld (lettergrepen): ssf (bijv. voet...pal voor voetbal)

NB.

- Bij een hoorbare pauze: u
- Bij spraak dat geen voorgelezen spraak is: xxx

#### 4. Tier 6: leesfouten

In deze tier worden de leesfouten geannoteerd. De segmentverdeling is identiek aan de verdeling in tier 1. Er wordt altijd een label toegekend in tier 6, behalve wanneer een woord correct werd gelezen tijdens de laatste poging ('g' of 'vkg').

Tier 6 kan worden aangepast door met de linkermuisknop te klikken in het betreffende segment. De inhoud kan door middel van de tekstbalk links boven worden aangepast.

#### Methode:

- Elke foute poging om een woord te lezen in tier 1 moet een label krijgen voor de leesfout die heeft plaatsgevonden. Het is mogelijk dat meerdere fouten optreden binnen één woord; meerdere labels kunnen dan toegekend worden. In dat geval worden de coderingen gescheiden door een '/'.  
• Als een bepaalde leesfout meer dan één keer voorkomt binnen hetzelfde woord dan wordt de codering voor die leesfout slechts één keer geannoteerd.  
• Het label '0' wordt gebruikt voor woorden die werden weggelaten.

#### Tip:

Aan de hand van de asterisks in tier 1, de orthografische transcriptie, wordt aangegeven welke orthografische fouten een persoon maakt, zie [corpus\\_orth-transcriptie.pdf](#), B. Hoofdregels B4. Dit is een indicatie waar de leesfouten zitten. Luister echter ook goed naar de Soundfile en vergelijk dit met de doelttekst. Hier kan ook een discrepantie in voorkomen die geannoteerd dient te worden.

#### 5.1 Verkeerde leesstrategie

##### 5.1.1 Decodering

- a. (foute of correcte) lettergreep-per-lettergreep decodering die niet werd gevolgd door een poging tot gesynthetiseerd lezen (bijv. voet...bal)
- b. (foute of correcte) letter-per-letter decodering die niet werd gevolgd door een poging tot gesynthetiseerd lezen (bijv. b...a...l)
- c. (foute of correcte) gedeeltelijke (bestaande uit (een combinatie van) enkele letters of lettergrepen) decodering die niet werd gevolgd door een poging tot gesynthetiseerd lezen (bijv. v...oe..t...bal)

##### 5.1.2. Directe woordherkenning

- e. een bestaand woord werd vervangen door een ander bestaand woord dat orthografisch gelijk is, waarbij maximaal één grafeem of -en werd weggelaten, vervangen of toegevoegd (bijv. staart/straat).
- f. een bestaand woord werd vervangen door een ander bestaand woord dat semantisch gelijk is (bijv. poes/kat) en/of semantisch passend is binnen het stuk tekst dat werd gelezen (bijv. kandelaar/vuur). Semantisch passende woorden zijn bijvoorbeeld synoniemen of meer algemene termen.
- g. een bestaand woord werd vervangen door een pseudowoord (bijv. straat/skraat)
- k. invoeging van een extra woord dat niet in de tekst staat en semantisch passend is binnen het stuk tekst dat werd gelezen (bijv. toevoeging van bruin bij kat)
- l. invoeging van een extra woord dat niet in de tekst staat en semantisch niet passend is binnen het stuk tekst dat werd gelezen (bijv. toevoeging van 'taart': het was zijn verjaardag/het was zijn verjaardagstaart)

m. een bestaand woord werd vervangen door een ander bestaand woord dat zowel orthografisch als semantisch niet gelijk is en niet semantisch passend is binnen het stuk tekst dat werd gelezen

## 5.2 Fouten in de volgorde

1. verandering in volgorde van aangrenzende klanken (bijv. doorn/droon of blik/bilk)
2. verandering in volgorde van klanken die van elkaar zijn gescheiden door ten minste één andere klank (bijv. slot/stol)
3. andere volgordefouten

## 5.3 Fouten door vervangen

### 5.3.1 Klinkers

4. vervanging van klinkers (bijv. bal/bel )
5. vervanging van korte klinkers naar lange klinkers of omgekeerd (bijv. man/maan) (opmerking: vervangen van /E/, /@/, /e/ behoren ook tot deze categorie)
6. onjuiste vervanging van (combinaties van) klinkers in leenwoorden (bijv. au in chauffeur of in toiletpapier)
7. uitspraak van tweeklanken (orthografisch weergegeven als een combinatie van twee verschillende klinkers) alsof het eigenlijk twee klinkers waren (bijv. au in auto, of omgekeerd, bijvoorbeeld het negeren van de diaeresis in poëzie)

### 5.3.2 Medeklinkers

9. spiegelen van medeklinkers: b/d/p/q/g, t/f, m/w
11. vervanging van medeklinkers die visueel gelijk zijn (bijv. c/G, m/n, k/h, t/l, l/k, r/z, v/w, y/j, h/b, f/s, y/ij, l/f, f/p, r/n, h/n)
12. vervanging van /N/ door /nx/, /nG/, /Nx/ of /NG/, vervanging van /Nk/ door /nk/ or /N/, vervanging van /S/ door /sj/
13. verkeerd lezen van leenwoorden:
  - vervanging van /S/ door /x/, /k/ of /G/ (bijv. chocolade, machine)
  - vervanging van /si/ door /ti/ (bijv. vakantie)
  - substitutie van /s/ door /k/ (bijv. circus)
  - substitutie van /x/ door /S/, /s/ of /k/ + /h/ (bijv. chaos)
14. vervanging van een originele medeklinker door een willekeurige andere medeklinker (bijv. knop/kjop)
15. vervanging van een conjugatie -t door een -n of vice versa (bijv. gaan/gaat)

### 5.3.3 Klinker-medeklinker

16. vervanging van een medeklinker door een klinker of vice versa (vaak zijn de letters visueel gelijk) (bijv. u/n, i/j, i/l)

### 5.3.4 Lettergrepen

17. vervanging van een lettergreep door een andere lettergreep of letter

## 5.4 Fouten door weglaten

### C.4.1 Klinkers

18. weglaten van een willekeurige klinker

### 5.4.2 Medeklinkers

20. weglaten van een meervoud -n (inclusief @) of -s (bijv. stoelen/stoel, stoel@ wordt wel goed gerekend), of deletie van een conjugatie -t of -n (inclusief @) (bijv. hij loopt/hij loop)

21. weglaten van een medeklinker in een medeklinkercluster (bijv. spin/pin)
22. weglaten van een n of g in ng of deletie van een n of k in nk (bijv. bank/ban)
23. weglaten van een willekeurige medeklinker (bijv. reis/rei)

#### 5.4.3 Lettergrepen

25. weglaten van het achtervoegsel -en of -er (bijv. kinderen/kinder, schoenenwinkel/schoenwinkel)
26. weglaten van een willekeurige lettergreep (bijv. Gerrit/Ger)

#### 5.4.4 Andere

28. het eind van het woord is niet volledig herhaald bij het corrigeren van een vorige poging (bijv. vuutbal... voet)
29. het begin van het woord is niet volledig herhaald bij het corrigeren van een vorige poging (bijv. voetbel ... bal)

### 5.5 Fouten door invoegen

#### C.5.1 Klinkers

30. invoegen van een willekeurige klinker (bijv. speelplaats/speleplaats)

#### 5.5.2 Medeklinkers

32. foutief invoegen van een meervouds -n (inclusief @) of -s (bijv. stoel/stoelen) of een conjugatie -t of -n (inclusief @) (bijv. ik loop/ik loopt)
33. foutief lezen van sch wanneer slechts één of twee van de drie grafemen zijn geschreven of onjuiste lezing van ch wanneer alleen c is geschreven
34. invoegen van een n voor een g of k of invoegen van een g of k na n, waardoor de fonemen /N/ of /Nk/ worden gelezen in plaats van een g, k of n
35. invoegen van een willekeurige medeklinker (bijv. poes/poest, gisteren/gisteren)

#### 5.5.3 Lettergrepen

38. invoegen van een willekeurige lettergreep

## 6. Opslaan

Wanneer alle tieren van een bestand zijn geannoteerd, moet het geheel worden opgeslagen.

- Klik op de menu 'file' in het PRAAT-schermbord
- Klik: 'Write TextGrid to Text file'
- Een nieuw scherm verschijnt, een bestandsnaam kan worden gekozen
- Klik op 'Ok'
- Het scherm kan nu worden gesloten.

## 7. Controleren

- Wanneer verschillende mensen binnen dezelfde dataset annoteren is het wel de bedoeling dat dit op dezelfde manier gebeurt. Om deze consistentie te waarborgen, dienen er verschillende zaken te worden gedaan.
- Iedereen dient zich aan het protocol te houden.
- Voordat de annoteerders individueel aan de slag kunnen, is het aan te raden om met elkaar de eerste uitingen te annoteren.
- Daarna gaan de annoteerders op individueel niveau de eerste tien uitingen annoteren. Deze annotaties dienen met elkaar te worden vergeleken.

- Het is praktisch wanneer alle annoteerders met elkaar in contact zijn. Wanneer er vragen zijn tijdens het annoteren, kan dit in overleg worden opgelost.
- Zorg ervoor dat de bestanden worden opgeslagen met dezelfde bestandsnaam.

## 8. Literatuur

- Au, K.H.-P. (1977). Analyzing Oral Reading Errors to Improve Instruction. *The Reading Teacher*, 31(1), 46-49. Geraadpleegd van <https://www.jstor.org/stable/20194454>
- Black, M., Tepperman, J., Lee, S., Price, P., & Narayanan, S. (2007). Automatic Detection and Classification of Disfluent Reading Miscues in Young Children's Speech for the Purpose of Assessment. Paper gepresenteerd op de *Annual Conference of the International Speech Communication Association*. Geraadpleegd van [https://sail.usc.edu/publications/files/black\\_interspeech\\_2007.pdf](https://sail.usc.edu/publications/files/black_interspeech_2007.pdf)
- Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer (Versie 6.0.52) [Computerprogramma]. Gedownload van <http://www.praat.org/>
- Cleuren, L., Duchateau, J., & Sips, A. (2008). *Annotation Protocol for CHOREC*. Leuven: Katholieke Universiteit Leuven.
- Cucchiari, C., Driesen, J., Van hamme, H., & Sanders, E. (Mei 2008). Recording Speech of Children, Non-Natives and Eldery People for HLT Applications: the JASMIN-CGN Corpus. Paper gepresenteerd op de *Language Resources and Evaluation Conference*. Geraadpleegd van <http://lands.let.ru.nl/literature/catia.2008.1.pdf>
- Donald, D.R. Analysis of children's oral reading errors: a current perspective. *Journal of Research in Reading*, 3(2), 106-114.
- Goikoetxea, E. (2006). Reading errors in first- and second-grade readers of a shallow orthography: Evidence from Spanish. *British Journal of Educational Psychology*, 76(2), 333-350. <http://doi.org/10.1348/000709905X52490>
- Goodman, K.S. (1969). Analysis of Oral Reading Miscues: Applied Psycholinguistics. *Reading Research Quarterly*, 5(1), 9-30. <http://doi.org/10.2307/747158>
- Laing, S.P. (2002). Miscue Analysis in School-Age Children. *American Journal of Speech-Language Pathology*, 11(4), 407-416. [http://doi.org/10.1044/1058-0360\(2002/044\)](http://doi.org/10.1044/1058-0360(2002/044))
- Stemberger, J. P. (1985). *The lexicon in a model of language production*. (Doctoral dissertation, University of California at San Diego.) New York: Garland



## Bijlage 2: PRAAT-script

```
Create Strings as file list... list *awd
numberOfFiles = Get number of strings
for ifile to numberOfFiles
  select Strings list
  filename$ = Get string... ifile
  basefilename$ = filename$ - ".awd"
  wavfile$ = basefilename$+".wav"

  tempfile$ = basefilename$ + ".TextGrid"
  targetfile$ = filename$ + ".tg"
  if not fileReadable(targetfile$)
    Open long sound file... 'wavfile$'
    if fileReadable(tempfile$)
      Read from file... 'tempfile$'
    else
      Read from file... 'filename$'
    endif
    select LongSound 'basefilename$'
    plus TextGrid 'basefilename$'
    Edit
    pause Klik continue wanneer je klaar bent. Tussentijds saven met ctrl-s!
    select TextGrid 'basefilename$'
    Write to short text file... 'targetfile$'
    plus LongSound 'basefilename$'
    Remove
    filedelete 'tempfile$'
  endif
endfor
select Strings list
Remove
```

Dit PRAAT-script is opgesteld door E.P. Sanders (onderzoeker bij Centre for Language Studies en Centre for Language and Speech Technology van de Radboud Universiteit Nijmegen).