

De ontwikkeling van woordbeelden bij beginnende lezers en spellers

A. M. T. BOSMAN en A. M. B. DE
GROOT *

Universiteit van Amsterdam

Samenvatting

In drie experimenten werden bij beginnende lezers de volgende vragen onderzocht: a) na hoeveel aanbiedingen van een woord is er een woordbeeld opgebouwd, b) hoe groot is de duurzaamheid van een representatie van een woord die door herhaald visueel aanbieden ervan opgebouwd wordt, c) is een door lezen opgebouwde representatie inzetbaar bij het spellen, en d) zijn de verschillen tussen goede en zwakke lezers/spellers kwalitatief, kwantitatief of beide. De resultaten geven geen antwoord op de vraag bij welke oefenfrequentie beginnende lezers een woordbeeld ontwikkelen, maar de resultaten van alle drie de experimenten tonen wel aan dat als gevolg van herhaalde aanbiedingen een bepaalde representatie wordt opgebouwd. De opgebouwde representatie bleek duurzaam, maar zij lijkt slechts beperkt inzetbaar te zijn voor spellen. Ten slotte bleek dat de gevonden verschillen tussen goede en zwakke lezers voornamelijk kwantitatief van aard zijn.

1 Inleiding

De eerste fase van de beginnende lezer van een alfabetisch schrift kenmerkt zich door een sequentiële verklanking van de schrifttekens in fonemen. In een taal waarin een sterke cor-

* Met dank aan M. Klønhammer en R. Snoek voor de hulp bij de voorbereiding en uitvoering van de experimenten en de kinderen en het onderwijzend personeel van de basisscholen 'De Flevo'. 'De Klimop', 'De Kraanvogel', 'Het Noorderlicht', 'De Voorhof', 'De Where' en 'De Zesde' uit Purmerend voor hun bereidwillige deelname. Tevens danken wij de reviewers voor de waardevolle opmerkingen bij een eerdere versie van dit artikel.

respondentie bestaat (zoals in het Fins) tussen de orthografie en fonologie ervan is deze methode altijd min of meer succesvol. Is de correspondentie echter veel minder groot (zoals in het Engels) of afwezig (zoals in het Chinees) dan kan de lezer niet altijd of helemaal niet op een systeem van grafeem-foneem correspondenties afgaan.

Voor de ervaren lezer blijkt dit echter nauwelijks een probleem te zijn, omdat deze in vele gevallen toch al niet verklankend, indirect leest, maar de lettersequentie onmiddellijk in zijn geheel (direct) herkent. Er wordt verondersteld dat de lezer door herhaalde confrontatie met een woord geleidelijk een 'woordbeeld' van het woord in kwestie vormt. Met de term 'woordbeeld' wordt verwezen naar een geheugenrepresentatie op een hoger niveau dan de representaties van grafemen en fonemen en die, zo wordt aangenomen, visueel-ruimtelijke informatie bevat over het erin gecodeerde woord. In plaats van de term 'woordbeeld' wordt, om de connotatie met iets visueels te vermijden, in de Nederlandse vakliteratuur ook wel gesproken van woordspecifieke kennis of woordunieke letterpatronen (Reitsma & Vinke, 1986). In navolging van de Engelse terminologie (visual or orthographic code), hebben wij hier gekozen voor de term woordbeeld.

Evidentie voor de veronderstelling dat geoefende lezers gebruik maken van woordbeelden werd aangedragen door onder meer Seidenberg, Waters, Barnes en Tanenhaus (1984). In dit onderzoek, aan volwassen proefpersonen, werden leestijden vergeleken van regelmatig en onregelmatig gespelde hoog- en laag-frequente Engelse woorden. Regelmatig gespelde woorden zijn woorden met een eenduidige, voorspelbare relatie tussen hun fonologische en orthografische vorm; bij onregelmatig gespelde woorden is deze relatie niet of minder eenduidig.

Uit het onderzoek van Seidenberg e.a. bleek dat de leestijden van laag-frequente onregelmatige woorden iets langer waren dan die van laag-frequente regelmatige. Daarentegen waren de leestijden van hoog-frequente

regelmatige en hoog-frequente onregelmatige woorden gelijk. Uit de interactie tussen gebruiksfrequentie en spellingregelmaat moet geconcludeerd worden dat geoefende lezers van het Engels bij het lezen van hoog-frequente woorden op een andere manier lezen dan door middel van grafeem-foneem omzetting. Als namelijk het leesproces uitsluitend via grafeem-foneem omzetting plaatsvindt dan zouden de leestijden van onregelmatige hoog-frequente woorden langer moeten zijn dan die van regelmatige hoog-frequente, omdat het proces van grafeem-foneem omzetting gestoord moet worden door elke onregelmatigheid die het op zijn weg tegenkomt. Dit blijkt echter niet het geval te zijn en Seidenberg e.a. concluderen dan ook dat bij geoefende lezers voor hoog-frequente woorden een woordbeeld is opgeslagen. In plaats van de indirecte leesweg hebben geoefende lezers bij hoog-frequente woorden blijkbaar ook de mogelijkheid om rechtstreeks contact te leggen tussen het geschreven woord en een mentale representatie in de vorm van een woordbeeld (directe leesweg).

Behalve dat het evidentie oplevert voor het bestaan van woordbeeldrepresentaties blijkt uit het onderzoek van Seidenberg e.a. tevens dat de gebruiksfrequentie van een woord van belang is voor de ontwikkeling van een woordbeeld. Dit blijkt uit het feit dat leestijden voor regelmatige laag-frequente woorden korter zijn dan voor onregelmatige, terwijl die voor de beide typen hoog-frequente woorden gelijk zijn.

Reitsma en Vinke (1986) onderzochten de relatie tussen frequentie en woordbeeldontwikkeling op een meer directe wijze dan Seidenberg e.a. (1984). Zij boden beginnende jonge lezers in een oefenfase een aantal woorden drie, negen of 18 keer aan. Na afloop van deze oefenfase werd de opleestijd van de geoefende woorden en van daarvan afgeleide 'pseudohomofonen' vastgesteld. Pseudohomofonen zijn lettersequenties die niet als woord bestaan, maar die een klankvorm hebben identiek aan die van bestaande woorden (bijvoorbeeld: wedstrijd-wetstret). Uit dit onderzoek bleek dat aanbiedingsfrequentie interacteerde met het type stimulus (geoefend woord vs pseudohomofoon). Bij drie of minder aanbiedingen bleek er geen verschil te zijn in opleestijd tussen geoefend woord en pseudohomofoon, maar bij negen en 18 aanbiedin-

gen was de opleestijd van geoefende woorden significant korter dan die van de corresponderende pseudohomofonen. Dit effect wijst erop dat er na drie confrontaties met een woord nog geen woordbeeldontwikkeling heeft plaatsgevonden, maar na negen aanbiedingen wel. Reitsma en Vinke vonden geen effect van stimulustype bij een groep kinderen uit het 'Speciaal Onderwijs', zelfs niet na het maximale aantal van 18 aanbiedingen. Bij deze kinderen lijkt er dan dus nog geen woordbeeldontwikkeling te hebben plaatsgevonden.

Behalve de vraag naar de functie van het woordbeeld bij het lezen is het tevens interessant om na te gaan wat de rol van het woordbeeld is bij het spellen. Waters, Bruck en Seidenberg (1985) concluderen uit hun onderzoek dat beginnende lezers (Grade 3) voor zowel lezen als spellen de informatie over grafeem-foneem correspondentie-regels gebruiken, en dat ze bij het spellen geen gebruik maken van het woordbeeld ("do not... 'read out' the orthographic form"). Uit het onderzoek van Waters e.a. en van Waters, Bruck en Malus-Abramowitz (1988) bleek tevens dat er geen kwalitatief verschil was tussen goede en zwakke lezers/spellers in de wijze waarop ze lezen en spellen. Volgens Kreiner en Gough (1990) spellen volwassenen soms volgens de fonologische route en soms door het woordbeeld uit te lezen.

Het feit dat de kinderen uit het onderzoek van Waters e.a. (1985) bij het spellen geen gebruik leken te maken van het woordbeeld is niet verwonderlijk, omdat de kinderen waarschijnlijk ook nog geen woordbeeld hadden opgebouwd. Ze bleken immers ook (nog) fonologisch te lezen. Dat volwassen lezers wel in staat zijn (Kreiner & Gough, 1990) het woordbeeld in te zetten bij het spellen, doet ons de vraag stellen in hoeverre jonge kinderen die wel een woordbeeld hebben opgebouwd dit ook daadwerkelijk gebruiken om te spellen.

De resultaten van Reitsma en Vinke (1986) waren aanleiding voor het uitvoeren van het hieronder beschreven onderzoek. Zoals hierboven reeds aangegeven vonden Reitsma en Vinke een pseudohomofoneffect bij negen aanbiedingen van een woord. De eerste vraag waarop hier geprobeerd werd een antwoord te krijgen was of het aantal aanbiedingen nodig voor de vorming van een woordbeeld dichter

bij drie of dichterbij negen aanbiedingen ligt. In hun onderzoek werden vier frequentieniveaus gebruikt (0, 3, 9, 18). In ons eerste experiment werden dezelfde frequentie-condities gebruikt met toevoeging van de niveaus 5 en 7. Een tweede vraag betrof de duurzaamheid van de opgebouwde representatie. Daartoe vond een week na de eerste meting een tweede meting plaats. Op de derde plaats wilden we door de afname van een woorddictee nagaan in hoeverre kinderen de opgebouwde representatie inzetten bij het spellen. Anders gezegd, wordt een woord dat vaker gelezen is (en waarvan mag worden aangenomen dat daarvan een vollediger representatie is opgebouwd) ook beter gespeld. Ten slotte wilden we nagaan of goede lezers zich kwalitatief en/of kwantitatief onderscheiden van zwakke lezers.

2 Experiment 1

In dit experiment werd bij basisschoolleerlingen het volgende onderzocht: 1) de invloed van visuele aanbiedingsfrequentie van een woord op de vorming van woordbeelden; 2) de invloed van visuele aanbiedingsfrequentie van een woord op spellingprestaties; 3) het effect van leesniveau op de snelheid waarmee een woordbeeld wordt gevormd; 4) de duurzaamheid van het woordbeeld bij lezen; 5) de duurzaamheid van spellingkennis.

2.1 Methode

Materiaal

Uit de 'Nieuwe streeflijst woordenschat voor 6-jarigen' van Kohnstamm, Schaerlaekens, De Vries, Akkerhuis en Frooninckx (1981) werden 15 concrete zelfstandige naamwoorden geselecteerd¹. De geselecteerde woorden, die in het vervolg 'oefenwoorden' heten, komen niet voor in de basisleesstof van Groep 3 van scholen die de methode "Veilig leren lezen" gebruiken².

Van deze 15 oefenwoorden werden 'pseudohomofonen' gevormd. Elk van de pseudohomofonen werd gevormd door in het corresponderende oefenwoord één of meer van de volgende vier veranderingen aan te brengen: de 'sch' werd vervangen door 'sg'; de 'au' door 'ou' en andersom, de 'ij' door 'ei' en omgekeerd, en de 'd' in een positie waarin deze

uitgesproken wordt als een /t/ door de 't'. De gemiddelde lengte van de oefenwoorden was 6.3 letters.

Naast de lijst met oefenwoorden werd een lijst (zoveel mogelijk gelijk aan de oefenwoorden qua lengte en letterclusters) met 15 woorden samengesteld die niet door de kinderen geoefend werden en verderop 'niet-geoefende woorden' worden genoemd. De gemiddelde lengte van de niet-geoefende woorden was 5.9 letters. Zie Bijlage A voor een volledig overzicht van het stimulusmateriaal.

Voorts werden bij elk oefenwoord 18 zinnen geconstrueerd. Deze zinnen bevatten elk één oefenwoord en bestonden voor het overige uitsluitend uit woorden die voorkomen in de deeltjes 1, 2 en 3 van de leesmethode "Veilig leren lezen". Op deze wijze ontstond er een verzameling van 270 (15 maal 18) zinnen. Deze verzameling zinnen vormde het experimentele oefenmateriaal. De zinnen waren zo samengesteld dat deze geen voorspelbare context vormden voor het oefenwoord. Uit deze verzameling van 270 zinnen werden op systematische wijze vijf oefenlijsten samengesteld, elk bestaande uit 126 zinnen. In elke oefenlijst kwamen telkens drie oefenwoorden driemaal voor (Frequentie 3), drie oefenwoorden vijfmaal (Frequentie 5), drie oefenwoorden zevenmaal (Frequentie 7), drie oefenwoorden negenmaal (Frequentie 9) en drie oefenwoorden 18 maal (Frequentie 18). Het resultaat was dat elk oefenwoord in elke frequentie voorkwam.

Opzet en procedure

In de oefenperiode (maand maart) kregen de kinderen drie opeenvolgende dagen (maandag, dinsdag en woensdag) één van de vijf oefenlijsten met 126 zinnen te lezen. Per dag lazen de kinderen dus 42 zinnen. De proefleider liet elk kind in een individuele zitting de betreffende lijst met zinnen hardop voorlezen.

Op de vierde dag (donderdag) werden de kinderen getest (eerste meting). Op de monitor van een Macintosh Plus werden de 15 oefenwoorden, de daarbij behorende pseudohomofonen en de 15 niet-geoefende woorden elk éénmaal aan de kinderen aangeboden. Hun werd gevraagd om elk van de aangeboden woorden zo snel mogelijk uit te spreken. Eén proefleider registreerde door middel van het indrukken van een knop op het toetsenbord de reactietijd van het kind op het aangeboden

woord. De reactie- of opleestijd was de tijd tussen het verschijnen van het woord op het scherm en het moment waarop het kind het woord helemaal uitgesproken had. Om te voorkomen dat de registratie van de opleestijden beïnvloed zou worden door de proefleider werd ervoor gezorgd dat deze het scherm niet kon zien. Een tweede proefleider keek mee naar de monitor en noteerde of het kind het woord correct opgelezen had. Er is gekozen voor deze registratie van reactietijden, en niet voor voice-key-registratie omdat dit laatste een bepaalde mate van discipline vereist die kinderen van deze leeftijd nog niet lijken te bezitten.

Om aan de taak te wennen kregen de kinderen eerst vijf woorden te lezen die ze al vaker in de les hadden gehad. Daarna werden in een pseudo-gerandomiseerde volgorde de experimentele woorden aangeboden. In totaal moesten de kinderen dus 50 woorden hardop voorlezen. Het experiment duurde gemiddeld 15 minuten.

Na afloop van het opleesexperiment nam een derde proefleider een schriftelijk woorddictee af, bestaande uit de 15 oefenwoorden. De proefleider las een zin voor waarin telkens één van de oefenwoorden voorkwam. Door herhaling van het oefenwoord werd duidelijk gemaakt welk woord uit de voorgelezen zin opgeschreven moest worden.

Precies één week later, zonder dat ze tussentijds enige extra oefening kregen in de 15 oefenwoorden, werden de kinderen opnieuw getest (tweede meting, ook op donderdag). De procedure was precies dezelfde als tijdens de eerste meting. Na deze meting werd opnieuw het dictee afgenomen.

Proefpersonen

Uit een populatie van 141 basisschoolleerlingen uit Groep 3 (gemiddelde leeftijd ten tijde van het experiment: 84 maanden), die allen onderwezen werden middels de leesmethode "Veilig leren lezen" van Caesar (1979), werden een groep goede en een groep zwakke technische lezers geselecteerd (gemiddelde leeftijd geselecteerde groep: 84 maanden). Drie weken voor de oefenfase en de daaropvolgende eerste meting werd een aantal tests afgenomen die van belang werden geacht bij de selectie van deze niveau-groepen. Naast een test voor technisch lezen en een spellingtest werden een non-verbale en een verbale intelligentie-test

afgenomen.

De indeling naar leesniveau vond plaats op grond van de score op de Eén-Minuut-Test van Caesar (1970). Kinderen die een andere taal dan het Nederlands als moedertaal hadden werden niet in de selectie opgenomen. De groep zwakke technische lezers ($n = 26$) bestond uit kinderen die een score hadden tussen 15 en 19 (gemiddelde score 17.0, sd 1.6) en de kinderen uit de groep goede technische lezers ($n = 24$) hadden een score tussen 29 en 36 (gemiddelde score 32.2, sd 2.4). De totale onderzoeksgroep bestond uit 50 kinderen. Kinderen met een score lager dan 15 of hoger dan 36 werden niet in de experimentele groep opgenomen, omdat van deze kinderen verondersteld kan worden dat ze een leesniveau hebben dat niet representatief is voor de beginnende lezer uit Groep 3.

De twee experimentele groepen onderscheiden zich niet alleen in niveau van technische leesvaardigheid, maar waren tevens significant verschillend van elkaar wat hun spellingprestaties betreft ($F_{1,48} = 15.2$, $p < .001$). Het spellingniveau werd bepaald aan de hand van het Woorddictee bij "Veilig leren lezen" van Mommers en Van Dongen (1986). De groep goede lezers had een gemiddelde score van 28.1 (sd 1.6) op het woorddictee, en de zwakke lezers 23.2 (sd 5.9). Om na te gaan of de twee leesgroepen op intelligentie van elkaar afweken werden de delen A, B, C en D van de Standard Progressive Matrices van Raven (1958), een non-verbale intelligentietest, en de vocabulaire-test uit de RAKIT (Bleichrodt, Drenth, Zaal & Resing, 1984) afgenomen. Er was geen significant verschil tussen de groepen goede en zwakke lezers op de Standard Progressive Matrices en een marginaal significant verschil op de vocabulaire-test ($F_{1,48} = 3.48$, $.05 < p < .10$). Mogelijke verschillen tussen de twee leesgroepen kunnen dus niet toegeschreven worden aan een verschil in intelligentie.

Van de hierboven beschreven 50 kinderen die deelnamen aan de eerste meting, konden er 15 om uiteenlopende redenen (o.a. een schoolreisje waarop niet was geanticipeerd) niet meer meedoen aan de tweede meting. Van de 35 kinderen die aan de tweede meting deelnamen waren er 17 goede en 18 zwakke technische lezers.

Om na te gaan in hoeverre er sprake was van selectieve uitval is een aantal gegevens van

de totale groep vergeleken met de groep van 35 kinderen die aan zowel de eerste als de tweede meting hebben meegedaan. Het verschil tussen de goede en zwakke lezers was ook bij deze groep significant ($F_{1,33} = 438.0, p < .001$). Ook op de spellingtest waren de groepen significant verschillend ($F_{1,33} = 11.58, p < .001$). Op de Raven en de vocabulaire test onderscheidde de groepen goede en zwakke lezers zich niet van elkaar. Deze gegevens laten zien dat deze groep niet afwijkt van de groep kinderen die aan de eerste meting deelnam. Anders gezegd, er is geen sprake geweest van selectieve uitval, zodat de resultaten uit de eerste en tweede meting met elkaar kunnen worden vergeleken.

2.2 Resultaten

Bij de bespreking van de resultaten zullen de gegevens van de eerste en de tweede meting geïntegreerd weergegeven worden, omdat deze vrijwel identiek zijn. Eerst worden de resultaten van het leesexperiment besproken en vervolgens van het spellingdictee. De afkortingen M1 en M2 geven aan dat de resultaten betrekking hebben op de eerste resp. de tweede meting.

Voorafgaande aan de analyse werden de reactietijden gebaseerd op foute reacties en extreme opleestijden verwijderd. Bij het geven van de reactie kon op twee manieren een fout ontstaan: ofwel de proefpersoon las niet wat er stond of de proefleider was bij de registratie van de reactie te snel of te traag. Een opleestijd werd als extreem lang beschouwd, wanneer deze langer was dan de langste opleestijd van de gemiddeld langzaamste proefpersoon. Een extreem korte opleestijd was een tijd korter dan de snelste opleestijd van de gemiddeld snelste proefpersoon.

Het gemiddelde percentage fouten in Meting 1 als gevolg van verkeerd oplezen door de proefpersoon bedroeg 12.2%, waarbij het aantal fouten van de groep goede lezers significant kleiner bleek te zijn dan dat van de zwakke (4.7% vs. 19.0%, $t = -5.73, df = 35.74, p < .001$). In Meting 2 was het gemiddelde percentage fouten 9.8%. Ook hier maakten goede lezers significant minder fouten dan zwakke (4.2% vs. 15.1%, $t = -4.35, df = 25.33, p < .001$). In Meting 1 werden in totaal acht extreme (vijf extreem korte en drie extreem lange; 0.3%) opleestijden uit de analyse verwijderd; in Meting 2 alleen twee

extreem lange. Door deze manier van databewerking kon één proefpersoon uit de eerste meting niet in de analyse worden opgenomen, omdat deze in één van de frequentie-condities geen gemiddelde waarde had. De nu volgende analyses van de eerste meting zijn gebaseerd op de resultaten van 49 proefpersonen en de analyses van de tweede meting op die van 35 proefpersonen.

Testfase

Uit een 2 (technisch leesniveau: goed vs. zwak) bij 2 (stimulustype: oefenwoorden vs. pseudohomofonen) bij 5 (oefenfrequentie: 3, 5, 7, 9, 18) variantie-analyse op de data van de eerste meting bleek dat het hoofdeffect van leesniveau significant was (M1: $F_{1,46} = 53.69, p < .001$; M2: $F_{1,33} = 36.23, p < .001$). De gemiddelde opleessnelheid in M1 en M2 voor de goede lezers bedroeg 2770 resp. 2581 msec. Voor de groep zwakke lezers was dit 5066 resp. 5110 msec. Het hoofdeffect van stimulustype was eveneens in beide metingen significant (M1: $F_{1,46} = 58.99, p < .001$; M2: $F_{1,33} = 28.58, p < .001$). De opleestijdverschillen tussen pseudohomofonen en woorden waren in elke frequentie-conditie significant. De gemiddelde opleessnelheid van de oefenwoorden in M1 en M2 bedroeg 3464 resp. 3440 msec. en van de pseudohomofonen 4373 resp. 4323 msec. Het hoofdeffect van frequentie was eveneens in beide metingen significant (M1: $F_{4,184} = 2.89, p < .05$; M2: $F_{4,132} = 3.69, p < .01$). Alleen in de tweede meting bleek dat de gemiddelde opleestijd in Frequentie 18 significant korter was dan in Frequentie 3. Dit werd uitsluitend veroorzaakt door de groep zwakke lezers (Newman-Keuls, $p < .01$). Er waren in de eerste meting geen significante interactie-effecten. In de tweede meting was er een significante interactie tussen leesniveau en oefenfrequentie ($F_{4,132} = 3.36, p < .05$). In de groep zwakke lezers was het hoofdeffect van frequentie significant ($F_{4,132} = 3.57, p < .01$), bij de goede lezers niet.

Om na te gaan of de gevonden resultaten een artefact waren van materiaalkeuze (i.e.: het voorkomen in de pseudohomofonen van de in het Nederlands niet-bestaande 'sg'-cluster in één lettergreep) werd een analyse uitgevoerd op de opleestijden van de woorden en pseudohomofonen uit de eerste meting met weglating van de 'sch' c.q. 'sg' stimuli³. Een 2 (stimulustype: oefenwoorden vs. pseudoho-

mofonen) bij 5 (oefenfrequentie: 3, 5, 7, 9, 18) ANOVA toonde aan dat ook nu het hoofdeffect van stimulistype significant was ($F_{1,44} = 7.1, p < .05$). Het frequentie-effect was echter nog slechts marginaal significant ($F_{4,176} = 2.2, .05 < p < .10$). Weglating van 33% van de responsen is hier zeer waarschijnlijk verantwoordelijk voor. Er was ook nu geen significant interactie-effect. Nu bleek echter dat het verschil tussen pseudohomofonen en woorden in Frequentie 3 niet significant was. Alleen in Frequentie 5 ($F_{1,44} = 5.6, p < .05$) was het verschil in opleestijd tussen woorden en pseudohomofonen significant, in de overige frequentie-condities niet. In Figuur 1a staan de resultaten van alle stimuli van de eerste meting weergegeven en in 1b die van de deelverzameling van stimuli (zonder opleestijden van 'sch'- en 'sg'-stimuli).

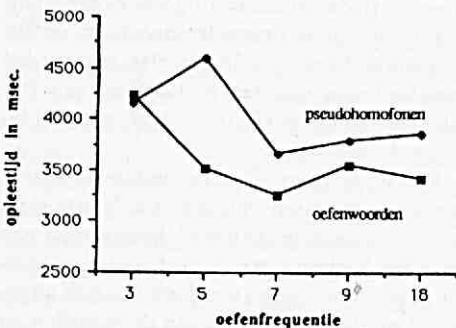
Uit een 2 (technisch leesniveau: goed vs. zwak) bij 6 (oefenfrequentie: 0, 3, 5, 7, 9, 18) variantie-analyse, waarin uitsluitend de oefenwoorden werden opgenomen (Frequentie 0 bevat de niet-geoefende woorden), bleek opnieuw in zowel de eerste als de tweede meting het hoofdeffect van leesniveau (M1: $F_{1,47} = 60.79, p < .001$; M2: $F_{1,33} = 47.68, p < .001$) en het hoofdeffect van frequentie (M1: $F_{5,235} = 22.29, p < .001$; M2: $F_{5,165} = 16.93, p < .001$) significant. De grootste afname van opleestijd in beide metingen vond plaats tussen 0 en 3 aanbiedingen (M1: 1543 msec., M2: 1229 msec.). De gemiddelde opleestijd in Frequentie 0 week in beide metingen significant af van alle andere frequenties (Newman-Keuls, $p < .01$). De gemiddelde opleestijd in Frequentie 18 was in beide metingen significant korter dan die van Frequentie 3

($p < .05$). Er was een significant interactie-effect tussen leesniveau en frequentie (M1: $F_{5,235} = 3.84, p < .01$; M2: $F_{5,165} = 3.11, p < .01$). Dit effect werd in beide metingen veroorzaakt door het feit dat de zwakke lezers, in met name de lage frequentie-condities, een grotere afname van de opleessnelheid lieten zien dan de goede lezers. Het totale opleestijdverschil (i.e.: gemiddelde opleestijd van Frequentie 0 - gemiddelde opleestijd van Frequentie 18) voor goede lezers in M1 was 1513 msec. (M2: 1194 msec.) en voor de zwakke 2980 msec. (M2: 3104 msec.).

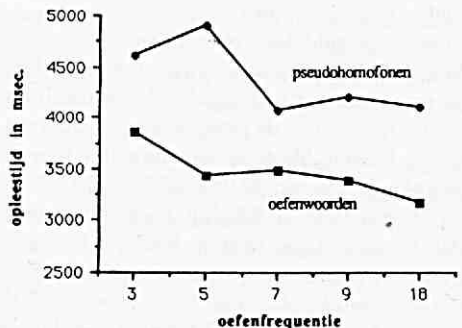
Uit een 2 (technisch leesniveau: goed vs. zwak) bij 5 (woordlengte: 4, 5, 6, 7, 8) variantie-analyse van de opleestijden uit de eerste meting bleek naast het hoofdeffect van leesniveau ook het hoofdeffect van woordlengte significant ($F_{4,180} = 10.43, p < .001$). De opleestijd van woorden met vier letters was significant korter dan die van woorden met vijf of meer letters (Newman-Keuls, $p < .01$). De gemiddelde opleestijd voor woorden met vijf letters was significant korter dan de woorden met acht letters (Newman-Keuls, $p < .01$). Het interactie-effect was niet significant.

Spellingresultaten

De prestaties op het spellingdictee werden op de volgende wijze beoordeeld. Van elk woord werd nagegaan of er een spelfout werd gemaakt in het zogenaamde kritische deel van het woord. Het kritische deel betrof hier de letters of lettercombinaties: 'd', 't', 'sch', 'au', 'ou', 'ei', 'ij' (kritische fout). Daarnaast werd gekeken of er op andere plaatsen dan de kritische spelfouten in het woord werden gemaakt (overige fouten).



Figuur 1a Opleestijden van alle oefenwoorden en pseudohomofonen naar oefenfrequentie van de totale groep kinderen uit de eerste meting van Experiment 1



Figuur 1b Opleestijden van de oefenwoorden en pseudohomofonen zonder 'sch'- en 'sg'-stimuli naar oefenfrequentie van de totale groep kinderen uit de eerste meting van Experiment 1

Een 2 (technisch leesniveau: goed vs. zwak) bij 2 (type fout: kritische vs. overige) bij 5 (oefenfrequentie: 3, 5, 7, 9, 18) variantie-analyse liet zien dat in beide metingen het hoofdeffect van leesniveau significant was (M1: $F_{1,45} = 16.31$, $p < .001$; M2: $F_{1,32} = 11.87$, $p < .01$). Het gemiddelde aantal spelfouten per woord voor goede lezers in M1 en M2 was .45 resp. .44 en voor de groep zwakke lezers in beide metingen .67. Het hoofdeffect van type fout was eveneens significant (M1: $F_{1,45} = 153.02$, $p < .001$; M2: $F_{1,32} = 97.18$, $p < .001$). Het gemiddelde aantal fouten dat gemaakt werd in het kritische deel van het woord was in M1 .81 en in M2 .80. Het gemiddelde aantal spelfouten in het overige deel bedroeg .32 in M1 en .31 in M2. Er bleek in beide metingen geen significant effect van frequentie (M1: $F_{4,180} = 1.00$, $p = .41$; M2: $F_{4,128} = .63$, $p = .64$). Alleen de eerste orde interactie tussen type fout en leesniveau in M1 bereikte significantie ($F_{1,45} = 4.30$, $p < .05$). Dit werd veroorzaakt door het feit dat de goede lezers significant minder overige fouten (.16 vs. .46) maakten dan de zwakke lezers ($F_{1,84} = 19.95$, $p < .001$). Het verschil tussen goede en zwakke lezers met betrekking tot aantal fouten in het kritische deel week niet van elkaar af (.75 resp. .87). In de tweede meting was het interactie-effect tussen type fout en leesniveau slechts marginaal significant ($F_{4,128} = 2.31$, $.05 < p < .10$). Wel bleek ook nu weer dat het verschil tussen goede en zwakke lezers met betrekking tot overige spelfouten significant was ($F_{1,59} = 15.18$, $p < .001$), maar niet voor kritische fouten.

Een 2 (technisch leesniveau: goed vs. zwak) bij 2 (type fout: kritische vs. overige) bij 3 (woordklasse: sch vs. au-ou vs. ei-ij) variantie-analyse gaf aan dat naast de significante hoofdeffecten van leesniveau en type fout ook het hoofdeffect van woordklasse significant was (M1: $F_{2,90} = 114.32$, $p < .001$; M2: $F_{2,64} = 70.10$, $p < .001$). De minste fouten werden gemaakt in de woorden waarin een 'sch' voorkwam (M1: .22; M2: .25). Het gemiddelde aantal fouten in woorden met 'au' of 'ou' bedroeg in beide metingen .65. Woorden waarin een 'ei' of 'ij' voorkwam gaven de meeste fouten te zien (M1: .83; M2: .78). In beide metingen weken alle drie de gemiddelden significant van elkaar af. (Newman-Keuls, $p < .01$). Tevens waren in beide metingen alle eerste orde interacties significant.

Hier zal alleen de relevante interactie tussen leesniveau en woordklasse besproken worden. De F-waarde van het interactie-effect tussen leesniveau en woordklasse was voor M1: $F_{2,90} = 7.04$, $p < .01$ en voor M2: $F_{2,64} = 4.26$, $p < .05$. Voor de zwakke lezers geldt dat zij woorden met een 'sch' erin even makkelijk vinden als de goede lezers, maar er worden door zwakke lezers meer spelfouten gemaakt in woorden met 'au'-ou' en 'ij'-ei' dan door de goede lezers. In beide metingen was het effect van leesniveau op woorden waarin een 'sch' voorkomt niet significant. Het effect van leesniveau op woorden die een 'au'-ou' hebben is wel significant. (M1: $F_{1,109} = 13.95$, $p < .001$; M2: $F_{1,71} = 10.56$, $p < .001$); evenals het effect van leesniveau op woorden met 'ei'-ij' (M1: $F_{1,109} = 22.62$, $p < .001$; M2: $F_{1,71} = 15.72$, $p < .001$). In beide gevallen maken goede lezers minder spelfouten dan zwakke.

2.3 Discussie

De hierboven gerapporteerde resultaten geven geen antwoord op de vraag na hoeveel aanbiedingen van een woord er woordbeeldontwikkeling optreedt. Anders dan bij Reitsma en Vinke (1986), die pas bij negen aanbiedingen van een woord een opleestijdverschil tussen woorden en pseudohomofonen vonden (wat erop duidt dat er zich dan een woordbeeld of woordspecifieke kennis heeft gevormd), was er in ons onderzoek, wanneer alle stimuli in de analyses werden opgenomen, bij een oefenfrequentie van drie al sprake van een opleestijdverschil. Dit verschil werd niet groter met toenemende aanbiedingsfrequentie. Uit de analyse op de deelverzameling van de stimuli bleek dat er geen verschillen waren tussen de opleestijden van de pseudohomofonen en de woorden in de frequentie-condities (behalve bij Frequentie 5). In beide gevallen is echter het effect van stimulustype en frequentie aanwezig, en is er geen interactie tussen beide.

Een mogelijke verklaring voor het door ons gevonden resultaat is dat er in minder dan drie aanbiedingen reeds een woordbeeld is opgebouwd, en dat dit woordbeeld zich bij toenemende aanbiedingsfrequentie tot 18 niet wijzigt. De verwachte interactie zou dan alleen aantoonbaar zijn bij een aanbiedingsfrequentie lager dan drie (in Experiment 2 zal deze mogelijke verklaring onderzocht worden).

Op de vraag of de opgebouwde representa-

tie van de oefenwoorden relatief duurzaam is kan bevestigend geantwoord worden. Dit blijkt uit het feit dat de analyses van Meting 1 en 2 vrijwel identiek zijn.

De spellingresultaten geven aan dat de frequentie waarmee een woord eerder gelezen is geen effect heeft op het correct spellen ervan. De opgebouwde representatie van een woord dat 18 maal gelezen is, is dus bijvoorbeeld niet kwalitatief beter voor het spellen dan die van een woord dat drie keer is gelezen. Tevens bleek dat de spellingprestaties uit Meting 1 en 2 vrijwel identiek waren. Dit gold voor zowel goede als zwakke lezers. Goede lezers maken minder spelfouten dan zwakke, maar dit viel te verwachten, omdat de groep goede lezers tevens de groep goede spellers is.

Zowel de lees- als de spellingresultaten duiden erop dat er geen kwalitatieve verschillen tussen goede en zwakke lezers zijn, slechts kwantitatieve. Goede lezers lezen bijna tweemaal zo snel als zwakke en zowel zwakke als goede lezers maken meer kritische spelfouten dan overige. Hoewel er een interactie optrad tussen leesniveau en type spelfout, is dit geen aanwijzing voor een kwalitatief verschil, omdat het maken van een fonologische fout in de lettercluster 'sch' heel onwaarschijnlijk is, terwijl die in de clusters 'au', 'ou', 'ei' en 'ij' wel voor de hand ligt. Een fonologisch juiste fout ('sg') in een 'sch'-lettercluster is een orthografisch niet bestaande lettercombinatie in het Nederlands. De kans dat kinderen deze woorden fout spellen is niet zo erg groot. In de andere gevallen is dat veel waarschijnlijker. Het interactie-effect werd voornamelijk verklaard door het feit dat binnen de woordklasse 'sch' geen verschillen tussen groepen voorkomen.

Zoals hiervoor al werd gesuggereerd is het mogelijk dat het woordbeeld in minder dan drie confrontaties al is opgebouwd. Om dit te onderzoeken zal in Experiment 2 nagegaan worden of de interactie tussen stimulustype en aanbiedingsfrequentie misschien wel optreedt tussen nul en drie aanbiedingen. Tevens zal nagegaan worden waar precies tussen nul en drie aanbiedingen de grootste afname van de opleestijd tot stand komt. Om dit vast te kunnen stellen zal in Experiment 2 de opleestijd gemeten worden van woorden met aanbiedingsfrequenties 0, 1, 2, 3, 6 en 9.

Om meer informatie over het leren lezen

van woorden te verkrijgen werd in Experiment 2 niet alleen de testfase, maar ook de oefenfase op een computer afgenomen, waarbij van alle gelezen woorden de opleestijden werden geregistreerd. Daarbij werd afgezien van aanbieding in een zin, omdat dit de registratie van de opleestijd van het oefenwoord zou bemoeilijken. Omdat immers de oefenzinnen uit Experiment 1 geen voorspelbare context vormden voor het oefenwoord wordt daarmee het lezen ervan dus een taak die vergelijkbaar is met het oplesen van losse woorden.

In Experiment 2 zal geen onderscheid gemaakt worden tussen goede en zwakke lezers, omdat hun lees- en spelgedrag geen kwalitatieve verschillen vertoonden.

3 Experiment 2

3.1 Methode

Materiaal

Opnieuw was de 'Nieuwe streeflijst woorden-schat' van Kohnstamm e.a. (1981) de bron voor het stimulusmateriaal. Er werden 12 zelfstandige naamwoorden als oefenwoorden geselecteerd. Ook deze woorden maken geen deel uit van de basisleesstof "Veilig leren lezen" van groep 3. De gemiddelde lengte van de oefenwoorden bedroeg 7.0 letters. Op dezelfde wijze als in Experiment 1 werden van deze woorden pseudohomofonen gevormd. Zowel de oefenwoorden als de daarvan afgeleide pseudohomofonen staan in Bijlage B.

Het oefenmateriaal bestond uit zes verschillende lijsten met oefenwoorden. Deze lijsten waren zo samengesteld dat in elke lijst twee woorden nooit voorkwamen (Frequentie 0), twee woorden éénmaal (Frequentie 1), twee woorden tweemaal (Frequentie 2), twee woorden driemaal (Frequentie 3), twee woorden zesmaal (Frequentie 6) en twee woorden negenmaal (Frequentie 9). Het resultaat was dat elk oefenwoord in elke frequentie voorkwam. Elke oefenlijst bevatte 42 woorden. Het testmateriaal (24 woorden) bestond uit een lijst met alle 12 oefenwoorden en de daarvan afgeleide pseudohomofonen. De verdeling van oefen- en testwoorden binnen een lijst was niet random, maar zodanig dat dezelfde woorden binnen de lijst op evenredige afstand van elkaar stonden, zonder dat er een voorspelbare volgorde ontstond. Het totaal aantal

experimentele stimuli bedroeg 66 (42 oefen- en 24 teststimuli).

Opzet en procedure

De afname van dit experiment (maand mei) vond net als het testgedeelte van Experiment 1 plaats op een Macintosh Plus computer. In dit geval echter werden zowel de oefen- als de testfase op de computer uitgevoerd. De oefen- en testfase vonden zonder onderbreking plaats en voor de kinderen was de overgang tussen de twee fasen niet merkbaar. Van alle aangeboden woorden werd door één proefleider de reactietijd geregistreerd. Een andere proefleider beoordeelde of het kind het woord juist had opgelezen. De sessie begon met vier bekende woorden om aan de taak te wennen. Net als in Experiment 1 werd ook hier, na afloop van het computerdeel, een dictee afgenomen, maar er vond geen tweede meting plaats (zie voor details het methode gedeelte van Experiment 1). Het experiment duurde gemiddeld tien minuten.

Proefpersonen

Uit dezelfde populatie als die van Experiment 1 werden 36 nieuwe kinderen geselecteerd (gemiddelde leeftijd ten tijde van het experiment: 88 maanden). Deze kinderen hadden een middelmatige score op de test voor technisch lezen (23.4, sd 2.6), waarbij de onder- en bovengrens precies lag tussen de scores van de groepen goede en zwakke lezers uit Experiment 1 (ondergrens 20, bovengrens 28). Ook hier werden kinderen met een andere moedertaal dan het Nederlands niet opgenomen in de onderzoeksgroep. In deze geselecteerde onderzoeksgroep bleek het aantal jongens aanzienlijk kleiner te zijn dan het aantal meisjes (12 resp. 24). Om na te gaan of de groep jongens op één of meer variabelen significant van de groep meisjes afweek zijn er t-tests uitgevoerd op de gemiddelde scores van de afgenomen tests (technisch lezen, spelling, vocabulaire en de Raven). De groep jongens bleek op geen enkele test significant af te wijken van de groep meisjes, zodat aangenomen mag worden dat de ongelijke verdeling van de variabele sekse niet problematisch is voor het experiment⁴.

3.2 Resultaten

De bespreking van de resultaten valt uiteen in de onderdelen oefenfase, testfase en spelling-

resultaten. Op dezelfde wijze als in Experiment 1 werden de extreme opleestijden verwijderd. Het aantal extremen bedroeg acht (vier extreem korte en vier extreem lange; 0.3%).

De vervanging van de foute responsen in de testfase was in dit experiment anders dan in Experiment 1. De gemiddelde opleestijd van woorden uit de oefenfase was gebaseerd op het gemiddelde van alle goede responsen. De opleestijden van de foute responsen van de verschillende frequentie-condities uit de testfase werden vervangen door de gemiddelde opleestijd van de proefpersoon van de betreffende frequentie uit de oefenfase. Het gemiddelde percentage fout in de testfase als gevolg van verkeerd oplezen door de proefpersoon bedroeg 8.3%.

Oefenfase

Uit een variantie-analyse op de gemiddelde opleestijden met herhaalde metingen op de variabele frequentie (frequenties: 0, 1, 2, 3, 4, 5, 6, 7, 8) bleek dat het frequentie-effect significant was ($F_{8,280} = 56.21, p < .001$). Opleestijden van woorden in Frequentie 0 waren significant langer dan die in alle andere frequentie-condities (Scheffé F-test, $p < .01$). Opleestijden van woorden in Frequentie 1 waren significant langer dan die in de frequentie-condities 5, 6, 7 en 8. De grootste afname van opleestijd trad op tussen Frequentie 0 en Frequentie 1 (1866 msec.).

Testfase

Een 2 (stimulustype: oefenwoorden vs. pseudohomofonen) bij 6 (oefenfrequentie: 0, 1, 2, 3, 6, 9,) variantie-analyse liet zien dat beide hoofdeffecten significant waren. Het hoofdeffect van stimulustype was $F_{1,32} = 19.39, p < .001$. De gemiddelde opleestijd van oefenwoorden bedroeg 2897 en die van pseudohomofonen 3712 msec.. Het hoofdeffect van oefenfrequentie was $F_{5,160} = 6.34, p < .001$. De opleestijd van woorden in Frequentie 0 was significant langer dan in de andere condities (Newman-Keuls, $p < .01$). Ook nu bleek de grootste afname van opleestijd te liggen tussen de Frequenties 0 en 1 (916 msec.).

Het interactie-effect was niet significant ($F_{5,160} = 1.82, p = .11$), maar de verschillen tussen de opleestijden van oefenwoorden en pseudohomofonen in de Frequenties 1, 3, 6 en 9 bleken wel significant ($p < .05$). Het verschil tussen de opleestijden van oefenwoorden en

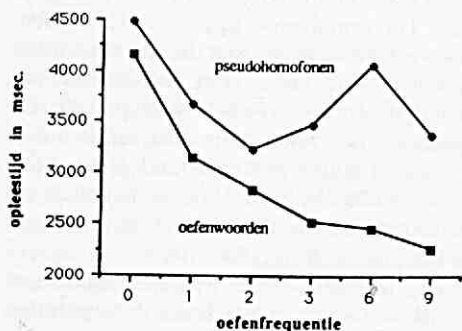
pseudohomofonen in de Frequenties 0 en 2 echter niet (de opleestijd van woorden in Frequentie 0 was 4161 msec. en van pseudohomofonen 4485 msec.)

Ook nu werd net als in Experiment 1 een analyse uitgevoerd op een deelverzameling van het totaal aantal responsen (de opleestijden gebaseerd op de woorden waarin een 'sch' en pseudohomofonen waarin een 'sg' voorkwam werden verwijderd). Uit de analyse van de opleestijden waarbij stimulustype (woorden vs pseudohomofonen) en oefenfrequentie (0, 1, 2, 3, 6, 9) werden getest, bleek dat het hoofdeffect van frequentie significant was ($F_{5,160} = 5.8, p < .001$) en dat van stimulustype marginaal ($F_{1,32} = 3.8, p < .059$). Ook nu was er geen significant interactie-effect. Nu bleek het opleesverschil tussen woorden en pseudohomofonen alleen nog maar significant te zijn in Frequentie 9 ($F_{1,32} = 5.8, p < .05$). De grootste afname van opleestijd lag overeenkomstig de eerder gevonden resultaten tussen de eerste en de tweede aanbieding (924 msec.) De resultaten zijn weergegeven in Figuur 2a en 2b.

Uit een variantie-analyse op de opleestijden met woordlengte als onafhankelijke variabele bleek dat het effect van woordlengte, net als in Experiment 1, significant was ($F_{4,140} = 7.15, p < .001$). De opleestijden van woorden met vier letters was significant korter dan van woorden met zes of meer letters (Scheffé F-test, $p < .05$).

Spellingresultaten

Een 2 (type fout: kritische vs. overige) bij 6 (oefenfrequentie: 0, 1, 2, 3, 6, 9) variantie-



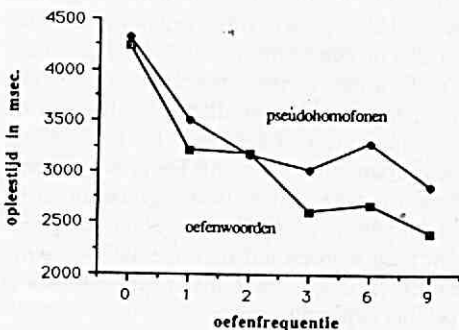
Figuur 2a Opleestijden van alle oefenwoorden en pseudohomofonen naar oefenfrequentie van de groep kinderen uit Experiment 2

analyse liet zien dat het hoofdeffect van type fout significant was ($F_{1,35} = 199.95, p < .001$). De gemiddelde fout in het kritische deel bedroeg 1.12 en in het overige .32. Het hoofdeffect van frequentie was niet significant. Er was wel een significant interactie tussen stimulustype en frequentie ($F_{5,175} = 3.11, p < .05$). Dit werd veroorzaakt door het feit dat het effect van oefenfrequentie op het niveau kritische fout wel significant was ($F_{5,175} = 2.28, p < .01$) maar niet op overige fout.

Een 2 (type fout: kritische vs. overige) bij 3 (woordklasse: sch vs. au-au vs. ei-ij) variantie-analyse liet zien dat zowel het hoofdeffect van type fout ($F_{1,35} = 170.97, p < .001$) als het hoofdeffect van woordklasse ($F_{2,70} = 46.29, p < .001$) significant was. Er werden significant minder spelfouten gemaakt in woorden met 'sch' (.52) dan in woorden met 'au'-'ou' (.65) of woorden met 'ei'-'ij' (1.01; Newman-Keuls, $p < .05$). Tevens was het verschil significant tussen woorden met 'au'-'ou' en woorden met 'ei'-'ij' (Newman-Keuls, $p < .01$). Het interactie-effect tussen type fout en woordklasse was eveneens significant ($F_{2,70} = 15.98, p < .001$). In vergelijking met de 'sch' en de 'au'-'ou' woorden zijn de woorden met 'ei'-'ij' wat hun kritisch deel betreft onevenredig veel moeilijker dan wat hun overig deel betreft.

3.3 Conclusie

Hoewel in Experiment 2 opnieuw geen interactie tussen stimulustype en aanbiedingsfrequentie werd gevonden laten de resultaten opnieuw wel een pseudohomofon-effect zien. Oefenwoorden worden immers sneller opgelezen dan pseudohomofonen. De inter-



Figuur 2b Opleestijden van de oefenwoorden en pseudohomofonen zonder 'sch'- en 'sg'-stimuli naar oefenfrequentie van de groep kinderen uit Experiment 2

pretatie van de resultaten is echter niet onproblematisch, omdat uit de analyse van de opleestijden van de totale stimulusverzameling is gebleken dat het pseudohomofoon-effect al optrad na de eerste aanbieding, terwijl uit de analyse op de deelverzameling van stimuli (zonder 'sch'- en 'sg'-stimuli) afgeleid kon worden dat dit pas bij negen aanbiedingen optrad. Dit laatste is een bevestiging van het resultaat van Reitsma en Vinke (1986); ook zij vonden een pseudohomofoon-effect bij negen aanbiedingen. Dat wellicht aan dit laatste resultaat meer waarde gehecht moet worden wordt ondersteund door het volgende. De analyse op de totale stimulusverzameling (inclusief woorden en pseudohomofonen met 'sch'- resp. 'sg'-clusters) liet een weliswaar niet-significant, maar toch substantieel verschil zien tussen de opleestijd van pseudohomofonen en woorden bij de eerste aanbieding (woorden die niet eerder door de kinderen gezien waren). Dit verschil was niet meer aantoonbaar in de analyse op de deelverzameling. Dit laatste resultaat viel te verwachten, omdat het voor kinderen bij de eerste lezing van een woord niet uit mag maken hoe dit woord geschreven is (wedstrijd of wetstrijd). Hieruit kan afgeleid worden dat door weglating van de 'sch' en 'sg'-stimuli er een valider meting van de effecten heeft plaatsgevonden.

De grootste afname van de opleestijd blijkt tussen nul en één aanbieding van een woord te liggen. Dit komt eveneens overeen met en nuanceert de resultaten van Experiment 1, waar geconstateerd werd dat de grootste afname plaatsvond tussen nul en drie aanbiedingen.

De analyses van de spellingprestaties komen voor het grootste deel overeen met die van Experiment 1. Er blijkt opnieuw geen effect van aantal malen lezen op het correct spellen van woorden te zijn. Maar in tegenstelling tot Experiment 1 werd hier een interactie gevonden tussen aanbiedingsfrequentie en type fout: er is wel een effect van aanbiedingsfrequentie op het gemiddeld aantal kritische fouten, maar niet op het aantal overige fouten. Dit duidt aan dat deze groep beginnende lezers, met een middelmatige score op de leestest, tijdens de oefenfase een representatie heeft opgebouwd die minstens voor een deel bij het spellen werd ingezet. Een punt dat hier wellicht van belang is, is het feit dat deze groep

kinderen drie maanden langer onderwijs had genoten dan de kinderen uit Experiment 1. Mogelijkerwijs spelen algemene rijpingsfactoren die samenhangen met schoolse vaardigheden een rol.

Om de bovenstaande conclusie te onderbouwen werd Experiment 3 uitgevoerd. Het mogelijk contaminerend effect van stimulusmateriaal werd beperkt door bij de samenstelling van de pseudohomofonen geen 'sg'-clusters meer op te nemen. Tevens werd het aantal aanbiedingen van een woord sterk vergroot (maximale oefenfrequentie 30) om daarmee mogelijkerwijs alsnog een interactie tussen stimulustype en oefenfrequentie te forceren. Naar aanleiding van de positieve ervaringen (met name het goede concentratievermogen) met jonge kinderen in Experiment 2 (88 maanden) is besloten om nu gebruik te maken van een voice key om opleestijden te registreren. Dit zou in ieder geval een reductie van de variantie tot gevolg moeten hebben. Net als in Experiment 1 en 2 waren we opnieuw ook geïnteresseerd in de invloed van oefenfrequentie van lezen op spellingprestaties.

4 Experiment 3

4.1 Methode

Materiaal

Uit de 'Nieuwe streefwoordenlijst woordenschat' van Kohnstamm e.a. (1981) werden 20 woorden geselecteerd (12 zelfstandige naamwoorden, drie werkwoorden, vier bijwoorden en één bijvoeglijk naamwoord). Deze woorden worden in het vervolg 'oefenwoorden' genoemd en maken geen deel uit van de basisleesstof uit "Veilig leren lezen" van groep 3. Van deze oefenwoorden werden net als in Experiment 1 en 2 pseudohomofonen gevormd door één of meer letters te wijzigen, dat wil zeggen: een 'ch' door een 'g', een 'ij' door een 'ei' of omgekeerd, een 'ou' door een 'au' of omgekeerd, en een 'd' op het eind van een woord waar deze uitgesproken wordt als een /t/ door een 't'. De gemiddelde lengte van de oefenwoorden bedroeg 6.0 letters. Een volledig overzicht van het gebruikte stimulusmateriaal staat in Bijlage C.

Het oefenmateriaal bestond uit vier verschillende lijsten, die zo waren samengesteld

dat in elke lijst vijf woorden niet voorkwamen (Frequentie 0), vijf woorden éénmaal (Frequentie 1), vijf woorden tweemaal (Frequentie 2) en vijf woorden 30 maal (Frequentie 30). Het resultaat was dat elk woord in elke frequentie voorkwam. Elke oefenlijst bestond uit 165 woorden. Het testmateriaal (totaal 40 stimuli) bestond uit alle 20 oefenwoorden en de daarbij behorende pseudohomofonen. De verdeling van oefen- en testwoorden binnen een lijst was niet random, maar zodanig dat dezelfde woorden binnen de lijst op evenredige afstand van elkaar stonden, zonder dat er een voorspelbare volgorde ontstond. Het totaal aantal experimentele stimuli bedroeg dus 205 (165 oefen- en 40 teststimuli).

Opzet en procedure

Ook dit experiment (maand mei) werd afgenomen op een Macintosh Plus computer. De oefenwoorden werden één voor één op het beeldscherm aangeboden en aan de kinderen werd gevraagd om deze zo snel mogelijk voor te lezen. Zonder merkbare onderbreking werden direct erna de testwoorden aangeboden. De opleestijden werden geregistreerd met behulp van een voice key. De opleestijd is in dit geval de tijd tussen de aanbieding op het scherm en de steminzet (aanzet van de respons). In Experiment 1 en 2 was de opleestijd de tijd tussen aanbieding en geheel uitgesproken hebben van het woord (beëindiging van de respons). De proefleider evalueerde elke respons op juistheid en eventuele voice key fouten. Voorafgaande aan de experimentele taak moesten de kinderen tien bekende woorden lezen om aan de taak (met name aan de voice key) te wennen. De duur van dit experiment was ongeveer 25 minuten. Na afloop van het lees-experiment werd, net als in Experiment 1 en 2, een dictiee afgenomen.

Proefpersonen

Op basis van de score op de Eén-Minut-Test (Caesar, 1970) werd uit een populatie (ongelijk aan die van Experiment 1 en 2) van 126 kinderen van Groep 3 (gemiddelde leeftijd ten tijde van het experiment: 87 maanden) een selectie gemaakt van 20 kinderen. Het gemiddelde van de totale groep kinderen op de Eén-Minut-Test bedroeg 28.8 (sd 15.1; min: 12 en max: 100). Het gemiddelde van de 20 kinderen die de geselecteerde groep (11 jongens en negen meisjes) vormden was 31.8 (sd 1.2; min: 30 en

max: 36). Tevens waren de resultaten op een spellingtest bekend (Woorddictee bij Veilig leren lezen; Mommers & Van Dongen, 1986). De gemiddelde score van de totale groep bedroeg 26.9 (sd 3.7; min: 9 en max: 30), en die van de geselecteerde groep 28.9 (sd 1.2; min: 26 en max: 30). Zowel uit de resultaten van de lees- als de spellingtest kan afgeleid worden dat de geselecteerde groep kinderen beschouwd kan worden als goede lezers en spellers.

4.2 Resultaten

De opleestijden van de oefen- en testfase zullen apart worden besproken. Opleestijden gebaseerd op foute responsen veroorzaakt door de proefpersoon, en door foutieve registratie van de voice key werden verwijderd, evenals extreem lange (> 7000 msec.) en extreem korte (< 200 msec.) opleestijden.

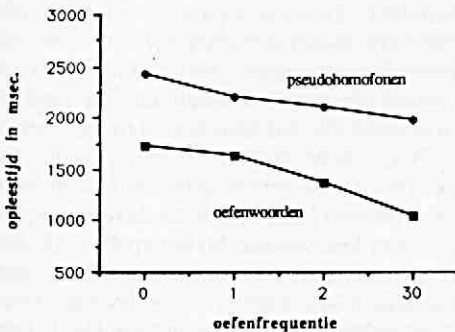
Oefenfase

Een ANOVA op de opleestijden van de gemiddelden van de Frequenties 1 tot 30 liet een significant effect van oefenfrequentie zien ($F_{29,551} = 19.99, p < .001$). De grootste afname in opleestijd (487 msec.) werd gevonden tussen de eerste (2047 msec.) en de tweede aanbieding (1560 msec.) Zowel het verschil in opleestijd tussen de eerste en de tweede ($F_{1,551} = 39.16, p < .001$) als die tussen de tweede en de derde ($F_{1,551} = 13.69, p < .001$) aanbieding was significant. De afname in opleestijd tussen de derde en de vierde presentatie bleek niet significant te zijn. Een analyse van de variabele woordlengte liet een significant hoofdeffect zien ($F_{2,38} = 11.52, p < .001$), de opleestijden van woorden met vier of vijf letters (1013 msec.) was significant korter dan die van woorden met acht letters (1227 msec.), en de opleestijd van woorden met zes of zeven letters (1058 msec.) was korter dan die met acht. Tussen de opleestijden van woorden met vier of vijf en zes of zeven letters was geen significant verschil.

Testfase

Uit een 2 (stimulustype: woorden vs. pseudohomofonen) bij 4 (oefenfrequentie: 0, 1, 2, 30) ANOVA bleek dat beide hoofdeffecten significant waren. Het hoofdeffect van stimulustype ($F_{1,19} = 37.76, p < .001$) liet zien dat de gemiddelde opleestijd voor woorden (1437 msec.) korter was dan die voor pseudohomo-

fonen (2181 msec.) Uit een post hoc analyse, gebaseerd op het significante hoofdeffect van oefenfrequentie ($F_{3,57} = 7.78, p < .05$) is gebleken dat de opleestijden van Frequentie 0 (2076 msec.) en 30 (1509 msec.), van Frequentie 1 (1912 msec.) en 30, en die van frequentie 2 (1739 msec.) en 30 significant van elkaar afwijken (Newman-Keuls 1%, 1%, resp. 5%). In dit geval werd de grootste afname in opleestijd gevonden tussen 2 en 30 aanbiedingen (230 msec.). De interactie tussen stimulustype en oefenfrequentie bleek niet significant te zijn. Het is hier wel van belang op te merken dat er een significant verschil was tussen de opleestijden van de woorden (1720 msec.) en de pseudohomofonen (2433 msec.) van Frequentie 0 (Newman-Keuls, $p < .01$). Frequentie 0 bevatte woorden die de kinderen voor de eerste keer lezen. Figuur 3 toont de resultaten van Experiment 3.



Figuur 3 Opleestijden van de oefenwoorden en pseudohomofonen naar oefenfrequentie van de groep kinderen uit Experiment 3

Spellingresultaten

Op dezelfde manier als in Experiment 1 en 2 werd ook hier bij de beoordeling van de fouten in het dictee een onderscheid gemaakt in kritische en overige fouten. Een 2 (type fout: kritische vs. overige) bij 4 (oefenfrequentie: 0, 1, 2, 30) ANOVA op de spellingfouten liet zien dat beide hoofdeffecten significant waren, maar het interactie-effect niet. Uit het hoofdeffect van type fout ($F_{1,19} = 19.57, p < .001$) kon afgeleid worden dat de kinderen meer kritische (.45) dan overige fouten (.25) maakten. Hoewel ook het hoofdeffect van oefenfrequentie ($F_{3,57} = 2.83, p < .05$) op het gemiddeld aantal spellingfouten significant was, konden uit een Newman-Keuls post hoc analyse geen significante verschillen tussen de

frequenties aangetoond worden. De proporties fout in de Frequenties 0, 1, 2 en 30 waren achtereenvolgens .39, .40, .31 en .29.

4.3 Conclusie

Ook nu kon een frequentie- en een pseudohomofoon-effect aangetoond worden, maar weer bleek er geen interactie tussen oefenfrequentie en stimulustype te bestaan. Net als in de twee vorige experimenten was de afname van opleestijd het grootst tussen de eerste en de tweede aanbieding. Het verschil in opleestijd tussen woorden en pseudohomofonen bij de eerste aanbieding toont aan dat, ondanks het feit dat de aangeboden woorden voor de kinderen nieuw waren, zij inmiddels zoveel orthografische kennis hadden opgebouwd dat ze de woorden sneller konden decoderen dan de pseudohomofonen. De gevormde pseudohomofonen hebben weliswaar een orthografisch toegestane structuur, maar de letterclusters binnen deze structuur komen minder voor in het Nederlandse schrift. Dit blijkt uit de resultaten van een post-hoc analyse op de gemiddelde cumulatieve positionele letterfrequentie (CPTF) van de experimentele stimuli volgens Rolf en Van Rijssoever (1984). De gemiddelde CPTF van de oefenwoorden bedroeg 759 en die van de pseudohomofonen 184, ($F_{1,38} = 13.62, p < .001$).

De resultaten op de spellingtest komen gedeeltelijk overeen met die van Experiment 1 en 2. Net als in de twee vorige experimenten blijken de kinderen meer kritische dan overige fouten te maken. Er was dit keer, anders dan in Experiment 1 en 2, een effect van oefenfrequentie op het aantal gemaakte fouten. Net als in Experiment 1, en in tegenstelling tot Experiment 2 was er geen interactie tussen oefenfrequentie en type fout. De kinderen uit dit experiment hadden evenveel onderwijs genoten als de kinderen uit Experiment 2, kinderen uit Experiment 1 hadden drie maanden korter onderwijs gehad. Dit zou er op kunnen wijzen dat het aantal maanden onderwijs in lezen en spellen mogelijk van belang is voor de inzetbaarheid van de door het lezen opgebouwde representatie bij het spellen.

5 Algemene discussie

In de inleiding werd een aantal vragen gesteld, waarop door uitvoering van drie boven-

staande experimenten geprobeerd werd een antwoord te vinden. De eerste vraag was hoeveel aanbiedingen een beginnende lezer nodig heeft om een woordbeeld op te bouwen, zodat hij of zij zou kunnen overgaan van indirect naar direct lezen. Het antwoord hierop is niet ondubbelzinnig. Op de eerste plaats bleek uit geen van de drie experimenten een interactie tussen oefenfrequentie en stimulistype. Er was echter wel in alle gevallen een effect van oefenfrequentie – een woord dat vaker gelezen is werd ook sneller gelezen – en een pseudohomofoneffect – woorden werden sneller gelezen dan pseudohomofonen. Uit de resultaten van het eerste experiment kon niet afgeleid worden of er woordbeeldontwikkeling had plaatsgevonden. Immers, wanneer uitgegaan werd van de totale stimulusverzameling leek er bij drie aanbiedingen reeds sprake te zijn van de opbouw van woordspecifieke kennis, maar uit de analyse van de deelverzameling kon deze conclusie niet getrokken worden. In experiment 2 leek, ook weer op grond van de totale stimulusset, dat er zelfs na de eerste aanbieding al woordbeeldontwikkeling was opgetreden, terwijl de analyse van de deelverzameling tot de conclusie leidde dat dit pas optrad bij negen aanbiedingen. Het laatste experiment kon helaas ook geen uitsluitsel geven over dit probleem, omdat de kinderen die aan dit experiment deelnamen reeds over zoveel orthografische kennis beschikten dat zij nieuwe bestaande woorden reeds bij de eerste aanbieding sneller lezen dan de daarvan afgeleide pseudohomofonen. Dat er als gevolg van herhaalde aanbiedingen wel degelijk een bepaalde representatie is opgebouwd (hoewel het op dit moment nog niet duidelijk is van welke aard die representatie is) kan afgeleid worden uit het feit dat in alle drie de experimenten oefening in snellere opleestijden resulteerde (frequentie-effect).

De kinderen die deelnamen aan Experiment 3 hadden evenveel onderwijs gehad als de kinderen uit Experiment 2, en deze beide groepen hadden drie maanden langer onderwijs gehad dan de kinderen uit Experiment 1. Maar de kinderen uit Experiment 2 hadden een middelmatige score op de leestest, terwijl de kinderen uit het laatste experiment beschouwd kunnen worden als goede lezers (overeenkomstig de goede lezers uit Experiment 1). Mogelijk dragen deze factoren ook bij aan de gevonden verschillen in uitkomsten

tussen de experimenten.

Een andere verklaring voor het niet vinden van een interactie tussen stimulistype en oefenfrequentie is dat het meten van direct lezen door het gebruik van pseudohomofonen moeilijk vaststelbaar is. Immers uit alle drie de experimenten is gebleken dat, hoewel de pseudohomofonen niet geoefend werden, ze desondanks onder invloed van de oefenfrequentie van de originele woorden stonden. Pseudohomofonen wijken uiteraard af van de corresponderende woorden, maar een groot deel ervan is gelijk aan deze woorden. De mogelijkheid bestaat dat er transfer is opgetreden, en als dat zo is dan is een andere operationalisatie van de vraagstelling wenselijk.

Een tweede vraag die werd onderzocht, was de duurzaamheid van de opgebouwde representatie. Uit het experiment is gebleken dat na een week nog niets van de opgebouwde representatie verloren is gegaan. In hoeverre de stabiliteit van de representatie verder reikt dan een week is niet te beantwoorden, maar ook moeilijk vast te stellen, omdat naarmate de tijd vordert de kans toeneemt dat een kind de experimentele woorden leest dan wel oefent.

Naast de invloed van oefening op de leesprestatie stond ook de vraag centraal in hoeverre de door lezen opgebouwde representatie gebruikt kon worden bij het spellen. Uit het eerste en tweede experiment bleek dat de kinderen geen betere spelprestaties leverden als ze de woorden vaker gelezen hadden. Het 18 keer aanbieden van een woord zorgde er niet voor dat ze het woord beter spelden dan wanneer ze een woord drie maal hadden gezien. Het derde experiment liet weliswaar een effect van oefenfrequentie zien, maar het was een zeer matig effect, omdat het niet vaststelbaar was bij welke aanbiedingsfrequentie de kinderen betere spelprestaties gingen leveren. Een reeds eerder besproken reden voor dit resultaat zou kunnen zijn dat de kinderen uit het eerste experiment drie maanden korter lees- en spelonderwijs hadden genoten dan de kinderen uit de laatste twee experimenten, en dat mogelijk deze drie extra maanden de kinderen dysdanig vaardig had gemaakt dat ze nu wel in staat waren om enigszins gebruik te maken van de door lezen opgebouwde representatie.

Overigens is het wel duidelijk dat de kinderen bij het spellen erg sterk leunen op een fonologische strategie, omdat ze veel meer kritische fouten maakten dan overige fouten.

Hierbij was er geen kwalitatief onderscheid te maken tussen goede en zwakke lezers/spellers. Dit resultaat komt overeen met de studies van Waters e.a. (1985, 1988). Goede lezers/spellers maken weliswaar minder fouten dan zwakke lezers/spellers, maar beide groepen maakten meer kritische dan overige fouten. Dit laatste is overigens in tegenstelling met onderzoek van Frith (1980), waaruit bleek dat zwakke lezers/spellers relatief meer niet-fonologische fouten maakten dan goede lezers/spellers. De resultaten van Frith zijn echter niet zonder meer vergelijkbaar met die van ons, omdat haar groep proefpersonen uit Groep 8 kwam en het leren van de te spellen woorden niet op experimentele wijze had plaatsgevonden. Een andere reden die de vergelijkbaarheid tussen Frith en ons werk bemoeilijkt, is het gegeven dat de orthografie van het Engels aanzienlijk ingewikkelder is dan die van het Nederlands.

Kort samengevat laat het bovenstaande de volgende conclusie toe. Op de eerste plaats is het nog onduidelijk bij hoeveel aanbiedingen van een woord er nu precies woordbeeldontwikkeling optreedt. In hoeverre operationalisatie van de vraagstelling hieraan debet is zal in vervolgonderzoek nader moeten worden onderzocht. Wel mogen we concluderen dat er als gevolg van de herhaalde aanbieding een representatie wordt opgebouwd. Vooral het robuuste frequentie-effect lijkt anders niet verklaard te kunnen worden. Ten tweede lijkt de duurzaamheid van de opgebouwde representatie, van welke aard die dan ook zal blijken te zijn, groot te zijn. Op de derde plaats werd er geen kwalitatief, maar een kwantitatief verschil gevonden tussen goede en zwakke lezers. Ten slotte is gebleken dat de door lezen opgebouwde representatie pas aan het eind van groep 3 inzetbaar is voor het spellen. Ook dit laatste gegeven zal in de toekomst specifiek aandacht krijgen.

Noten

1. De 'Nieuwe streeflijst woordenschat' geeft van een groot aantal Nederlandse en Vlaamse woorden aan welk percentage van de leerkrachten uit het basisonderwijs van mening is dat de betrokken woorden door 6-jarigen passief gekend moeten worden. Van de 15 geselecteerde woorden waren er negen zogenaamde unaniemen, dat wil zeggen meer dan 90% van de leerkrachten vond

dat deze woorden bij de kinderen bekend moesten zijn. Voor het onderzoek werd nagegaan in hoeverre de geselecteerde woorden onbekend waren en waar nodig werd uitleg over de betekenis van het woord gegeven.

2. Het feit dat de experimentele oefenwoorden niet in de leesboekjes van 'Veilig leren lezen' voorkomen biedt uiteraard geen volledige garantie voor de veronderstelling dat de kinderen de woorden ook niet eerder hebben gelezen, maar is er wel een goede indicatie voor.
3. Naar aanleiding van een opmerking van een anonieme reviewer op een eerdere versie van dit artikel werd op een deelverzameling van de stimuli (namelijk zonder de woorden met een 'sch'-en de daarvan afgeleide pseudohomofonen met 'sg'-cluster) opnieuw een analyse uitgevoerd. Hiermee werd het mogelijk differentiële effect van niet bestaande letterclusters ('sg' in één lettergreep komt immers in het Nederlands niet voor) nagegaan.
4. Uit een variantie-analyse waarbij de variabele sekse als factor in de analyse werd opgenomen bleek dat de gemiddelde opleestijd van de groep meisjes (3097 msec.) van de oefenwoorden in de testfase korter was dan die van de jongens (3720 msec.). Dit verschil was echter marginaal significant ($F_{1,31} = 3.86, .05 < p < .10$).

Literatuur

- Bleichrodt, N., P. J. D. Drenth, J. N. Zaal & W. C. M. Resing, *Revisie Amsterdamse Kinder Intelligentie Test, RAKIT*. Lisse: Swets & Zeitlinger, 1984.
- Caesar, F. B., *Eén-minuut-Test voor de technische leesvaardigheid*. Tilburg: Zwijssen, 1970.
- Caesar, F. B., *Veilig leren lezen: Structuurmethode voor het aanvankelijk leesonderwijs*. Tilburg: Zwijssen, 1979.
- Frith, U., Unexpected spelling problems. In: U. Frith (Ed.), *Cognitive processes in spelling*. London: Academic Press, 1980.
- Kohnstamm, G. A., A. M. Schaerlaekens, A. K. de Vries, G. W. Akkerhuis & M. Frooninckx, *Nieuwe streeflijst woordenschat voor 6-jarigen*. Lisse: Swets & Zeitlinger, 1981.
- Kreiner, D. S. & P. B. Gough, Two ideas about spelling: rules and word-specific memory. *Journal of Memory and Language*, 1990, 29, 103-118.
- Mommers, M. & D. van Dongen, *Zorgverbreding bij het leren lezen*. Tilburg: Zwijssen, 1986.
- Raven, J. C., *Standard Progressive Matrices*. London: H. K. Lewis, 1958.
- Reitsma, P. & J. Vinke, Individuele verschillen in direct en indirect lezen. In: P. Reitsma, A. G. Bus & W. H. J. van Bon (red.), *Leren lezen en spellen*. Lisse: Swets & Zeitlinger, 1986.

Rolf, P. C. & R. J. van Rijnsoever, *Positionele letterfrequenties van het Nederlands*. Lisse: Swets & Zeitlinger, 1984.

Seidenberg, M. S., G. S. Waters, M. A. Barnes & M. K. Tanenhaus. When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, 1984, 23, 383-404.

Waters, G. S., M. Bruck & M. Malus-Abramowitz. The role of linguistic and visual information in spelling: a developmental study. *Journal of Experimental Child Psychology*, 1988, 45, 400-421.

Waters, G. S., M. Bruck & M. Seidenberg. Do children use similar processes to read and spell words? *Journal of Experimental Child Psychology*, 1985, 39, 511-530.

Curricula vitae

A. M. T. Bosman studeerde Psychologische Functieleer te Nijmegen en is thans als a.i.o. werkzaam bij de vakgroep Psychonomie van de Universiteit van Amsterdam.

A. M. B. de Groot studeerde Algemene Taalwetenschap in Nijmegen. Van 1979 tot 1987 was zij als wetenschappelijk onderzoeker verbonden aan de vakgroep Psychologische Functieleer van de K.U. Nijmegen. Sindsdien werkt zij als hoofddocent aan de vakgroep Psychonomie van de Universiteit van Amsterdam.

Adres: Universiteit van Amsterdam, Faculteit der Psychologie Vakgroep Psychonomie, Roetersstraat 15, 1018 WB Amsterdam.

Manuscript aanvaard 22-5-'91

Summary

Bosman, A. M. T. & A. M. B. de Groot. 'The development of orthographic images in beginning readers and spellers.' *Pedagogische Studiën*, 1991, 68, 199-215.

In three experiments we investigated the following questions pertaining to young beginning readers: a) how many word presentations does a reader need to form an orthographic image of the word, b) how permanent is a representation acquired through visual presentation, c) is this representation useful for spelling, and d) are the differences between good and poor readers/spellers qualitative, quantitative or both. The results did not answer our first questions, but all experiments showed that some representation had been formed through repeated presentation. The acquired representation turned out to be durable, but hardly useful for spelling. Finally, our results indicated that differences among young beginning readers and spellers are mainly quantitative.

Bijlage A

Stimulusmateriaal van Experiment 1

Oefenwoord	Pseudohomofon	Niet-geoeffend woord
pauw	pouw	schat
schep	sgep	schaar
tijger	teiger	schoen
schoot	sgoot	schuim
schuur	sguur	schrift
augurk	ougurk	auto
schort	sgort	vouw
paleis	palijs	mouw
eiland	ijlant	gebouw
schuld	sgult	kauwgom
oerwoud	oerwaut	vijver
goudvis	gautvis	bedtijd
nijlpaard	neilpaart	glijbaan
stouterd	stautert	wijsneus
wedstrijd	wetstreit	breinaald

Bijlage B

Stimulusmateriaal van Experiment 2

Oefenwoord	Pseudohomofon
schuim	sguim
scheur	sgeur
augurk	ougurk
auto	outo
schuld	sgult
stouterd	stautert
oerwoud	oerwaut
aardbei	aartbij
nijlpaard	neilpaart
breinaald	brijnaalt
schildpad	sgiltpat
wedstrijd	wetstreit

Bijlage C

Stimulusmateriaal van Experiment 3

Oefenwoord	Pseudohomofon
jouw	jauw
wijzen	weizun
meisje	mijsju
dichtbij	digtbei
handdoek	hantdoek
vouwblad	vauwblat
altijd	alteit
glijbaan	gleibaan
tijger	teigur
zacht	zagt
rondje	rontju
lijmen	leimun
stouterd	stautert
nacht	nagt
einde	ijnde
bladzijde	blatzeide
flauw	flouw
krijgen	kreigun
optocht	optogt
paleis	palijs