Reading and Spelling in Children and Adults: Evidence for a Single-Route Model

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr P. W. M. de Meijer

ten overstaan van een door het college van dekanen ingestelde

commissie in het openbaar te verdedigen in de Aula der Universiteit

op woensdag 22 juni 1994 te 15.00 uur

door

Anna Maria Theodora Bosman

geboren te Beneden-Leeuwen (Gld)

".....and she thought that the most beautiful words were
those which were not needed."

Ayn Rand (1943), The Fountainhead, page 361.

Voor:

Kees, die me mijn intellectuele zelfvertrouwen teruggaf

Norma, die erop toezag dat ik het niet weer verloor

Theses

accompanying the dissertation

Reading and Spelling in Children and Adults: Evidence for a Single-Route Model

Anna Maria Theodora Bosman

1. A 'Single-Route Model' (compare Van Orden, Pennington en Stone, 1990) provides a more parsimonious than, but as powerful an explanation for well-known phenomena in the domain of visual word recognition as the classic 'Dual-Route-Model'.

2. The appropriateness of the here applied pseudohomophone paradigm for research of visual word recognition deserves further investigation. (this dissertation)

3. Although it is not clear yet which spelling-instruction method is the most effective one, it is, however, obvious that reading is the least appropriate way to learn the spelling of a word. (this dissertation)

4. The effects established with the first-letter task are better explained in terms of response competition than in terms of contextual facilitation. (this dissertation)

5. The poorer performance of beginning readers executing a proofreading task compared to experienced readers may be explained by assuming a less efficient spelling-verification mechanism. (this dissertation)

6. Reading and spelling are two independent skills. (this dissertation)

7. "The fact of inner speech forming a part of silent reading has not been disputed, so far as I am aware, by any one who has experimentally investigated the process of reading. Its presence has been established, for most readers, when adequate tests have been made." (E. B. Huey, 1908, p. 117). Did we come full circle?

8. The complexity of visual word recognition becomes self evident when one is confronted with the reading behaviour of beginning readers or readers with severe reading problems.

9. The popularity of the 'Dual-Route Model' in psycholinguistic research would not have reached such a high level if the dominant language in scientific research (English) was one with a more shallow orthography like the Serbo-Croatian or Dutch.

10. The mental lexicon is not to be known, only how it behaves, namely, as an emerging transient dynamical property, or as an explicit symbolic representation.

11. All researchers should take the statement of astrophysicist Dr. Ewine van Dishoeck very seriously: "It is essential to turn the switches yourself". (Volkskrant; January, 29th, 1994)

12. Psychological science is one of the most creative art forms: the number of degrees of freedom pertaining to theories, paradigms and research topics is almost innumerable.

13. Working hard and thinking are competing activities if and only if one is equipped with a serial operating mental processor.
    (reply to H. van der Maas' "Catastrophe Analysis of Stagewise Cognitive Development", Doctoral Dissertation, University of Amsterdam, The Netherlands, September 1993).

14. Everything is relative, and even that!
    (reply to K. R. Ridderinkhofs "Interference from Irrelevant Information", Doctoral Dissertation, University of Amsterdam, The Netherlands, October 1993, thanks to E. D. Dekker)

15. Every dissertation contains a sentence that should not be in there. (this dissertation)

Stellingen

behorend bij het proefschrift

Reading and Spelling in Children and Adults: Evidence for a Single-Route Model

Anna Maria Theodora Bosman

1. Het 'Eén-Route-Model' van Van Orden, Pennington en Stone (1990) verschaft een even krachtige, maar zuiniger verklaring voor de bekende fenomenen op het terrein van de visuele woordherkenning dan het klassieke 'Twee-Route-Model'.

2. De geschiktheid van het hier gebruikte 'pseudohomofoon' paradigma voor onderzoek naar visuele woordherkenning verdient nader onderzoek. (dit proefschrift)

3. Hoewel het nog niet duidelijk is wat de meest effectieve spellinginstructie-methode is, is het wel zeker dat lezen de minst geschikte methode is om een woord te leren spellen. (dit proefschrift)

4. De effecten gevonden met de eerste-letter taak kunnen beter verklaard worden in termen van respons competitie dan in termen van contextuele facilitatie. (dit proefschrift)

5. De oorzaak van de slechtere prestaties van beginnende lezers bij het uitvoeren van een tekstcorrectietaak dan van ervaren lezers moet vooral gezocht worden in een minder geautomatiseerd spellingverificatie- mechanisme. (dit proefschrift)

6. Lezen en spellen zijn twee onafhankelijke vaardigheden. (dit proefschrift)

7. "The fact of inner speech forming a part of silent reading has not been disputed, so far as I am aware, by any one who has experimentally investigated the process of reading. Its presence has been established, for most readers, when adequate tests have been made." (E. B. Huey, 1908, p. 117). Did we come full circle?

8. De complexiteit van het proces van visuele woordherkenning wordt pas echt duidelijk wanneer onderzoekers geconfronteerd worden met het leesgedrag van beginnende lezers of lezers met ernstige leesproblemen.

9. Het 'Twee-Route-Model' in psycholinguïstisch onderzoek zou veel minder populair zijn als de dominante taal in wetenschappelijk onderzoek (het Engels) er een was geweest met een eenduidiger relatie tusseen grafemen en fonemen, zoals het Servo-Kroatisch of het Nederlands.

10. Het mentale lexicon kan niet gekend worden, alleen hoe het zich gedraagt, namelijk als een vluchtige dynamische eigenschap of als een expliciet symbolische representatie.

11. Alle onderzoekers zouden de uitspraak van de astrofysicus Dr. Ewine van Dishoeck: "Het is essentieel om zelf aan de knoppen te zitten", ter harte moeten nemen. (Volkskrant 29 januari 1994)

12. De psychologische wetenschap is een van de creatiefste kunstvormen: het aantal vrijheidsgraden wat theorieën, paradigma's en onderzoeksthema's betreft is vrijwel oneindig.

13. Hard werken en nadenken zijn dan en alleen dan concurrerende activiteiten wanneer je uitgerust bent met een serieel werkende mentale processor.
(antwoord op H. van der Maas' "Catastrophe Analysis of Stagewise Cognitive Development", Academisch Proefschrift, september 1993)

14. Alles is betrekkelijk en zelfs dat!
(antwoord op K. R. Ridderinkhofs "Interference from Irrelevant Information", Academisch Proefschrift, oktober 1993, met dank aan E. D. Dekker)

15. In elk proefschrift komt een zin voor die er niet in thuis hoort. (dit proefschrift)

Contents

0

DANKBETUIGING

Er is niet veel in het leven en nog minder in dit proefschrift waar ik volledig zeker van ben, maar één ding staat vast, ik had me geen betere begeleidster kunnen wensen dan Annette de Groot. Ze gaf me niet alleen van het begin af aan alle vrijheid om mijn eigen ideeën te ontwikkelen, maar ze wist bovendien precies wanneer ik meer behoefte aan begeleiding had. In mijn geval was dat bij het rapporteren van de bevindingen zodat het acceptabel Engels werd. In de loop van het project is er een vriendschapsrelatie ontstaan die nooit verstrengeld raakte met de werkrelatie. Ze wist de scheiding van de twee verschillende petten die ze moest dragen tot het eind toe vol te houden. Met liefde en respect: Annette, Dank Je Wel!

Mijn promotor Jan Elshout dank ik ook voor het vertrouwen dat hij in mij stelde en de vrijheid die hij me gaf. Ik mocht mij immers bezighouden met "woordjes", een onderwerp waar alleen de ware 'word nerd' door gefascineerd wordt.

Bij de vakgroep Psychonomie waar ik mijn onderzoek uitvoerde is de eerste die ik wil bedanken mijn sparring partner Martin van Leerdam. Hij leverde een significante bijdrage aan dit proefschrift, en deelde mijn enthousiasme voor het onderzoek naar woordjes. Tot voor kort (sinds de liefde haar intrede heeft gedaan zijn Martin's prioriteiten wel enigszins verschoven) kon ik hem er altijd toe verleiden om tot sluitingstijd in het café te discussiëren over de rol van fonologie bij lezen en spellen. Janet van Hell, mijn huidige kamergenote, maakte mijn verblijf op de vakgroep niet alleen wetenschappelijk interessanter, maar heeft vooral ook bijgedragen aan een bijzonder gezellige sfeer. Janet, ik zal onze bouwkeet (kamer 6.01) en je humor missen. Diane Pecher bleek als stagiaire een betrouwbaar proefleidster, de hypothesen werden immers bevestigd, maar belangrijker waren de altijd levendige discussies die we met ons drieën voerden tijdens Diane's tijdelijk verblijf op kamer 6.01 (over het nut van de psychonomie).

I would also like to thank Greg Stone of Arizona State University, who invited me over for an unforgettable visit to the South-West, and first got me interested in dynamical systems theory. Vervolgens wijdde Maartje Raijmakers van de vakgroep Ontwikkelingspsychologie me op zeer plezierige wijze in in de beginselen ervan.

Dan waren er natuurlijk de studenten die mij met name in het uitvoerend werk hebben bijgestaan: Marjolein Gompel, altijd bereid om te helpen, zonder dat er punten tegenover hoefden te staan; Rineke Keijzer, precies als een boekhoudster; Anna-Titia Goutbeek, niet-te-stuiten enthousiast en altijd wel iets te vertellen, en voor Michiel KlØnhammer was niets te zot. Op de fiets naar de scholen, een Mac Plus en randapparatuur in een tas op het stuur en ik achterop de bagagedrager, stoep op stoep af (niet tegen Nico Nootebaart zeggen hoor). René Snoek behoorde samen met Michiel tot de studenten van het eerste uur. Op René

rustte de vervelende taak om bij hele lieve kinderen uit Groep 3 hele moeilijke dictees af te nemen.

De leden van de Cognitie Groep van de vakgroep Psychonomie bedank ik voor hun kritische opmerkingen tijdens presentaties van mijn onderzoek. In Gerard van Essen en later Elly Zeef heb ik fijne kamergenoten gehad. Was er een probleem met de Mac dan gaf Volkert Balk mij altijd antwoord op de vraag die ik stelde, en dus niet op een vraag die ik niet stelde (een zeldzaamheid onder computerdeskundigen). Hij zorgde er tevens voor dat mijn losse files binnen twee uur zodanig op elkaar afgestemd waren dat ik het opmaken van het manuscript bijna leuk begon te vinden (ik zei bijna hoor!). Reynier Molenaar en Janet van Hell (of was 't Frans) waren zo lief om hun auto voor langere tijd aan mij af te staan, zodat het reizen met de computer naar de scholen niet meer per fiets (of trein) hoefde. In Hansje Ehbets had en in Anja Vendrig heeft de vakgroep een zeer capabele en vriendelijke secretaresse. Beiden gaven mij altijd het gevoel dat alles op het gevraagde moment mogelijk was, en als er iets mis ging dan lag dat nooit aan het secretariaat.

Zonder de onvoorwaardelijke en trouwe deelname van een aantal scholen uit Purmerend en Almere was het onderzoek minder vlot, maar zeker ook minder plezierig verlopen. In Purmerend hielden Anneke Visser en Els Droog van "De Flevo" (nu "De Carrousel"), Ria Slijkhuis en Ingrid van Koningsbruggen van "Het Hoefblad", Jeroen Kramer, Mariëtte Valentijn, en Wil Boersma van "De Klimop", Natascha Koene van "De Kraanvogel" (nu "De Carrousel"), Gea Rattink, Yvon Vellinga en Mirjam Bakker van "De Voorhof", Carlien Symor van "De Where", en in Almere Peggy Rodijk en Tineke Kraak van "De Scheepskameel", Monica van Mierlo, Karin van Herten en Loes Haringsma van "De Vlieter", Lucia Tevreden, Sjoerd-Jan Visser, Mirjam Rappel, en Ingrid van Dijk van "De Waterschapsheuvel", Luciënne Lubbers, Hilda Reitsma, Jackie Nienkemper en Monique Blom van "De Zeetuin" gedurende het hele project hun klasdeur zeer gastvrij voor mij open.

De deelname van de basisscholen, "Ichthus" in Lelystad, "Het Noorderlicht" en "De Zesde" (nu "De Graeffschool") in Purmerend, "Voorweg" en "Sint Antonius" in Heemstede, "De Eenhoorn", "Sint Lucia", "Don Bosco" en "Sint Franciscus Xaverius" in Haarlem, "De Meerpaal", "In den Bongerd", "De Middelburg", "De Rank", "Tesselschade" en "De Vogelboom" in Alkmaar mag ook niet onvermeld blijven.

Tenslotte wil ik mijn ouders (voor wie het niet zo nodig hoefde, maar die het achteraf toch leuk meegenomen vinden zo'n doctor in de familie), mijn zussen Thea en Marianne (die diverse hand- en spandiensten verrichtten), en mijn beste vrienden Kees Bisseling en Norma Montulet, Susan Leloup (Bed & Breakfast faciliteiten, en gedeelde smart is......), Reynier Molenaar (voor alle praktische hulp), en Sjanneke Broekmeulen (altijd geïnteresseerd) bedanken voor de diverse vormen van steun die zij me gaven. And,.......when I was too (pre)occupied with the dissertation, Greg's decision to travel into the heart of Borneo stopped my feeling so guilty.

VEELLL DANK

1

GENERAL INTRODUCTION

"And so to completely analyze what we do
when we read would almost be the acme of a
psychologist's achievements, for it would be
to describe very many of the most intricate
workings of the human mind, as well as to
unravel the tangled story of the most
remarkable specific performance that
civilization has learned in all its history."
(E. B. Huey, 1908, p. 6)

This dissertation is a reflection of the development of experimental research on the reading and spelling of single isolated words in beginning and fluent readers, which I carried out over a period of five years. At the outset (mid 1988) of the project a research program had been developed, which took the 'dual-route' model of word recognition (Coltheart, 1978) as its theoretical starting point. During the project it gradually became clear that the model that had dominated the eighties and which had been one of the cornerstones of psycholinguistic research might not survive the nineties. An experimental study that seriously countered the assumptions of dual-route theory and that appeared to be in hindsight the first of a large number of studies that did the same, was an article by Van Orden (1987). In 1990 Van Orden, Pennington, and Stone published an impressive paper, in which they convincingly showed, at least in my opinion, that the assumptions of dual-route theory do not stand up to the empirical facts. They not only criticised the dual-route model, but also supplied a single-route alternative. While this theoretical revolution was taking place in the literature, the accumulating results of my own experiments told me that word-recognition phenomena indeed do not need a model in which two independent routes are assumed. A more parsimonious single-route model seemed (and still appears) to be sufficient to explain the data. In this part of the dissertation I will describe the theoretical course the research has taken, and explain along the way the rationale of the studies that are presented in the Chapters 2 through 7.

Theoretical background

The major assumption of dual-route theory is the idea that skilled readers have at their disposal two different, independent processes (or routes) to read a word. The most important process is the lexical or direct route. A high-frequent word is supposed to have a representation in a mental lexicon that contains its orthography (spelling pattern), its

phonology (how it sounds), its meaning (a kind of dictionary description) and its grammatical class (for instance, verb, noun, or adjective). A reader who is presented with a word that has an orthographic representation in the mental lexicon can map the input pattern onto the orthographic pattern already present and activate its meaning without prior activation of its phonology. Phonologic activation is mandatory in case a reader has to pronounce the word, but if silent reading is required it is optional. The second, more subsidiary process, is the nonlexical or indirect route to the mental lexicon. A reader presented with a low-frequent word, lacking an orthographic representation in the mental lexicon, has to recode the graphemes of the word into phonemes in order to compute the word's phonology. The phonologic representation can in turn activate its meaning. This process of recoding is also referred to as phonologic mediation, because the phonologic property of the word serves as an intermediary agent between the orthographic pattern and the semantic representation.

According to Van Orden et al. (1990) three hypotheses are interwoven with the independent-processes assumption. They used these to organise their paper. I will discuss them to clarify why dual-route theory had to postulate two independent routes, and how the developmental shift from a novice reader into a highly skilled expert reader is supposed to take place. The GPC (Grapheme-Phoneme Correspondence) hypothesis states that the phonology of a word will be computed by means of grapheme-phoneme correspondence rules. These GPC-rules are discrete all-or-none rules (Venezky, 1970 in Van Orden et al., 1990). The GPC hypothesis necessitates the distinction between two types of words, namely, words that follow GPC-rules (for instance, 'put'), and words that do not (for instance, 'aisle'). Words that obey GPC-rules can be read via the nonlexical or phonologic route, whereas words that violate the GPC-rules have to be read by means of the lexical route. Because of the existence of a large number of so-called irregular words in the English language the postulation of a lexical route seemed inevitable. On the other hand, the fact that subjects are capable of reading nonwords, of which no representation is present in the mental lexicon, made it necessary to assume the nonlexical route.

A beginning reader who is let into the secrets of the alphabetic principle will initially read all regular words by means of recoding graphemes into phonemes. Only after substantial practice will the beginner be able to read these words lexically, that is, without phonologic mediation, because an orthographic representation will have been developed. Thus, when subjects have gained experience in reading they will be able to bypass the nonlexical route. This is known as the bypass hypothesis. The use of phonology in order to read words is either an indication of limited reading experience, or in the case of experienced readers a sign that a low-frequent word has been presented. The third hypothesis that is connected to dual-route theory is the delayed phonology hypothesis. Computing the phonology according to GPC-rules is relatively time consuming as compared to the process of direct access to the mental lexicon. Thus, the outcome of the nonlexical process in experienced readers will hardly ever interfere with that of the lexical process, because the nonlexical route usually lags behind. It is a well-known fact that reading times of beginning readers are considerably slower than those of fluent readers, a finding that can be explained by the assumption that word reading in beginning readers is predominantly phonologically mediated.

The initial research question that emerged from the theoretical framework described above can be summarised as follows: How many presentations of a word does a beginning reader need to develop an orthographic representation (also referred to as orthographic image) that enables the child to bypass phonologic mediation? An additional question that arose was whether a beginning reader who has developed an orthographic image will use this representation to spell the word. These were the major issues investigated in Chapter 2. Beginning readers, with less than a year of formal reading and spelling instruction, had to learn to read new words. Some of these words were presented relatively often (high-frequent condition) and the others were presented less often (low-frequent condition). It was expected that after a certain number of presentations the children would have developed an orthographic representation that enabled them to read the words without the application of grapheme-phoneme correspondence rules. To find out after how many presentations of a word this would happen, the children were to read, in a test stage, both the trained words of the high-frequency and low-frequency conditions and pseudohomophones that were derived from these words. A pseudohomophone is a letter string with a phonology identical to a word, but with a different orthography. For instance, the letter string 'FONEAM is a pseudohomophone of the word 'PHONEME'. Pseudohomophones are not supposed to have a representation in the mental lexicon and therefore have to be read through the nonlexical route. An interaction between type of stimulus and frequency was expected. That is, after a certain number of presentations (frequency factor) of a word the subject will read the word faster than its corresponding pseudohomophone (which has not been practised), because the word is read through the fast lexical route, whereas its pseudohomophone has to be read via the slower nonlexical route. The interaction, thus, indicates the number of presentations necessary to form an orthographic image used in the reading process. Three experiments were executed in which frequency was manipulated. No interaction emerged, not even after 30 presentations of a word. It was clear that repeated readings of a word speeded up the naming process, but it also shortened the naming times of the pseudohomophones that were derived from these words. Transfer from the word to the pseudohomophone seemed to have occurred. This is plausible, because despite the orthographic difference between a word and its pseudohomophone there is also considerable overlap between the two types of stimuli. The results of the study described in Chapter 2, thus showed that some kind of mental representation had developed, but its nature remained unknown. It was hypothesised that the pseudohomophone operationalisation may not have been an adequate tool to investigate the shift from indirect to direct reading. Below, I will introduce an alternative research paradigm that was used in Chapter 5 to investigate the issue of the presumed shift in reading.

Because it was not possible to assess whether or not the children had developed an orthographic representation, it was also impossible to test whether these beginning readers used the orthographic image in spelling. The data did show, however, that these subjects had not learned the spelling of words as a result of the reading training, and their spelling products revealed the subjects' heavy reliance on phonology. The experiments undertaken in Chapter 2 suggest that reading is not a very effective method to teach beginning readers the spelling of words. It is, however, not clear whether the subjects would have learned the spelling of the words with a spelling instruction method other than a reading training. In Chapter 3, precisely this issue was subject of investigation. The effects of a reading training

on spelling performance was compared to the effectiveness of three non-reading spelling-instruction methods.

As was indicated by the studies of Chapter 2, beginning literates rely on a phonologic strategy to spell words. If beginning reading is indeed characterised by a shift from indirect to direct reading, it is not unlikely that the same shift occurs in beginning spelling. Kreiner (1992) and Kreiner and Gough (1990) argue that highly skilled readers use both direct (word-specific strategy) and indirect (rule strategy) strategies to spell words, and neither of the two has temporal primacy. Chapter 4 deals with the nature and the development of the spelling process in beginning spellers. The issue of the nature of the spelling process will be picked up again in Chapter 7, in which the relationship between reading and spelling will be clarified. Before doing so I will introduce the Chapters 5 and 6. These chapters reveal the theoretical turning point of this dissertation.

I hypothesised earlier that the pseudohomophone manipulation applied in Chapter 2 was inappropriate to investigate the shift from indirect to direct reading in beginning readers. The paradigm that was applied in Chapter 5 is the first-letter-naming task developed by Rossmeissl and Theios (1982). In their study they presented fluent readers with words, orthographically legal pseudowords and orthographically illegal pseudowords (anagrams). Subjects were required to name the first letter of the visually presented stimuli. First-letter-naming times of words and legal pseudowords were shorter than those of orthographically illegal pseudowords. They explained the faster naming times of the former groups of stimuli in terms of contextual facilitation. Fluent readers not only perceive letters in a letter string in a parallel fashion but they employ the orthographic context as an independent source of information. This enables them to identify the first letter of legal letter strings (words and legal pseudowords) faster than of illegal pseudowords. According to Rossmeissl and Theios (1982), this effect is another indication that fluent readers use the direct route in reading. If they had read the words by means of the indirect route the orthographic context would not have influenced the naming times of the first letter, because this process is supposed to proceed from left to right.

In Chapter 5, I tested the hypothesis that beginning readers in Grade 1 do not show the first-letter effect, that is, name the first letter of legal letter strings faster than those of illegal letter strings, because they still read indirectly. Much to my surprise, the group of beginning readers showed the same effect as the fluent readers of Rossmeissl and Theios (1982). They were also faster on the word stimuli than on the nonword stimuli. A straightforward interpretation of this result is that these readers already went through the transition from indirect to direct reading. This, however, is not very likely given classroom observations. The fact that no interaction occurred was the first hint that a major qualitative difference between the reading behaviour of beginning and fluent readers may after all not exist, and hence that a developmental shift from indirect to direct reading should not be expected. The experiments in Chapter 5 show that the first-letter effect should not be interpreted in terms of context facilitation at the identification level, but are better explained in terms of response competition. The results of Chapter 5, with both beginning and fluent readers, appear to be compatible with a single-route model of reading, that only retains the indirect or phonologic route. At this point in the development of my research I started to acknowledge the merits of the model of Van Orden, Pennington and Stone (1990), which I will describe in some more detail shortly.

In Chapter 6, three experimental tasks that were used by Van Orden (1991) and Van Orden, Stone, Garlington, Markson, Pinnt, Simonfy, and Brichetto (1992) to investigate the reading behaviour of fluent readers were used to investigate the reading process of beginning readers. I reasoned that if the reading process of fluent readers was not fundamentally different from that of beginning readers, the results they found should emerge in my group of beginners as well. This indeed was the case. Both the beginning readers who participated in Chapter 6 and the fluent readers of Van Orden et al. (1992) showed huge and significant phonologic effects, suggesting that the reading process of all readers is phonologically mediated, or in terms of Van Orden et al. (1990) 'phonologically constrained'.

This is probably a suitable place to elucidate the single-route model that I would advocate as one that can explain word-recognition phenomena more parsimoniously than the dual-route model. Van Orden and Goldinger (in press a, b) do not refer to their ideas on printed word perception as a model, but rather as a theoretical framework. A general description of the time course of word-recognition processes is presented in Van Orden (1987). It is assumed that an orthographic input pattern activates its phonology, which in turn activates its meaning. The meaning-activation process is followed by a verification process, which is in fact a spelling-check mechanism that operates on the orthographic input pattern. The most controversial assumption of the model is the idea that the presentation of an orthographic pattern always causes phonologic activation to emerge. The proposed verification process is not unique to the model. It has also been included in other word-recognition models (see Paap, Newsome, McDonald, & Schvaneveldt, 1982; Grossberg & Stone, 1986). The theoretical framework of Van Orden and his co-workers (Van Orden & Goldinger, in press a, b; Van Orden et al. 1990) is not based on symbolic ("flow-chart") cognitivism in which representations are assumed that serve as the symbols on which rule-governed operations are performed, but it is rooted in dynamical systems theory. The general ideas of dynamical systems theory as applied to cognitive processes are described in detail in Stone (in press), and in Van Orden and Goldinger (in press a, b). I will restrict myself to an explanation of how phonology is thought to constrain word recognition at an early stage according to the phonologic coherence hypothesis, which sprouts from the broader adaptive dynamic systems metaphor. The following description draws heavily from the work of Van Orden and Goldinger (in press a).

In dynamical systems theory, in contrast to information processing theory, no explicit representations (i.e., symbols) are assumed. Instead, patterns of activation act as the systems' key notions, and these patterns of activation are only maintained when they are active. In visual word recognition an orthographic pattern emerges across visual features, and induces an activation pattern across linguistic features. If the linguistic features' feedback pattern is matched adequately to the stimulus-driven visual (orthographic) pattern the cycle becomes self-perpetuating and a coherent dynamic whole emerges. Performance of the cognitive system can be described by using an adequate level of 'representation'. Van Orden et al. (1990), and Van Orden and Goldinger (in press a) argue that pragmatic substitutes, referred to as subsymbols, are used for the initial conditions. The subsymbols, which are the nodes that interact in the dynamical approach, do not refer to lower level symbols in the sense used in representational theory, they are not "psychologically real", but have "an exclusively narrative function", sic (Van Orden & Goldinger, in press a). In the

dynamical systems approach of word recognition it is assumed that activation from visual features will spread to linguistic features, comprising orthographic, phonologic and semantic subsymbols, which are interconnected. Coherence amongst all three kinds of subsymbols will be established, but the order in which coherence emerges is determined by the level of self-consistency between them. Stated differently, subsymbols sharing a relatively consistent relation (between orthographic and phonologic subsymbols) are supposed to cohere earlier than subsymbols sharing a less consistent relation (between orthographic and semantic subsymbols). The relations between words and their meanings varies more than between words and their phonology (there are only a few exceptions, for instance, in English: 'wound' and 'tear'; in Dutch: 'regent', 'bedelen' and 'kantelen'). A word's meaning is largely determined by the context in which it appears, and high-frequent words tend to be polysemous. This explains why phonology constrains (determines) word recognition at an early stage.

The proposed dynamics that establishes the connection between orthography and phonology will be described now, but my account remains necessarily simple and incomplete. Connections in a network exist between orthographic subsymbols (graphemes) and phonologic subsymbols (phonemes). After being exposed to a great variety of words, the network learns that covariations exist between orthographic and phonologic subsymbols. When learning (i.e., covariant learning) occurs, the connection weights between orthographic and phonologic nodes are adjusted in such a fashion that their size reflects the degree of covariation. The connection weight between, for instance, the orthographic subsymbol 'D' and the phonologic subsymbol [d] will be increased every time 'D' occurs with [d]. Here the orthographic-phonologic correspondence is shared consistently across contexts, because in most cases a 'D' codes for [d]. But in the case of an 'O' the picture changes, because in some circumstances the 'O' codes into [uː], as in "DO", whereas in "NO" the 'O' codes into [oʊ]. This particular orthographic-phonologic correspondence is inconsistent across contexts. The only way the system can solve the ambiguity of coding the orthographic subsymbol 'O' is by taking into consideration the context in which the particular subsymbol occurs. In case the orthographic subsymbol 'O' occurs in the word context "DO", not only the connection weights between 'D' and [d], and 'O' and [uː] increase, but also the ones between 'D' and [uː], and 'O' and [d]. If, on the other hand, the orthographic subsymbol 'O' occurs in the word context "NO", the connection weights between 'N' and [n], 'O' and [oʊ] increase, and also the ones between 'N' and [oʊ], and 'O' and [n]. The fact that each orthographic subsymbol in the word contributes to the coding of each phonologic subsymbol in that word results in the proper pronunciation of relatively inconsistent words.

The naming of both regular words (with consistent grapheme-phoneme correspondences; e.g., 'bill' is as 'fill', 'gill', 'hill', 'kill', 'mill', 'pill', 'rill', 'still', 'till', 'vill', and 'will') and irregular words (with inconsistent grapheme-phoneme correspondences; e.g., 'have' is unlike 'cave', 'dave', 'gave', 'lave', 'pave', 'rave', 'save', 'wave') occur by means of the same dynamics. Equally fast reading times for regular and irregular high-frequent words are expected, because the connections of high-frequent words are updated relatively often. Thus, frequent updating in case of irregular high-frequent words with inconsistent orthographic-phonologic correspondences overcomes the noise in the coding. Therefore no regularity effect is expected when high-frequent words are presented. In contrast, if the

system is presented with low-frequent words, a regularity effect is supposed to come about. The connections of low-frequent irregular words are coded less precisely than those of low-frequent regular words, and are therefore vulnerable to inconsistent crosstalk, and resonance will thus be slower. This is how the phonologic coherence hypothesis explains the regularity by frequency interaction as reported by Seidenberg, Waters, Barnes, and Tanenhaus (1984). The distinction between two independent routes for word recognition as is postulated in the symbolic dual-route model is thus unnecessary given the phonologic coherence hypothesis. In Chapter 7 the dual-route model is contrasted with the phonologic coherence hypothesis. From both models predictions are derived and tested in four experiments. These predictions concern both the reading and spelling processes in beginning and highly skilled literates.

In the course of this dissertation I hope to make clear, particularly in the final chapter, that a single-route model, the 'phonologic coherence hypothesis', suffices to explain the results of all the present six studies on reading and spelling.

General procedural information on subjects and methods used.

The topic of this dissertation, reading and spelling in Dutch-speaking children and adults, necessitates that some additional information is presented here on various aspects of the experimental studies. I will describe very briefly the orthography of the Dutch language, the subject groups, the reading instruction method used by the subjects, the tests administered to the beginning readers, the type of stimuli used, and the experimental tasks that were applied.

Dutch orthography is fairly shallow. That is, the relationship between graphemes and phonemes is in most cases regular, which makes sounding out Dutch words a reliable reading strategy. On a scale that represents degree of regularity between graphemes and phonemes, it is often assumed that the Dutch orthography is near the regular end, whereas the English orthography (also referred to as a deep orthography) is the quintessential example of an orthography that is near the irregular end (Van Heuven, 1980). It is probably fair to say that Dutch orthography is far more regular than English, but less so than Finnish, Spanish, or Italian. I refer to Van Heuven (1980) or Reitsma and Verhoeven (1990) for a more detailed description of Dutch orthography.

The subjects who participated in this study were groups of beginning and fluent readers. The fluent readers were undergraduates from the Faculty of Psychology of the University of Amsterdam with Dutch as mother tongue. The beginning readers attended Grade 1 (their mean age at the time of experimental testing was 87-88 months), and were all instructed according to the reading instruction method "Veilig leren lezen: structuurmethode voor het aanvankelijk leesonderwijs" (Learning to read safely: structure method for teaching beginning reading, Caesar, 1979). This method of teaching reading is called 'structure method', because the emphasis in this curriculum is on the structure of the orthographic system. It takes an intermediate position between whole word instruction and phonics instruction (Rayner & Pollatsek, 1989, pp. 348-349). Initially only regular words are used and after four months of instruction the children are familiar with the main grapheme-phoneme correspondence rules. It is a fairly rigid pre-programmed curriculum, which imposes a strict day-by-day and week-by-week progression. Assessing the reading and

spelling level of all children, who attended different schools, could therefore be reliably executed.

The majority of the experiments were executed nine or ten months (in May or June) after formal reading and spelling instruction had started. The subjects who participated in the first experiment of Chapter 2 and the first experiment of Chapter 4 had had seven (March) and six (February) months of formal reading and spelling instruction respectively. Reading and spelling skill, as well as verbal and non-verbal intelligence, of all subjects was assessed prior to the experiments. Reading level was measured by means of a standardised test for reading decoding ('Eén-minuut-test voor de technische leesvaardigheid', which translates into 'One-minute-test for reading decoding', Caesar, 1975). The reading test consists of a list of unrelated words, which have to be read aloud. The score on the reading test is the number of words read correctly in one minute. The average is 28 (*SD* = 15) after six months in first grade. All subjects were tested on reading decoding in February, three to four months before the experiments were performed. This, however, does not pose a problem, because the stability of this test is relatively high. The correlation between test scores assessed in March and the beginning of July is .85. This reading test is keyed to the reading curriculum "Veilig leren lezen" (see above) .

Spelling level was measured with a standardised word-dictation test ('Woorddictee bij Veilig leren lezen', which translates into 'Word dictation to Learning to read safely', Mommers & Van Dongen, 1986), that is also keyed to the reading curriculum 'Veilig leren lezen'. The internal consistency of this test is high, alpha is .91. The test consists of 30 items (i.e., words from the reading method). The score is the number of correctly spelled words in a dictation task. The non-verbal intelligence test was the Standard Progressive Matrices (Raven, 1958), a test which is regarded as a reliable measure of g (Vodegel Matzen, Van der Molen, & Dudink, 1994). The items are all analogy problems, and the idea is to supplement a missing figure from a set of alternatives. The reliability (Cronbach's alpha) of this test, based on a Dutch sample of 219 children, age 84 months, was .87.

The subtest 'woordbetekenis' (word meaning) of the RAKIT ('Revisie Amsterdamse Kinder Intelligentie Test' which translates into 'Revised Amsterdam Children's Intelligence Test', Bleichrodt, Drenth, Zaal, & Resing, 1984) was used to assess verbal intelligence. This subtest can be characterised as a test for passive vocabulary (similar to the "Peabody Picture Vocabulary test"). The test consists of four-choice picture items. The person administering the test names a word and the child then has to choose the picture that best represents the word that was named. The split-half reliability of the test for the subject group under investigation here is .82.

In some of the experiments the group of beginning readers is divided in good and poor readers, and in a few cases also in good and poor spellers. Because of the high correlation between reading and spelling, it appeared that the distinction between good and poor readers automatically implied one between good and poor spellers. I would like to emphasise that this distinction is not used as an absolute criterion, but purely as a means to compare the reading or spelling behaviour of the better readers with that of their less well performing classmates. Care was always taken to exclude children with severe reading and spelling problems from the experimental tasks. The reading decoding test appeared to be a reliable instrument for this selection. Children with a reading score less than 15 were never included in the experimental sample; experience taught that a score of minimally 15 was required to

perform the experimental tasks satisfactorily. Children with a mother tongue other than Dutch were also excluded from the experiments. Both the children with severely limited reading skills and the non-native Dutch children were tested nevertheless, but purely for social reasons.

The stimuli used in the experiments were either single words or word-like letter strings. The words were either chosen from the books of the reading curriculum or selected from the list "Nieuwe streeflijst woordenschat voor 6-jarigen" (which translates into: New target vocabulary for 6-year olds; Kohnstamm, Schaerlaekens, De Vries, Akkerhuis, & Froonincksx, 1981), depending on the goal of the experiment. The list of Kohnstamm et al. (1981) presents the percentage of Dutch and Flemish teachers from regular primary schools who take the view that these words should be known passively by 6-year olds. The majority of the words that were used as stimuli in this dissertation were so-called 'Unaniemen', that is more than 90% of the teachers believe that these words should be known by the children. I used the word list of Kohnstamm et al. (1981) for the selection of stimuli in case the experiment required words that were semantically familiar, but orthographically unfamiliar to the children. Thus, words that do not occur in the reading materials of the children. Using words that do not occur in the reading books obviously does not guarantee that the children have never seen the words before. It is, however, a good indication. Furthermore, all stimuli, whether presented on paper or on the screen of a Macintosh computer were printed in font Helvetica. Font Helvetica is highly familiar to the subjects who participated, because it is used in the first eleven volumes of the reading curriculum "Veilig leren lezen". Describing the stimuli sometimes necessitated a phonetic transcription. An English-Dutch dictionary (Van Dale, 1984) and Booij (1981) were consulted for these transcriptions.

One could argue that reading is more than decoding single words. I do not dispute this, but my main concern in this dissertation is with how single words are processed. In some of the experiments learning to read or spell new words was explicitly aimed at (see for instance Chapters 2, 3, 4 and 7). It has been debated whether words are better learned when they are presented in context or in isolation (Goodman, 1965, in Nicholson, 1991, and Nicholson, 1991). Ehri and Roberts (1979) showed that children who were to read words in printed sentence contexts learned more about the semantics of the presented words, whereas children who were to read words in isolation (flash-card method) showed superior knowledge of the orthographic form. Semantic processing of the words to be read was not a topic of this study. Moreover, knowledge of the meaning of the stimulus words was a prerequisite, which I tried to guarantee by using the word list of Kohnstamm et al. (1981). The use of single words is thus warranted.

It is also important to note that in investigating the spelling process in this dissertation, I was solely interested in words with a unique spelling, for instance: is it GEIT or GIJT (an English example: HARASS or HARRASS). Homophones, words identical in phonology, but different in spelling, for instance: PEIL or PIJL (Assink, 1982; an English example: DEAR or DEER), are not considered, and neither are grammatically based verb spellings, which constitute a spelling problem specific to the Dutch language (Assink, 1990).

Initially it was thought that the age of the subjects might severely restrict the choice of experimental tasks and the precision of measuring the dependent variables. In the first two experiments of Chapter 2, for instance, it was mistakenly assumed that it would not be possible to use a voice-key with children seven years of age. It later turned out that,

nevertheless, it was possible to do so. Therefore, in the remainder of the experiments all latencies based on verbal responses were registered with a voice-key. It goes without saying that the children produced somewhat more errors due to activating the voice-key inadvertently than did the adult subjects. When a time response was required the stimuli were presented on the screen of a Macintosh computer, and registration was controlled by a computer program. Throughout this dissertation a number of different experimental tasks has been used avoiding that conclusions concerning word-recognition processes would have to be based on a single experimental paradigm. The experimental tasks comprise: naming, i.e., reading aloud the presented stimuli (Chapters 2, and 7); lexical decision, i.e., deciding by means of a button press whether a presented stimulus constitutes a word (Chapter 6); semantic categorisation, i.e., deciding by means of a button press whether a presented stimulus is a member of a pre-designated category (Chapter 6); proofreading, i.e., reading a text silently or aloud and marking wrongly spelled words (Chapter 6); first-letter naming or first-phoneme naming, i.e., naming the first letter or phoneme of presented stimuli (Chapter 5); spelling recognition, i.e., deciding which of two presented stimuli denotes the correct spelling of a target word (Chapters 3, 4, and 7); and spelling production, i.e., a dictation task (Chapters 2, 3, 4, and 7).

A final remark concerns the application of the statistics used here. After Clark's (1973) seminal article on the use of the proper *F*-test in psycholinguistics, researchers started to report *min F'*, which enabled them to generalise beyond the materials used. This test treats both the subjects and the items as random effects in one and the same analysis. Shortly after Clark's publication Wike and Church (1976) criticised the *min F'*-test, because of being too conservative, and therefore leading to possible Type II errors. The result of this was that researchers started to report both F_1 (subject analysis) and F_2 (item analysis) separately, without computing *min F'*. Schrijnemakers (in press) showed that this is statistically incorrect. He also made clear that if proper measures are taken, that is, stimulus materials are matched or each stimulus appears in each condition (either by random assignment or by experimental manipulation) the proper test is the subject analysis (F_1). In the present dissertation, in most cases either the stimuli appeared in all conditions, or they were matched, or the materials were actually a fixed effect, because no more stimuli could be generated given the limitations of the set of stimuli that had to be developed. This, therefore, justifies the mention of F_1 only in this dissertation.

2

THE DEVELOPMENT OF ORTHOGRAPHIC IMAGES IN BEGINNING READERS AND SPELLERS*

Summary

In three experiments with young beginning readers as subjects we investigated: a) how many word presentations a reader needs to form an orthographic image of the word, b) how permanent is a word representation acquired through repeated visual presentation of the word, c) whether a representation acquired through reading is useful for spelling, and d) whether the differences between good and poor readers/spellers are qualitative, quantitative or both. The results did not answer our first question, but all experiments showed that some or other 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.

The first stage of beginning reading in an alphabetic writing system is characterised by sequential decoding of graphemes into phonemes. In languages with a strict correspondence (as in Finnish) between orthography and phonology this strategy is always more or less successful. If this correspondence is less strict (as in English) or absent (as in Chinese) a reader cannot rely on the system of grapheme-phoneme correspondences.

This, however, does hardly appear to be a problem for an experienced reader, because he or she generally does not read by means of phonology (indirect), but recognises the letter string as a whole (direct). It is assumed that a reader because of repeated presentations with a word gradually develops an 'orthographic image' of the word in question. The term 'orthographic image' refers to a mental representation on a higher level than the representations of graphemes and phonemes, and is assumed to contain visuo-spatial information of the coded word. To avoid the connotation with the representation being visual, instead of the phrase 'orthographic image', the terms 'word-specific knowledge' or 'word unique letter patterns' have been used (Reitsma & Vinke, 1986) in the Dutch literature on reading. However, in accordance with the English terminology (visual or orthographic code) we have chosen to use the term 'orthographic image' here.

Evidence for the assumption that fluent readers make use of orthographic images was provided, amongst others, by Seidenberg, Waters, Barnes and Tanenhaus (1984). In this study, with adult subjects, naming times for regular and irregular high-frequent and low-frequent English words were compared. Regular words are words with a univocal relation

---

* This chapter is based on an article co-authored with A. M. B. De Groot (1991), De ontwikkeling van woordbeelden bij beginnende lezers en spellers, *Pedagogische Studiën, 68*, 199-215.

between their phonologic and orthographic form; in irregular words this relation is not univocal or less so.

Seidenberg et al.'s study showed that naming times for low-frequent irregular words were longer than those for low-frequent regular words. But the naming times of high-frequent regular and high-frequent irregular words were equal. From the interaction between frequency and regularity it has to be concluded that experienced readers of English read high-frequent words in a different way than by coding graphemes one by one into phonemes. If the reading process would always only involve the recoding of graphemes into phonemes than naming times for irregular high-frequent words should have been longer than those for regular high-frequent ones, because the process of recoding graphemes into phonemes should be slowed down by each irregularity it encounters. This, however, turns out not to be the case, and Seidenberg et al. conclude that experienced readers have stored an orthographic image of high-frequent words. In the case of high-frequent words, instead of using the indirect route experienced readers can match the written word with a mental representation, the orthographic image (direct reading), which then directly activates a representation of how to pronounce the word.

The research of Seidenberg et al. not only supplied evidence for the existence of orthographic images, but it also showed that the frequency of word use is important for the development of orthographic images. This appears from the shorter naming times of regular low-frequent words than of irregular ones, whereas those of both types of high-frequent words were equal.

Reitsma & Vinke (1986) investigated the relation between frequency and the development of orthographic images in a more direct way than Seidenberg et al. (1984) did. In a training session young beginning readers were presented with words either three, nine or 18 times. After the training the naming times were registered of the trained words and of the 'pseudohomophones' derived from these words. Pseudohomophones are letter strings that do not exist as words as such, but have a phonology that is identical to that of existing words (for instance, the letter string "sute" for 'suit'). This study showed that presentation frequency interacted with stimulus type (training word vs. pseudohomophone). Three or less presentations of a stimulus did not render differences in naming times between training words and pseudohomophones, but after nine and 18 presentations the naming times of training words were significantly shorter than those of the corresponding pseudohomophones. This effect shows that after three presentations of a word no orthographic image has developed yet, but that it has after nine presentations. Reitsma and Vinke did not find an effect of stimulus type with a group of children attending special education, not even after the maximum of 18 presentations. It thus seems that by then these children have not developed an orthographic image yet.

It is not only interesting to elucidate the role of the orthographic image in reading, but it is also important to reveal its function in spelling. Waters, Bruck, and Seidenberg (1985) concluded from their study that beginning readers (Grade 3) use grapheme-phoneme correspondence rules in both reading and spelling, and that they do not use the orthographic image in spelling ("do not...'read out' the orthographic form"). The studies of Waters et al. and of Waters, Bruck, and Malus-Abramowitz (1988) also showed that no qualitative differences existed between good and poor readers/spellers in the way they read and spell.

According to Kreiner and Gough (1990) adult readers in some cases spell words using phonology and in other cases by means of reading out the orthographic image.

That the children in the study of Waters et al. (1985) did not seem to use the orthographic image in order to spell a word is not at all remarkable, because these children presumably had not developed one. This is suggested by the fact that they also appeared to read phonologically. The conclusion that in contrast, adult readers are capable of using the orthographic image in spelling (Kreiner & Gough, 1990), raises the question whether young children who have developed an orthographic image actually use this representation in spelling.

The results of Reitsma and Vinke (1986) led us to perform the present experiments. As was said before, Reitsma and Vinke found a pseudohomophone effect after nine presentations with a word. The first question posed here was whether the number of presentations necessary for the development of an orthographic image is nearer to three presentations or closer to nine. In their study they used four levels of frequency (0, 3, 9, 18). In our first experiment the same frequency conditions were used with the addition of the levels five and seven. A second question concerned the permanence of the developed representation. To answer this question, a week after the first test a second test took place. A third issue was to what extent the developed representation would be used in spelling. Phrased differently, will a word read more often (and of which it can be assumed that a more or less complete representation has been formed) be spelled better than a word less often encountered. A writing down to dictation test was administered to answer this question. Finally we wanted to know whether the reading of good readers is qualitatively and/or quantitatively different from that of poor readers.

Experiment 2.1

In this experiment we investigated the following issues pertaining to first graders: 1) the effect of visual presentation frequency on the development of orthographic images; 2) the effect of visual presentation frequency of a word on spelling performance; 3) the effect of reading level on the rate of development of orthographic images; 4) the permanence of the orthographic image in reading; 5) the permanence of spelling knowledge.

Method

Materials. From the word list of Kohnstamm et al. (1981) 15 concrete nouns were selected. From these 15 words pseudohomophones were derived. Each pseudohomophone was created by making one or more (to a maximum of three per word) of the following changes in the corresponding word: the 'sch' was changed into 'sg'; the 'au' into 'ou' or vice versa, the 'ij' into 'ei' or vice versa, and a 'd' in a position where it is pronounced as a [t] by a 't'. The mean length of the training words was 6.3 letters. Furthermore, we developed a list of 15 words (similar in length and letter clusters to the training words) that were not practised by the children. These words will be referred to as non-training words. The materials are presented in Appendix A.

Finally 18 sentences were constructed with each training word. These sentences contained one training word each and the remaining words were all taken from the first

three books of the reading curriculum "Veilig leren lezen" (Caesar, 1979). This way a set of 270 (15 times 18) sentences was created. This set of sentences constituted the experimental training materials. Sentences were developed such that they did not constitute a predictive context for the training word. From this set of 270 sentences 5 training lists were made up systematically; each list contained 126 sentences. In each training list three training words appeared three times (Frequency 3), three training words five times (Frequency 5), three training words seven times (Frequency 7), three training words nine times (Frequency 9), and three training words 18 times (Frequency 18). As a result, each training word appeared in each frequency condition.

Procedure. During the training period the children read on three consecutive days (Monday, Tuesday, and Wednesday), one of the five training lists containing the 126 sentences. Thus, each day the children read 42 sentences. The experimenter had each child read out loud the sentences individually.

On the fourth day (Thursday) the children were tested (first measurement). All 15 training words, their 15 derived pseudohomophones and the 15 non-training words were presented once on the monitor of a Macintosh Plus. The children were instructed to name each of the presented words as quickly as they could. One experimenter registered the children's naming times by pushing a key on the keyboard. The reaction time (i.e., naming time) was the time that expired between the presentation of the word on the screen and the moment the child had pronounced the complete word. To make sure that registration of naming times was not influenced by the experimenter, the screen was positioned such that it was invisible to this experimenter. A second experimenter watched the monitor en registered the errors made by the child. We chose this way of response registration, and not for voice-key registration, because the latter procedure requires a degree of discipline not to be expected of children this age. To get used to the task, the children received five words they had encountered before in their reading curriculum. Subsequently, the experimental words were presented in a pseudo-randomised order. The children thus had to read out loud a total of 50 words.

After completion of the naming task a third experimenter presented a spelling test to the children, that consisted of the 15 training words. On each trial the experimenter read a sentence containing one training word. By repeating the training word the children learned which word from the sentence to write down. The experiment lasted about 15 minutes on average.

Exactly one week later, without extra training on the 15 training words in the mean time, the children were tested again on the naming task (second measurement, again on Thursday). The procedure was exactly the same as on the first naming test. The spelling test was also administered once more.

Subjects. From a population of 141 children from Grade 1 a group of good and a group of poor readers were selected (mean age at the time of experimental testing, March: 85 months). Three weeks before the training and the first naming test several pre-tests were administered that were considered important for the selection of these groups. In addition to the reading-decoding test (Caesar, 1975) and the spelling test (Mommers & Van Dongen, 1986), Raven's Standard Progressive Matrices (the sets A, B, C, and D were used), a test for

non-verbal intelligence (Raven, 1958), and the vocabulary test from the RAKIT, a test for verbal intelligence (Bleichrodt et al., 1984) were administered. Assignment of subjects to the two reading levels was based on their scores on the reading test. The group of poor readers (*N* = 26) consisted of children with a score between 15 and 19 (mean: 17.0, *SD* = 1.6), and the group of good readers (*N* = 24) had a score between 29 and 36 (mean: 32.3, *SD* = 2.4). The total experimental group thus included 50 children. Children with a reading score lower than 15 or above 36 were not included in the experimental group, because it was assumed that they did not represent the average beginning reader of Grade 1.

The two experimental groups not only differed in reading-decoding skill, but were also significantly different from each other in spelling ability, $F(1, 48) = 15.2$, $p < .001$. The group of good readers had a mean score of 28.1 (*SD* = 1.6) on the word dictation test, and the poor readers 23.2 (*SD* = 5.9).There were no significant differences between the groups of good and poor readers on the verbal and non-verbal intelligence tests. Therefore, differences that may occur in this study cannot be attributed to a difference in intelligence.

For different reasons (amongst others a school trip), of the 50 children who took part in the first measurement 15 could not participate in the second measurement. Of the 35 remaining children who took part in the second stage 17 were good and 18 were poor readers.

To check for selective dropout, a number of pre-test results of the total group was compared with those of the group of 35 children who took part in both tests 1 and 2. The difference between good and poor readers was also significant in the smaller group, $F(1, 33) = 438.0$, $p < .001$. The groups also differed on the spelling test, $F(1, 33) = 11.58$, $p < .001$. The scores on the verbal and non-verbal intelligence tests were the same for both groups of readers. These results show that the group that took part in both tests was not different from the total group that only participated in the first test. Stated differently, no selective dropout occurred, indicating that the results of the first and second measurement can be compared reliably.

Results

In discussing the findings we will present the results of the first and second measurement conjointly, because they were almost identical. The results of the naming test will be discussed first, followed by those of the spelling test. The abbreviations M1 and M2 refer to the first and second measurement respectively.

Prior to the analyses reaction times based on erroneous responses and outlier naming times were discarded. Erroneous responses were either caused by the subject not pronouncing the presented word correctly or by the experimenter registering the response either too fast or too slow. Naming times were considered to be extremely long when they were longer than the longest naming time of the subject with the slowest average naming time. An extremely short naming time was one shorter than the shortest naming time of the fastest subject.

The mean percentage errors in Measurement 1 due to erroneous pronunciation by the subjects was 12.2%. The mean number of errors of the good readers appeared to be significantly smaller than that of the poor readers (4.7% and 19.0% respectively, $t(35.74) = -5.73$, $p < .001$). In Measurement 2 the mean percentage of pronunciation errors was 9.8%. Again, the good readers made less errors than the poor readers (4.2% and 15.1%

respectively, $t(25.33) = -4.35$, $p <$ 001). In Measurement 1 a total of eight outliers (five extremely short and three extremely long; .3%) were removed from the data set. In Measurement 2 only two extremely long naming times were excluded. As a result of the response screening procedure that was applied, one subject in the first measurement lacked a mean in one of the frequency conditions and thus had to be removed from the analysis. Thus, the analysis of the first measurement was based on the results of 49 subjects and that of the second test on all 35 subjects.

Test stage. A 2 (reading level: good vs. poor) by 2 (stimulus type: training words vs. pseudohomophones) by 5 (presentation frequency: 3, 5, 7, 9, 18) analysis of variance on the subject means showed a significant main effect of reading level (M1: $F(1, 46) = 53.69$, $p < .001$; M2: $F(1, 33) = 36.23$, $p < .001$). The mean naming times in M1 and M2 for the good readers were 2770 and 2581 ms respectively, and for the group of poor readers they were 5066 and 5110 ms respectively. The main effect of stimulus type was also significant in both measurements (M1: $F(1, 46) = 58.99$, $p < .001$; M2: $F(1, 33) = 28.58$, $p < .001$). The differences in naming times between training words and pseudohomophones were significant in all frequency conditions. The mean naming times of the training words in M1 and M2 were 3464 and 3440 ms respectively, and of the pseudohomophones they were 4373 and 4323 ms respectively. The main effect of presentation frequency was also significant in both tests (M1: $F(4, 184) = 2.89$, $p < .05$; M2: $F(4, 132) = 3.69$, $p < .01$). Only on the second test, however, naming times in frequency condition 18 were significantly shorter than in frequency condition 3. This was caused by the group of poor readers exclusively (Newman-Keuls, $p < .01$). The interaction effect between presentation frequency and stimulus type was not significant ($p > .15$). Figure 2.1 presents the results of the first measurement. In fact, no significant interactions emerged on the first test. On the second test a significant interaction showed up between reading level and presentation frequency ($F(4, 132) = 3.36$, $p < .05$). The main effect of presentation frequency was significant in the group of poor readers ($F(4, 132) = 3.57$, $p < .01$), but not in that of the good readers[1]

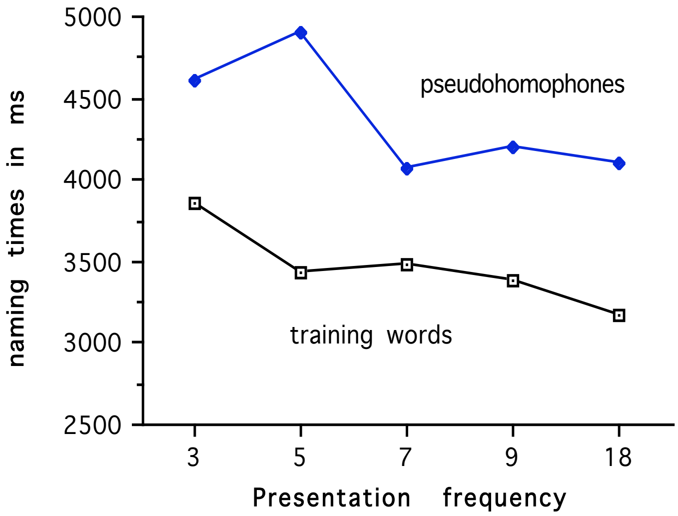

Figure 2.1. Mean naming times of the training words and the pseudohomophones of Experiment 2.1.

A 2 (reading level: good vs. poor) by 6 (presentation frequency: 0, 3, 5, 7, 9, 18) ANOVA on the naming times for the training words only (Frequency 0 contained the words that were not practised) was performed to assess where the largest decrease in naming times occurred. The main effect of reading level was again significant in both measurements (M1: $F(1, 47) = 60.79$, $p < .001$; M2: $F(1, 33) = 47.68$, $p < .001$), and also the main effect of presentation frequency (M1: $F(5, 235) = 22.29$, $p < .001$; M2: $F(5, 165) = 16.93$, $p < .001$). The largest decrease in naming times in both measurements occurred between 0 and 3 presentations (M1: 1543 ms; M2: 1229 ms). In both measurements the mean naming times of Frequency 0 was significantly longer than of all other conditions (Newman-Keuls, $p < .01$). In both measurements the mean naming time of Frequency 18 was significantly shorter than that of Frequency 3 ($p < .05$). A significant interaction effect emerged between reading level and presentation frequency (M1: $F(5, 235) = 3.84$, $p < .01$; M2: $F(5, 165) = 3.11$, $p < .01$) This effect was mainly caused by the fact that the poor readers, especially in the lower frequency conditions, showed a larger decrease in naming times than good readers. The total naming time difference (that is, mean naming times of Frequency 0 minus mean naming times of Frequency 18) was 1513 ms for the good readers in M1 (M2: 1194 ms) and for the poor readers it was 2980 ms (M2: 3104 ms).

Spelling results. Performance on the spelling test was judged as follows. It was established for each word whether an error occurred in the critical part of the word. The critical part refers to the letters or letter clusters: 'd', 't', 'sch', 'au', 'ou', 'ei', 'ij' (critical error). It was also assessed whether errors occurred in the non-critical part of the word (non-critical errors).

A 2 (reading level: good vs. poor) by 2 (error type: critical vs. non-critical) by 2 (presentation frequency: 3, 5, 7, 9, 18) analysis of variance showed a significant main effect of reading level (M1: $F(1, 45) = 16.31$, $p < .001$; M2: $F(1, 32) = 11.87$, $p < .01$). The mean number of errors for the good readers in M1 and M2 was .45 and .44 respectively, and for the groups of poor readers it was .67 in both measurements. The main effect of error type was also significant (M1: $F(1, 45) = 153.02$, $p < .001$; M2: $F(1, 32) = 97.18$, $p < .001$). The mean number of errors in the critical part of the word was .81 in M1 and .80 in M2. The mean number of errors in the non-critical part was .32 in M1 and .31 in M2. In none of the measurements a significant effect of presentation frequency emerged (M1: $p > .40$; M2: $F < 1$). In M1 only the first order interaction between error type and reading level reached a significant level ($F(1, 45) = 4.30$, $p < .05$). This was due to the fact that the good readers made significantly less non-critical errors (.16 vs. .46) than the poor readers ($F(1, 84) = 19.95$, $p < .001$). The difference in critical errors between good and poor readers was not significant (.75 and .87 respectively). In the second measurement the interaction effect between reading level and error type was only marginally significant ($F(1, 32) = 3.62$, $.05 < p < .10$). But the difference between good and poor readers on non-critical errors (.14 and .47 respectively) was significant ($F(1, 59) = 15.18$, $p < .001$), but the difference on critical errors was not (.73 and .87 respectively).

Discussion

The results reported above do not answer the question as to how many presentations of a word are needed to develop an orthographic image. Unlike Reitsma and Vinke (1986), who found a naming time difference between words and pseudohomophones only after nine

presentations (which suggests that only then an orthographic image or word specific knowledge has developed), we already found, - in the analysis on the complete stimulus set - a naming time difference after three presentations of a word. This difference did not become larger with increasing presentation frequency. The analysis on the subset of the stimuli (see Note 1) showed the absence of a difference between words and pseudohomophones in all but one frequency condition (Frequency 5 being the exception). In both cases, however, the effects of stimulus type emerged, but no interaction. The effect of presentation frequency was present on the analysis of the total set, but appeared to be somewhat weaker on the analysis of the subset.

A possible explanation for this result is that already in less than 3 presentations an orthographic image has developed, and that this orthographic image does not change with increasing presentation frequency (up to 18). In that case the expected interaction would only emerge with a presentation frequency of less than 3 (this possibility will be investigated in Experiment 2.2). The question whether the developed representation of the training words was relatively permanent can be answered affirmatively. This can be concluded from the nearly identical findings on M1 and M2.

The spelling results indicate that the frequency with which words have been read before does not affect their spelling. For example, the representation of a word that has been read 18 times is not qualitatively better for spelling than that of a word that was read three times. Moreover, the spelling results in M1 were almost identical to those in M2. This was the case for both the good and the poor readers. Good readers made less spelling errors than poor readers, but this was to be expected, because the group of good readers is also the group of good spellers.

Both the naming and the spelling results indicated that differences between good and poor readers were merely quantitative in nature, not qualitative. Good readers named words almost twice as fast as poor readers, and both good and poor readers made more critical than non-critical errors. The only qualitative difference that emerged between good and poor readers was that good readers made less non-critical errors than poor readers, whereas they had equal numbers of critical errors.

As was suggested earlier, it is possible that the development of the orthographic image took less than three presentations with the word. This possibility will be studied in Experiment 2.2 by investigating whether an interaction emerges between stimulus type and presentation frequency between zero and three presentations. Furthermore, we will investigate where exactly between zero and three presentations the largest decrease in naming times occurs. Naming times of words with presentation frequencies 0, 1, 2, 3, 6, and 9 will be assessed.

In Experiment 2.2 not only the test, but also the training stage was executed on the computer. This enables us to get more information on the development of word reading. Naming times of all words on all presentations during training were registered. It was decided not to present the training words in a sentence, because that would have complicated the registration of the naming times. Because the training sentences of Experiment 2.1 did not constitute a predictive context for the training words, it can be safely assumed that the task used in Experiment 2.1 is comparable to that of reading isolated words.

The reading and spelling behaviour of the good and poor readers in Experiment 2.1 did not show important qualitative differences. We, therefore, decided not to manipulate reading level in Experiment 2.2.

Experiment 2.2

Method

Materials. The word list of Kohnstamm et al. (1981) was again used for the selection of the stimuli. Twelve nouns were selected as training words. The mean length of the training words was 7.0 letters. The procedure as used in Experiment 2.1 was applied here to derive pseudohomophones from the training words. Appendix B presents the training words and the pseudohomophones.

The training materials consisted of six different lists with training words. Each of these lists contained two words that appeared only once (Frequency 1), two words that appeared twice (Frequency 2), two words three times (Frequency 3), two words six times (Frequency 6), and two words nine times (Frequency 9). The remaining two words did not appear at all. The result was that each training word occurred in each condition. Each training list contained 42 words. The test materials (24 stimuli) consisted of a list of all 12 training words and their corresponding pseudohomophones. The distribution of the stimuli within a list was not random, but such that a word's repetitions were evenly distributed over the list, without the order being predictable. The total number of stimuli was 66 (42 training words and 24 test stimuli).

Procedure. The test part of Experiment 2.2 was executed on a Macintosh Plus computer, the same way as Experiment 2.1. In Experiment 2.2, however, also the training stage was performed on the computer. The training stage and the test stage were performed without interruption, and the transition from training to test was not noticeable for the children. One experimenter registered the reaction times of the presented words. A second experimenter evaluated the response of the child. The session started with four well-known words to familiarise the subjects with the task. As in Experiment 2.1, a dictation test was administered after completing the computer part of the experiment, but no second measurement was carried out (see for details the 'Method' Section of Experiment 2.1). The experiment lasted about 10 minutes.

Subjects. From the same population and according to the same criterion, as that of Experiment 1, 36 new children were selected (mean age at the time of experimental testing, May-June: 87-88 months). These children had a mean score on the reading test (23.4; *SD* = 2.6). The lower and upper bound of the scores was exactly between those of the good and poor readers of Experiment 2.1 (lower bound 20, upper bound 28).

Results

De discussion of the results is split up in that of the training results, the test results, and the spelling results. According to the procedure used in Experiment 2.1, extreme response times

were discarded. The number of extreme responses was eight (four extremely short, and four extremely long; .3%).

The replacement of the latencies based on errors was different from that in Experiment 2.1. The mean naming times of words in the training stage was based on the latencies of the correct responses. The naming times of the error responses in the various frequency conditions of the test stage were replaced by the mean naming time of that subject in the corresponding frequency condition of the training stage. The mean percentage of naming errors in the test stage was 8.3%.

Training stage. An analysis of variance on the mean naming times with repeated measures on the frequency variable (frequencies: 0, 1, 2, 3, 4, 5, 6, 7, 8) showed a significant effect of frequency, $F(8, 280) = 56.21$, $p < .001$. Naming times of words in Frequency 0 were significantly longer than those in all other frequency conditions (Scheffé F-test, $p < .01$). Naming times of words in Frequency condition 1 were significantly longer than those in Frequency conditions 5, 6, 7, and 8. The largest decrease in naming time emerged between Frequency 0 and Frequency 1 (1866 ms).

Test stage. A 2 (stimulus type: training words vs. pseudohomophones) by 6 (presentation frequency: 0, 1, 2, 3, 6, 9) analysis of variance revealed that both main effects reached significant levels. The main effect of stimulus type was $F(1, 32) = 19.39$, $p < .001$, indicating a shorter mean naming time of training words (2897 ms) than of pseudohomophones (3712 ms). The main effect of presentation frequency was $F(5, 160) = 6.34$, $p < .001$. The mean naming time of words in Frequency 0 was significantly longer than in all other frequency conditions (Newman-Keuls, $p < .01$). Again the largest decrease in naming times occurred between the Frequencies 0 and 1 (916 ms).

The interaction effect between presentation frequency and stimulus type was not significant ($p > .10$), but the differences between naming times of training words and pseudohomophones in Frequencies 1, 3, 6, and 9 appeared to be (all p's < .05). However, the differences between naming times of training words and pseudohomophones in the Frequency conditions 0 and 2 was not significant[2]. The results are depicted in Figure 2.2.

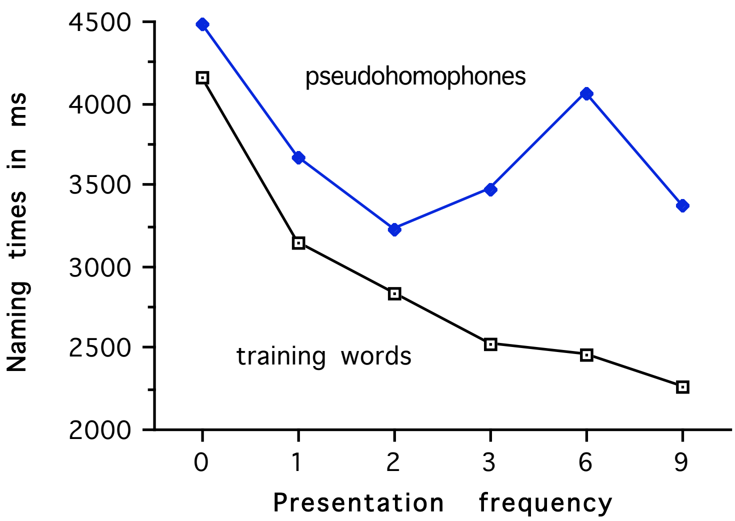

Figure 2.2. Mean naming times of the training words and pseudohomophones of Experiment 2.2.

Spelling results. A 2 (error type: critical vs. non-critical) by 6 (presentation frequency: 0, 1, 2, 3, 6, 9) analysis of variance showed a significant main effect of error type, $F(1, 35) = 199.95$, $p < .001$. The mean number of errors in the critical part of the word was 1.12, and in the non-critical it was .32. The main effect of presentation frequency was not significant. The interaction effect between error type and presentation frequency, however, was significant, $F(5, 175) = 3.11$, $p < .05$. The effect of presentation frequency on critical errors appeared to be significant ($F(5, 175) = 2.28$, $p < .01$), whereas the effect in the non-critical condition was not.

Discussion

As in Experiment 2.1, a pseudohomophone effect emerged - training words were named faster than pseudohomophones - but again no interaction between stimulus type and presentation frequency was found. The interpretation of the results, however, is not without problems, because from the analysis of the naming times of the total set of stimuli it appeared that the pseudohomophone effect already emerged after the first presentation. From the same analysis on a subset of the stimuli (without 'sch' and 'sg' stimuli, see Note 2), however, it could be inferred that the effect did not show up until nine presentations, which confirms the results of Reitsma & Vinke (1986), who also found a pseudohomophone effect after nine presentations. The analysis on the total set of stimuli (including words and pseudohomophones with 'sch' and 'sg' clusters) revealed a substantial, though not significant, difference between the naming times of pseudohomophones and words on the first presentation (words the children had not seen before). This difference was not apparent any more in the analysis on the subset. This latter finding was to be expected, because when a child is presented with a word for the first time it should not make any difference for them how it is spelled (in Dutch: wedstrijd or wetstreit; an English example is: speak or speek). It seems that discarding the 'sch' and 'sg' stimuli led to a more reliable assessment of the effect (but see Experiment 2.3).

The largest decrease in naming time appears to be between zero and one presentations of a word. This agrees with and qualifies the results of Experiment 2.1, in which it was found in the test stage that the largest decrease emerged between zero and three presentations.

The analyses of the spelling results are largely in accordance with those of Experiment 2.1. Again, no effect of reading frequency was apparent on the scores of the spelling tests. But unlike in Experiment 2.1, an interaction was found between presentation frequency and error type: An effect of presentation frequency on the mean number of critical errors showed up, whereas no such effect appeared on the number of non-critical errors. This suggest that this group of beginning readers, with an average score on the reading test, developed a representation during the training stage that was used partly while spelling. It is important to note here, that this group of readers had had three more months of reading instruction than the children from Experiment 2.1. General maturational factors that are related to school skills may have played a role in obtaining this effect.

To substantiate the above conclusion, Experiment 2.3 was executed. The possibility of a contaminating effect of stimulus materials was reduced by excluding pseudohomophone stimuli containing 'sg' clusters. Furthermore, the number of presentations of a word was

dramatically increased (maximum presentation frequency was 30) to try to enforce an interaction between stimulus type and presentation frequency. We decided to use a voice-key to register naming latencies, because of the positive experiences (especially with their abilities to concentrate) with these young children in Experiment 2.2. It should in any case reduce the variance. As in Experiments 2.1 and 2.2 we were again interested in the effect of presentation frequency on spelling performance.

Experiment 2.3

Method

Materials. The word list by Kohnstamm et al. (1981) was used again for the selection of 20 words (12 nouns, three verbs, four adverbs and one adjective). These words will henceforth be called 'training words'. As in Experiments 2.1 and 2.2, pseudohomophones were derived from these training words by changing one or more letters: a 'ch' was replaced by a 'g', an 'ij' by an 'ei' or vice versa, an 'ou' by an 'au' or vice versa, and a 'd' in the word position where it is pronounced as a [t] by a 't'. The mean length of the training words was 6.0 letters. All materials are presented in Appendix C.

The training materials consisted of four lists. In each list five words did not appear at all (Frequency 0), five words appeared once (Frequency 1), five words appeared twice (Frequency 2) and five words appeared 30 times (Frequency 30). Each word appeared in each frequency condition. Each training list contained 165 words. The test materials (a total of 40 stimuli) consisted of all 20 training words and their corresponding pseudohomophones. The distribution of training stimuli and test stimuli within a list was not random. Instead the words were distributed evenly within a list, without their order in the list being predictable. The total number of experimental stimuli was 205 (165 training and 40 test stimuli).

Procedure. This experiment was also run on a Macintosh Plus computer. The training words were presented on the screen one by one, and the children were asked to read them as quickly as possible. The test words were presented immediately after the training words, without the transition being noticeable for the subjects. The naming latencies were registered with a voice-key. Here, the naming time was the time between the stimulus appearing on the screen and the voice time. In Experiments 2.1 and 2.2 the naming time had been the time between the stimulus presentation and the moment the response had been completely pronounced. The experimenter evaluated each response on correctness and voice-key errors. Prior to the experimental task the children had to read ten well-known words to get familiarised with the task (specifically with the voice-key). The experiment lasted about 25 minutes. A dictation test followed the naming task.

Subjects. A group of 20 children from a population (different from the one of Experiments 1 and 2) of 126 children of Grade 1 (mean age at the time of experimental testing, May-June: 87-88 months) was chosen according to their scores on the reading test. The mean score of the 20 children from the selected group was 31.8 (*SD* = 1.6; min: 30, max: 36). The mean score of the selected group on the spelling test (Mommers & Van Dongen, 1986) was 28.9

($SD = 1.2$). The scores on both the reading and spelling tests indicate that the selected group may be considered good readers and spellers.

Results

The naming times of the training and the test stage will be discussed separately. Responses based on errors made by the subject, those due to erroneous registration by the voice-key, as well as extremely long (> 7000 ms) and extremely short (< 200 ms) responses, were discarded.

Training stage. An ANOVA on the naming times of the means of the Frequencies 1 to 30 revealed a significant effect of training frequency, $F(29, 551) = 19.99$, $p < .001$. The largest decrease (487 ms) in naming time emerged between the first (2047 ms) and second presentation (1560 ms). The differences in naming times between the first and second presentation ($F(1, 551) = 39.16$, $p < .001$) and between the second and third presentation ($F(1, 551) = 13.69$, $p < .001$) were significant. The decrease in naming times between the third and the fourth presentation did not reach a significant level.

Test stage. A 2 (stimulus type: training words vs. pseudohomophones) by 4 (presentation frequency: 0, 1, 2, 30) ANOVA revealed that both main effects reached significant levels. The main effect of stimulus type ($F(1, 19) = 37.76$, $p < .001$) showed a shorter mean naming time for training words (1437 ms) than for pseudohomophones (2181 ms). A post-hoc analysis, based on the significant main effect of presentation frequency ($F(3, 57) = 7.78$, $p < .05$), revealed that the difference in naming times between Frequency 0 (2076 ms) and Frequency 30 (1509 ms), between Frequency 1 (1912 ms) and Frequency 30, and between Frequency 2 (1739 ms) and Frequency 30 all differed significantly from one another (Newman-Keuls, $p < .01$, $p < .01$, and $p < .05$, respectively). In this case, the largest decrease in naming time was found between two and 30 presentations (230 ms). The interaction between stimulus type and frequency did not reach significance ($p > .25$). However, it is important to note that a significant difference occurred between the naming times of training words (1720 ms) and pseudohomophones (2433 ms) in Frequency condition 0 (Newman-Keuls, $p < .01$). Frequency condition 0 contained words the children had not seen before. The results of Experiment 2.3 are presented in Figure 2.3.

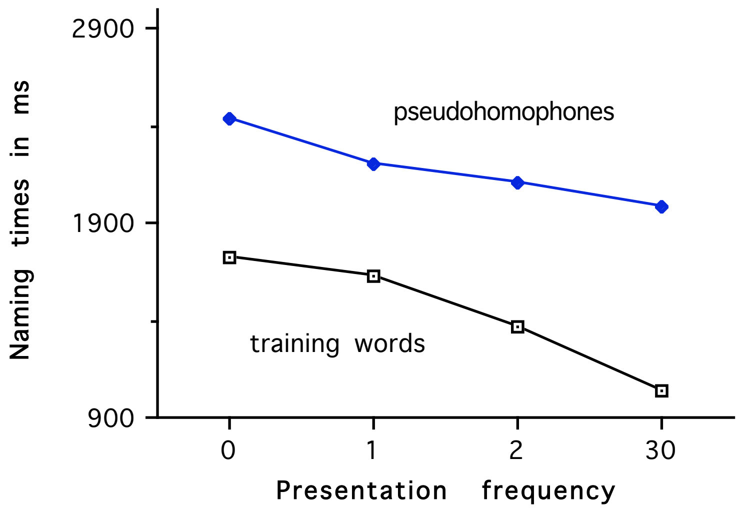

Figure 2.3. Childrens' naming times of the training words and pseudohomophones in each frequency of Experiment 2.3.

Spelling results. As in Experiments 1 and 2 the errors on the dictation test were evaluated according to whether they were critical or non-critical. A 2 (error type: critical vs. non-critical) by 4 (presentation frequency: 0, 1, 2, and 30) ANOVA on the spelling errors revealed that both main effects were significant. The interaction effect was not significant. The main effect of error type ($F(1, 19) = 19.57$, $p < .001$) showed that the children made more critical errors (.45) than non-critical errors (.25). Although also the main effect of presentation frequency was significant ($F(3, 57) = 2.83$, $p < .05$), a post-hoc Newman-Keuls analysis did not lead to significant differences between the frequencies conditions. The proportions of errors in the Frequency conditions 0, 1, 2, and 30 were .39, .40, .31, and .29 respectively.

Discussion

Again, a frequency and pseudohomophone effect emerged, but no interaction between frequency and stimulus type. As in the two foregoing experiments, the largest decrease in naming time showed up between zero and one presentations. The difference between naming times of words and pseudohomophones on the first presentation indicates that, despite the fact that the words were new for the children, they had already acquired sufficient orthographic knowledge to be able to decode words faster than pseudohomophones. Even though the pseudohomophones are orthographically legal, the letter clusters within these letter strings do not appear as frequently in the Dutch writing system as those within actual words. This is shown from the results of a post-hoc analysis on the mean cumulative positional trigram frequency (CPTF) of the experimental stimuli (Rolf & Van Rijnsoever, 1984). The mean CPTF of the training words was 759 and that of the pseudohomophones was 184, $F(1, 38) = 13.62$, $p < .001$.

The results of the spelling test are partly in accordance with those of Experiments 2.1 and 2.2. As in the two earlier experiments the children made more critical than non-critical errors. Unlike in the Experiments 2.1 and 2.2, however, an effect of presentation frequency

on the mean number of errors emerged. The children from this experiment, who may be considered good readers, had had the same amount of education as the children from Experiment 2.2, who may be considered medium readers, but the children from Experiment 2.1, in which both good and poor readers participated had had three months less education. This suggests that number of months of schooling in reading and spelling, and the reading level are both important factors for the usefulness for spelling of the representation developed during reading.

General Discussion

A number of questions was raised in the 'Introduction' of this chapter. The experiments reported here were designed to find an answer to those questions. The first question was how many presentations a beginning reader does need to build an orthographic image, which would enable her or him to switch from indirect to direct reading. This question cannot be answered unequivocally: In none of the three experiments an interaction between presentation frequency and stimulus type emerged. However, in all cases an effect of presentation frequency showed up - a word that had been presented often was read relatively fast .

A pseudohomophone effect was also always apparent - words were read faster than pseudohomophones. From the results of the first experiment it could not be decided whether or not an orthographic image had been established. The reason is that when considering the total set of stimuli it seemed that after three presentations word-specific knowledge had already been acquired, but the analysis on the subset of the materials, that is, without the words containing 'sch' and sg', did not support this conclusion. In Experiment 2.2, given the total set of stimuli, it even seemed that already after the first presentation an orthographic image had been established, whereas the same analysis on the subset seemed to indicate that this only happened on the ninth presentation. The last experiment could not answer the question either, because the children who participated in this experiment already possessed so much orthographic knowledge that already at the first presentation they read the new words faster than the pseudohomophones derived from these words. The fact that in all three experiments training resulted in faster naming times (frequency effect) indicates that as a result of frequent presentations indeed a representation has been established (although the nature of this representation is hitherto unclear).

The children who took part in Experiment 2.3 had had the same amount of schooling in reading and spelling as the ones in Experiment 2.2, and both these groups had had more (by three months) education than the children from Experiment 2.1. The children in Experiment 2.2, however, had a medium score on the reading test, whereas the children from the last experiment may be considered good readers (as the good readers of Experiment 1). It is possible that these factors contributed to the slightly differential findings of the experiments.

An explanation for the failure to find an interaction between stimulus type and presentation frequency is that it is perhaps inappropriate to assess direct reading by looking at the subjects' performance to pseudohomophones. The reason is that in all three experiments it appeared that, notwithstanding the fact that pseudohomophones were not

practised, they were affected by the presentation frequency of the corresponding words. Pseudohomophones are clearly different from their corresponding words, but to a large extent they are similar to these words. Transfer may have taken place. If that was the case, a different operationalisation of the question is desirable.

A second question that was investigated was the permanence of the established representation. The results of the first experiment suggested that after a week nothing of this representation had been lost. Whether this representation remains stable after this first week cannot be answered, but is also difficult to determine, because as time progresses the chances will increase that the child will see and practice the words.

Besides the effect of frequency on reading performance, the question was raised to what extent the representation established during reading would be useful for spelling. The first two experiments showed that the children did not show superior spelling performance after having read the words relatively often. That is, having read a word 18 times did not cause them to spell the word better than having read it only three times. The third experiment did show an effect of presentation frequency. But this effect was hardly substantial, because it could not be decided in what frequency condition the children's spelling performance started to improve. The cause of finding an effect of presentation frequency in this third experiment and not in the other two experiments, may be that the children of the first experiment had had three months less of reading and spelling instruction than the children of Experiments 2.2 and 2.3. For children of better mean average reading level (Experiment 2.3), three more months of instruction may have improved the skill of the children from Experiment 2.3 to such an extent that they were capable of using the representation that was induced by reading.

It is clear, however, that all children predominantly use a phonologic strategy in spelling, because more critical than non-critical errors occurred. This was the case both for good and for poor readers/spellers. This result coincides with that of Waters et al. (1985; 1988). Good readers/spellers made less errors than poor readers/spellers, but both groups made more critical than non-critical errors. Another finding that confirms earlier results of Frith (1980) is that poor readers/spellers made relatively more phonologically incorrect errors than good readers/spellers.

To summarise briefly, the above permits us to draw the following conclusions. First, it is still unclear after how many presentations of a word an orthographical image has developed. Future experiments will have to clarify whether the way this experimental question was operationalised is to blame for that. We may conclude, however, that a representation has been formed as a result of frequent word presentation. Otherwise, it seems difficult to explain the robust frequency effect. Secondly, the permanence of the representation that has been formed, of whatever kind it may turn out to be, seems fairly large. Thirdly, we found quantitative, but no qualitative differences between good and poor readers. Finally, it appeared that the representation induced by reading was useful for spelling only at the end of Grade 1. This finding will also be focused on in the future.

3

DIFFERENTIAL EFFECTIVENESS OF READING AND NON-READING TASKS IN LEARNING TO SPELL*

Summary

The effectiveness of a reading and three non-reading tasks: copying, problem naming and oral spelling, in learning to spell was investigated. The results on a spelling test following the training phase, indicated that words practised in the oral-spelling condition were spelled better than words practised in the copying and the problem-naming condition, and that words practised in the three non-reading tasks were spelled better than words practised in the reading task.

In psycholinguistic research spelling never received the attention reading did and reports of studies on the effectiveness of direct spelling instruction are hard to come by. The emphasis on reading may have two major causes. One is mainly technical: It is easier to execute experimentally well controlled studies on reading than on spelling. The second pertains to the assumption that reading and spelling are closely related (Ehri, 1980) and that facts established in reading research also hold for spelling.

A number of observations actually suggest that spelling is not simply the reverse of reading and that it is not the case that what one can read one can spell and vice versa. Classroom experiences tell us that spelling is more difficult to master than reading, and various empirically established facts show the discrepancy between reading and spelling skills. First, the correlation between reading and spelling scores is not perfect, but varies between .50 and .80 (according to Malmquist, 1958, in Frith, 1980). Secondly, one can easily find children who are good readers but poor spellers, whereas children who are poor readers but good spellers are hardly found (Frith, 1980). Thirdly, experiments reported in Chapter 2 indicated that beginning readers did not spell words they had read 18 times any better than words they only had read three times.

The conclusion from the above must be that reading and spelling, although closely related, are not each other's reverse, and that reading might thus not be the best way to learn to spell. There are only a few experimental studies in which reading as a spelling-instruction method is compared with other such methods. We will discuss these in the following sections.

A Dutch study by Van Doorn-Van Eijsden (1984) investigated the effectiveness of writing as compared to reading in a group of subjects from Grade 5. One group was instructed to read the stimuli (words and pseudowords), and a second group had to copy

---

* This chapter is published in F. Satow and P. Gatherer (Eds.). *Literacy without Frontiers*: Proceedings of the 7th European and 28th United Kingdom Reading Association Annual Conference at Heriot-Watt University Edinburgh, July, 1991 (pp. 279-289). Widnes: Chesire. It is co-authored by A. M. B. De Groot.

them on paper. The copy group was significantly better than the reading group on a post-test that assessed the subjects' spelling skill on both types of stimuli. The reason why the copy group performed better is not clear. It could be the motor aspect involved (development of an "integrated movement sequence"; Lashley, 1951; see also Smith, 1973), or it might be that the subjects in the copy group read with 'full cues' instead of 'partial cues' more often than the subjects in the reading group did (Frith, 1980).

From the work of Cunningham and Stanovich (1990) with Grade 1 subjects one would have to conclude that the motor activity involved in the copying task was the critical aspect. They compared the effectiveness of training words using handwriting, letter tiles, or the computer keyboard. The handwriting activity was superior in learning to spell, even when the test situation was congruous with letter and keyboard usage.

However, Sears and Johnson (1986), testing a group of subjects from Grades 4, 5 and 6, did not find superior spelling knowledge in the copying condition as compared to a treatment that involved visualising and the use of a computer. But because their sample consisted of subjects much older than those tested by Cunningham and Stanovich, these results might not be commensurable.

Two studies examined the effect on spelling performances of committing letters or letter strings to memory. Murphy, Hern, Williams, and McLaughlin (1990) compared a daily copy, cover, and compare approach with a traditional one (pre-test on Monday, followed by various activities of no clear relevance to spelling during the week, and a post-test on Friday). In the copy, cover, and compare approach children were instructed to look at a word carefully, copy it, then cover the word and try to write it down (by heart), and finally compare the product with the original word. When an error was made the process had to be repeated until the word was spelled correctly. Subsequently, the procedure was applied to the next word. The copy, cover, and compare approach seemed to lead to higher scores on a spelling test (no tests of significance were performed on the data).

An experiment by Roberts and Ehri (1983) showed that children from Grade 2 who learned the pseudowords, and later had to reproduce the letters by heart, acquired a superior representation for spelling than those just rehearsing or naming the letters while being allowed to look at the words. The authors assumed that the former strategy leads to the formation of an orthographic image. A major problem of both the study of Murphy et al. (1990) and that of Roberts and Ehri (1983) is that it is unclear whether the children in the conditions that led to superior spelling skills had not actually had more training in the words. Therefore their results may not be conclusive.

Research investigating which spelling strategies are spontaneously used by having the subjects "think aloud", indicates that neither children nor adults apply very effective means to remember the correct spelling of a word. Ormrod and Jenkins (1989) found that subjects (undergraduates, subjects of Grade 3/4 and 7/8) mainly used word pronunciation and letter rehearsal, but these turned out not to be very effective as they did not significantly correlate with post-test spelling scores. The only effective strategy, used by a small minority of 14% in the group of undergraduates, was overpronunciation, that is, pronouncing the word such that it reflects the actual spelling more closely (/ai/, /s/ /le/ for "aisle", or /ser/, /ge/, /ant/ for "sergeant").

The results of the above studies are inconclusive with respect to the role of the motor aspect in spelling and the effectiveness of earlier reading on subsequent spelling of the

words. It is also not clear whether letter rehearsal like in the study of Roberts and Ehri (1983) is a better way to teach spelling than copying is. Moreover, it is still an open question whether it simply is the sheer amount of time invested in the task that causes superior spelling skills.

The present experiment aimed at clarifying these issues. Young beginning readers and spellers were presented with words they had never seen prior to the experiment and that were likely to cause spelling mistakes because the correct spelling is contingent and can therefore not be derived. To assess the effectiveness of copying and reading, these spelling-instruction methods were included in the experiment, together with two methods that emphasised the rehearsal of letters or letter clusters. By comparing scores on a spelling test following the training session with the time it took to complete the session it is possible to ascertain whether it is merely the amount of time that is responsible for the improvement in spelling skill, or whether it is the quality of processing that causes this instead. Testing the subjects' spelling skills after training both with a dictation and a forced-choice task allowed the assessment of differential ease of spelling production and spelling recognition.

An additional goal, of practical importance, was to develop spelling-instruction methods that are not only theoretically interesting but are also of relevance to the classroom and are easily applicable.

Experiment 3.1

Method

Materials. Twelve Dutch words (five nouns, five adverbs and two verbs) were chosen from the word list by Kohnstamm et al. (1981) to serve as the experimental materials. The mean length in letters was 6.8, with the shortest word having five letters and the longest nine (*SD* = 1.36). Appendix D presents the experimental materials.

The training materials consisted of four different lists. In every list six words made up the reading condition and six the non-reading condition. Three words of both the reading and the non-reading condition were trained twice (Frequency 2) and three words six times (Frequency 6). All four lists thus contained 24 words in the reading and 24 in the non-reading condition. In every list a training word only appeared in either the reading or the non-reading condition. In constructing four lists this way, every word appeared in every frequency and in both conditions. All words were printed on a A4-size piece of paper with one word per line, one side containing the reading stimuli and the other the non-reading stimuli.

Procedure. A child was assigned to one of the three non-reading conditions, namely: 'copying', 'problem naming', or 'oral spelling'. Every child read all 24 words of one of the lists and practised 24 words according to one of the non-reading conditions. In the copying condition a child simply had to copy the 24 words from one of the lists onto a note book, using a new piece of paper for every new word. In the problem-naming condition a child was instructed to explicate the spelling ambiguities in every word, indicated by underlining of the ambiguous parts. In the Dutch language there are four clear cases of ambiguous phonemes. The schwa can be written by an 'e', 'u', 'ij' and 'o', [ɑu] by 'ou(w)' and 'au(w)', [g]

by 'g' and 'ch', and a final [t] by 't' or 'd'. In the oral-spelling condition the subject was asked to read the word aloud and then spell it aloud by heart. Because children in Grade 1 tend to analyse a word by naming the phonemes instead of the actual letters, the experimenter has to make sure that every ambiguous phoneme was clarified by the subject. Feedback was provided during the training. Half of the subjects participated in the reading condition first. The other half performed the non-reading task first. The time it took the subjects to complete both the reading and the non-reading task was measured separately with a stopwatch.

The training stage was followed by a test stage, which consisted of two parts. All subjects performed a dictation task first, followed by a forced-choice spelling task. In the dictation task, the experimenter read all twelve trained words one by one and the child was asked to write them down. The forced-choice task was presented on the screen of a Macintosh Classic personal computer. The child was presented with two spellings of a word, one was the correct spelling and the other contained a phonologically correct spelling error (an example: 'blauw' and 'blouw'; the first one is the correct spelling for the Dutch word 'blue'). By using the mouse the child could indicate (by clicking on the appropriate place) which of the two she or he thought was the correct spelling. The responses were registered by the computer. A complete experimental session took between 15 and 40 minutes.

Subjects. From a population of 246 children of Grade 1, a sample of 57 children was drawn with an average score (27.3; SD = 4.8) on the reading test (Caesar, 1975). The mean score on the spelling test (Mommers & Van Dongen, 1986) was 27.9 (SD = 2.3). Scores on the non-verbal (Raven, 1958) and the verbal intelligence tests (Bleichrodt et al., 1984) were also known. The three experimental groups were made up such that they did not differ significantly from each other on any of the four tests (all F's < 1), but the group was divided up into a group of good spellers (mean 29.5; SD = .51; N= 26) and a group of poor spellers (mean 26.6; SD = 2.3; N= 31). A group of 26 children (from one classroom) served as the control group. These children simply took the spelling test without being trained in the selected material. The mean scores on the spelling test and on both the verbal and non-verbal intelligence tests of this group did not differ significantly from those of the experimental groups. The mean score on the reading pre-test of the control group was significantly higher than the one of the experimental group (32.9 and 27.3 respectively, $F(1, 76) = 6.23$, $p < .05$).

Results

Training stage. In both the copying condition and the oral-spelling condition it was possible to register the total number and the diversity of errors (that is, the number of different words in which an error was made) occurring during training. The mean number of errors was 4.16 (SD = 2.3) in the oral-spelling condition and it was 2.63 (SD = 4.5) in the copying condition, but this difference was not statistically significant, $F(1, 36) = 1.72$, $p > .10$. The mean diversity score in the oral-spelling condition (2.63; SD = 1.1) was significantly higher than in the copying condition (1.10; SD = 1.3), $F(1, 36) = 14.70$, $p < .001$. Correlations between scores on the spelling test and number of errors (r = -.15) or diversity of errors (r = -.16) were both low and non-significant.

A 3 (non-reading condition: copying vs. problem naming vs. oral spelling) by 2 (task: reading vs. non-reading) ANOVA of the time-on-task data showed significant main and interaction effects. The effect of the non-reading condition ($F(2, 54) = 30.85$, $p < .001$) revealed that the problem-naming (5.43 min) task took significantly shorter (Newman-Keuls, $p < .01$) to complete than the copying (13.03 min) and oral-spelling (11.19 min) tasks. The time to complete the reading task (1.21 min) was significantly shorter than to complete the non-reading task (8.41 min); $F(1, 54) = 538.31$, $p < .001$). The significant interaction effect ($F(2, 54) = 53.92$, $p < .001$) between non-reading condition and task indicated no differences in the time to complete the reading task for the different non-reading conditions, but a highly significant difference between the non-reading conditions themselves, with the problem- naming task taking less time than the copying and oral-spelling tasks, $F(2, 54) = 42.94$, $p < .001$.

Test stage. Two statistical analyses were performed. In the first one, only the results of the dictation test were analysed. In the second analysis the results of the dictation test were compared with those of the forced-choice test.

The responses of the subjects on the dictation test were evaluated on the basis of correctness of target-letter(s) or target-clusters in the word. A 3 (non-reading condition: copying vs. problem naming vs. oral spelling) by 2 (spelling level: good vs. poor) by 2 (task: reading vs. non-reading) by 2 (frequency: 2 vs. 6) ANOVA was performed on the mean number of target errors. All four main effects were significant. The effect of spelling level revealed that good spellers made less errors (.55) than poor spellers (.98), $F(1, 51) = 24.05$, $p < .001$. The main effect of task showed that more errors were made in the reading condition (.96) than in the non-reading condition (.57), $F(1, 51) = 51.23$, $p < .001$. The main effect of frequency ($F(1, 51) = 4.07$, $p < .05$) indicated that subjects made more errors in words that were practised twice (Frequency 2: .82) than in words practised six times (Frequency 6: .71).

The main effect of non-reading condition was, $F(2, 51) = 6.12$, p < .01. According to a Newman-Keuls analysis, subjects in the oral-spelling condition (.56) made significantly less errors than those in the copying condition (.91; $p < .01$), and those in the problem-naming condition (.83; $p < .05$). No significant difference was apparent between the coping and problem-naming conditions.

Only one interaction effect was significant and one marginally so. The one between non-reading condition and spelling level was significant, $F(2, 51) = 5.15$, $p < .01$. This was mainly due to the fact that the poor spellers in the copying condition made significantly more errors (1.32) than those in any of the other conditions (Newman-Keuls, $p < .05$). The poor spellers in the problem-naming condition had also made significantly more errors than the good spellers in the copying and the oral-spelling conditions (Newman-Keuls, p < .05). Figure 3.1 summarises the results of this analysis. The interaction between task and frequency ($F(1, 51) = 3.81$, $.05 < p < .10$), although only marginally significant, indicated that all means differed from each other significantly (Newman-Keuls, $p < .01$) except those of the two reading conditions (see Figure 3.2).

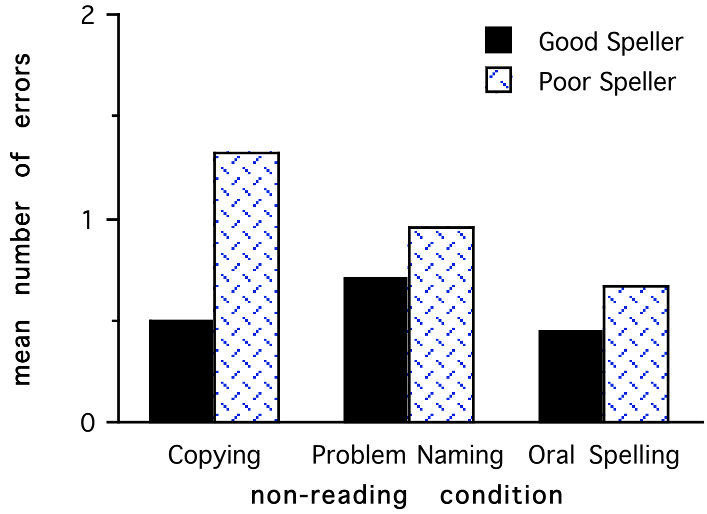

Figure 3.1. Mean number of target errors of good and poor spellers in the three conditions of Experiment 3.1.

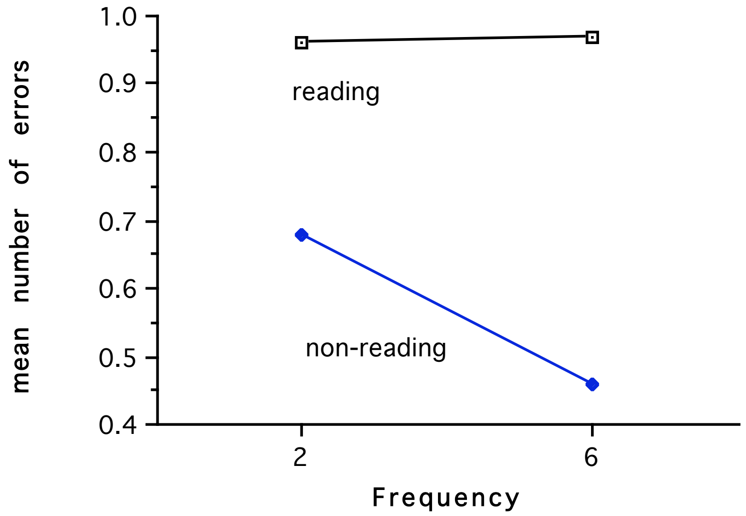

Figure 3.2. Mean number of target errors of all subjects in the reading and non-reading conditions of Experiment 3.1.

A comparison of the mean error scores on the spelling test (of the stimuli in the reading condition only) of the subjects of the experimental group ($N = 57$) with those of the control group ($N = 26$) showed a statistically significant result, $F(1, 81) = 6.48$, $p < .05$. The mean error score of the control group was higher (1.24) than that of the experimental group (.96).

In the analysis in which the results of the dictation test (mean number of target errors) were compared with the results of the forced-choice test, the mean error scores per subject on the forced-choice test were adjusted by adding .50 to their means. It would be trivial to compare the actual means, because of the fact that a subject in the forced test already has a chance of 50% correct on the basis of guessing.

A 3 (non-reading condition: copying vs. problem naming vs. oral spelling) by 2 (spelling level: good vs. poor) by 2 (type of test: forced-choice vs. dictation) analysis of variance revealed that two main effects and all first and second order interactions were significant. The significant effects were, again, those of non-reading condition and of spelling level (already discussed above). The main effect of type of test was marginally significant ($F(1, 51) = 3.84$, $.05 < p < .10$), with the mean error score on the forced-choice test (.69) being smaller than the one on the dictation test (.76). Two relevant interaction effects will be discussed here. One of them was the interaction between non-reading condition and type of test, $F(2, 51) = 4.90$, $p < .05$. This interaction was caused by the fact that the numbers of errors of the three non-reading conditions on the forced-choice test did not differ significantly, but on the dictation test they did, $F(2, 51) = 6.25$, $p < .01$. Figure 3.3 presents the results. The second relevant interaction was the one between spelling level and type of test, $F(1, 51) = 15.28$, $p < .001$. It showed that the poor spellers made more errors on the dictation test than on the forced-choice test ($F(1, 51) = 19.89$, $p < .001$), whereas good spellers did not differ on the factor type of test.

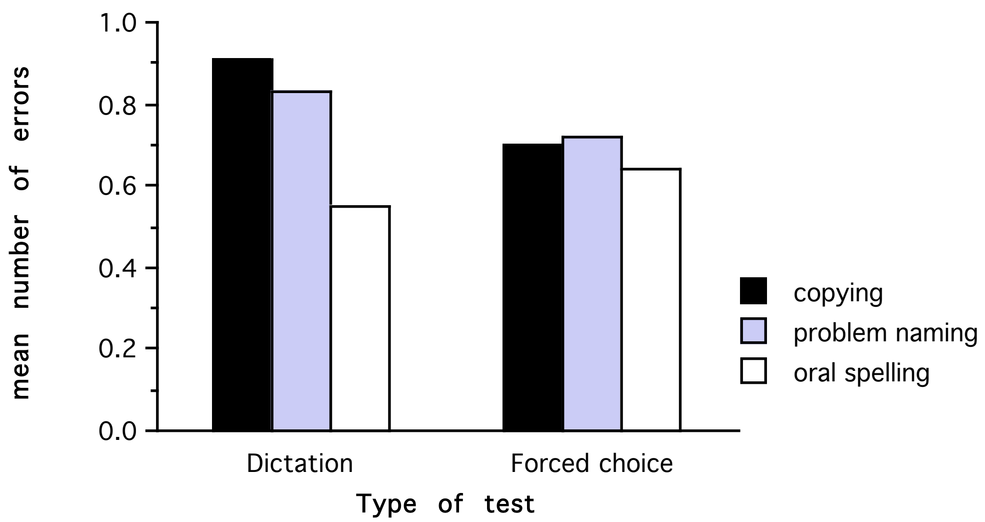

Figure 3.3. Mean number of errors of all the subjects on the two spelling tests of Experiment 3.1.

Discussion

The main purpose of this study was to gain insight into the effectiveness of four spelling-instruction methods. Before attending to this matter the results of the training stage will be discussed.

It does not appear to be more difficult to perform the oral-spelling task than the copying task, because the number of errors the children made during training was the same in both tasks. However, copying does lead to a lower diversity score than oral spelling does. That is, when an error is made on a particular word the subjects in the copying condition tend to make this same error again when copying this word the next time, while in the oral-spelling condition subjects made as many errors as in the copying condition, but these errors concerned a larger variety of words. Copying, unlike oral spelling, could thus lead to learning an incorrect spelling of the word.

From the results of the time-on-task variable it is clear that reading is the fastest way to process words. The data also showed that the oral-spelling and copying tasks take considerably more time than the problem-naming task. This result in itself is not interesting, but it is of major importance when the results of the spelling test are also taken into account. Subjects in the oral-spelling condition showed superior spelling skills on the dictation test after training. The performance of the children in the copying and problem-naming conditions did not differ from one another. If we take into account the time-on-task variable, then copying, particularly in the case of poor spellers, appears to be the least efficient method: It takes more time to finish the copying than the problem-naming task, whereas they are equally effective when considering the scores on the dictation test. This result is in accordance with that of Roberts and Ehri (1983). They found superior spelling knowledge (for pseudowords) for children who trained committing letters to memory. Our findings also clearly indicate that learning the spelling without being able to see the correct spelling (oral spelling) is more effective than just explicating the ambiguous letters or letter clusters (problem naming).

Our experiment permitted us to answer the question as to the role of the motor aspect in spelling (e.g., Van Doorn-Van Eijsden, 1984). It is evident from our study that the actual writing component is not essential in the process of learning to spell, because copying was less effective than oral spelling, and it was not more effective than problem naming. This fact sheds a different light on the results of Cunningham and Stanovich (1990). From their finding that writing (that is, copying) beat the use of letter tiles or the computer keyboard they concluded that its motor aspect caused the superior spelling performance. An explanation more in agreement with our results is that the handwriting condition led to superior performance because the focus of attention was directed more on processing the word itself than in the conditions with letter tiles and keyboard. These last two conditions both involve a search for the right letter among a large number of distracting ones. The focus was probably not on the structure of the word, but on each individual letter. To look for every single letter might have become a goal in itself, while the overall goal (easier to retain with handwriting) received less attention. It may thus be that not the differential motor aspect involved in their tasks caused their result but the fact that copying resembles the essential aspect of the spelling process more than choosing the correct sequence of letters does.

Reading as a means to learn the spelling of words is the least effective method. The spelling performance of the experimental group after reading the stimuli was better than that of the control group, who had not seen the stimuli before. But the effect of reading on spelling is not very large, because words that were read six times were not spelled any better than words that were read only twice. Practising the words six times in the three non-reading conditions, however, <u>was</u> more beneficial for the spelling results than practising the words only twice.

The good spellers outperformed the poor spellers on both the dictation test and the forced-choice test, suggesting that pre-experimental spelling skill is an independent and powerful factor that accounts for differences between groups (see also Roberts and Ehri, 1983). It is interesting to note that the type of spelling test, dictation or forced choice, determines level of spelling skill differently for poor spellers, but not for good spellers. The poor spellers' spelling skill was higher on the forced-choice test than on the dictation test,

whereas good spellers' spelling skill was similar on both tests. It seems that good spellers are as capable to produce as to recognise the correct spelling, whereas poor spellers are better recognising the correct spelling than producing it.

In sum, oral spelling for a group of beginning literates with average reading and spelling scores is the most effective instruction method (out of those studied here) to learn to spell, whereas reading is the least effective. Although copying, the most common practice at school, leads to better spelling than reading does, it is actually a very inefficient means (especially for poor spellers) to learn to spell, because it takes up quite a lot of time. When the time aspect is taken into account, problem naming is preferable to copying.

4

BEGINNING SPELLING: PREVALENCE OF THE PHONOLOGIC STRATEGY IN SPELLING AND THE LIMITED EFFECT OF READING AS A SPELLING-INSTRUCTION METHOD*

Summary

This article focused on spelling-instruction methods and spelling processes in beginning spellers (students from Grade 1). After a short training, the students predominantly used a phonologic way of spelling and did not convert to a lexical strategy. This conclusion was based on the results of two experiments, in which the subjects made proportionally more phonologically correct than phonologically incorrect errors. Of the spelling-instruction methods studied here reading appeared to be the least effective one. The two non-reading instruction methods, problem naming and word composition had the same effects on spelling performance, but were both superior to reading as instruction method. Problem naming appeared to be more efficient than word composition, because it took longer for the subjects in the word-composition condition to acquire the same spelling level as those in the problem-naming condition.

Spelling turns out to be more difficult than reading. This asymmetry (or was it assymetry) is not restricted to the acquisition of reading and spelling, because even skilled language users read more words correctly than they are able to spell. In dyslectics spelling problems also prove to be more persistent than reading problems. Whether spelling is a skill that is per se more difficult to master than reading is difficult to establish, because generally in education spelling receives less attention than reading, and learning the spelling of a word will mostly occur after the student has shown to be able to read the word.

In the present article the following issues on learning to spell will be investigated. The first issue concerns the development of the spelling process, that is, has the spelling process of beginning spellers changed from a phonologic strategy into a lexical strategy (reading out the orthographic image), after participation of subjects in a spelling training? The assumption is that the spelling process of beginning spellers is characterised by sequential recoding of phonemes into the corresponding graphemes, whereas the mature spelling process is mainly characterised by reading out the orthographic image from the mental lexicon (e.g.: Assink, 1981; Frith, 1980; Sloboda, 1980; Tenney, 1980, but see Assink, Van Doorn-Van Eijsden & Van Hees, 1982; Kreiner, 1992). The second issue pertains to the effectiveness of several spelling-instruction methods, and the level of stimulus processing will receive special attention. The third question is whether different spelling-instruction

* This chapter is based on an article co-authored with M. Van Leerdam (1993), Aanvankelijk spellen: de dominantie van de verklankende spelwijze en de geringe effectiviteit van lezen als spelling-instructie methode. *Pedagogische Studiën, 70*, 28-45.

methods already have differential effects at an early stage of reading and spelling instruction.

That reading and spelling are somehow related is not disputed, but to what extent they are the reverse of each other (as is often assumed) is a question that remains unsolved in the literature and will not be settled in this study either. Nevertheless, the 'dual-route model' of visual word recognition provides a theoretical framework for explaining spelling performance. According to the dual-route model, reading in an alphabetic writing system can occur in two distinct ways. One way to read a word (mainly employed by beginning readers) is sounding out the graphemes of a word. This process is referred to as the phonologic reading route or indirect reading. The latter term brings out the contrast with the second way a word can be read, namely, via the direct route (also called the lexical reading route). This route involves that a direct match is made between the written word and a mental representation (the 'orthographic image'). Several presentations with a word are needed to enable a reader to use the lexical strategy. It is assumed that skilled readers predominantly use the direct route. Only when they are presented with low-frequency words will they convert to a phonologic strategy (Seidenberg, Waters, Barnes, & Tanenhaus, 1984).

If reading and spelling are opposite processes, then it seems only natural to assume that there are also two ways to spell a word. A phonologic or indirect strategy of spelling would involve a successive conversion of each phoneme into its grapheme. This seems relatively straightforward, but it is not, because a large number of phonemes map onto more than one grapheme (a Dutch example: the phoneme [ɛi] can be represented by the graphemes 'ij' and 'ei'; an English example: the phoneme [iː] can be represented by the graphemes 'y', 'ey', 'e', 'ee', and 'ea'). A grapheme to phoneme ambiguity is also present, but to a lesser extent (an example that holds for both English and Dutch: the grapheme 'c' can represent the phonemes [s] and [k]). This asymmetry between grapheme-phoneme conversion and phoneme-grapheme conversion has been noticed before by Waters, Bruck, and Seidenberg (1985, p. 513). Analogous to the lexical reading strategy, it should be possible to spell familiar words by reading out their spellings from their orthographic images in lexical memory. The foregoing allows the inference that, because of the ambiguity in the selection of graphemes, the phonologic strategy is rather error-prone. On the other hand, a strategy in which the letters of a word are read out from memory is probably less error-prone, unless the representation is incomplete or contains errors.

To find out whether after a spelling training beginning spellers use a lexical or a phonologic strategy, an analysis will be made of the type of spelling errors they produce. Spelling errors can be divided in errors which are phonologically correct, and errors which are phonologically incorrect. Phonologically correct spelling errors are errors in which the phoneme is represented by an incorrect grapheme, but in which its sound is retained; in other words, a pseudohomophone emerges (a Dutch example: 'haut' instead of 'hout'; an English example: 'leef' instead of 'leaf'). A phonologically incorrect error occurs when the chosen grapheme is phonologically dissimilar from the one intended (a Dutch example: 'walt' instead of 'valt'; an English example: 'kired' instead of 'tired'). If spelling occurs by reading out the mental representation, the proportion of phonologically correct errors should not exceed the proportion of phonologically incorrect errors. If, on the other hand, the proportion of phonologically correct errors surpasses the proportion of phonologically

incorrect errors, then we would have an indication that the population under investigation predominantly uses a phonologic strategy (see also: Frith, 1980; Jorm, 1981; Nelson & Warrington, 1974).

The second issue is whether a differential effect of reading and non-reading tasks as spelling-instruction methods comes about in beginning spelling. The spelling-instruction methods under investigation differ in their depth of processing words, that is, some of the instruction methods require deep and some more shallow processing. We assume, that deep processing of a word leads to the development of a more or less complete representation, whereas shallow processing does not. It is also assumed that only when a word's representation is complete a lexical spelling strategy can be effective. Thus, the question that emerges from the above is, whether deep processing induces a lexical spelling strategy more easily than does shallow processing. In statistical terms, an interaction, thus may be expected to occur between the variables spelling-instruction method (see below) and error type (phonologically correct versus phonologically incorrect).

As implied above, some investigators regard reading and spelling as two sides of one coin (Dumont, 1984, p. 74). Frith (1980) also seems to hold this opinion, because she denotes a child that is a proficient reader but at the same time a poor speller as exhibiting 'unexpected spelling problems' (see p. 496-497). A reciprocal relation of that type is also referred to as 'Principle of Associative Symmetry' (Baron, Treiman, Wilf, & Kellman, 1980, p. 165), which indicates that associations obtained in one direction can also be used in the opposite way. In terms of reading and spelling this entails that if one can read a word one should also able to spell it and the other way round. Earlier research, however, showed that the relation between reading and spelling does not satisfy the 'Principle of Associative Symmetry' unequivocally (Baron et al., 1980; Van Doorn-Van Eijsden, 1984; Chapters 2 and 3).

Van Doorn-Van Eijsden compared the spelling performance of two groups of Grade 5 students. One group was instructed to read a number of words, and a second group had to copy them on paper. The copy group performed significantly better than the reading group. In Chapter 3 we also compared the spelling performance of children who learned to spell words according to different instruction methods. Again, the three non-reading methods led to better spelling performance than the reading training. The reading training, however, was more effective than the absence of any training, but spelling performance was equally good after the children had read the words only twice as after having read them six times. This is in accordance with a result obtained earlier (Chapter 2). In that study, reading a word often (18 presentations) was no more effective than reading it only a few times (three presentations).

Of the three non-reading tasks used in Chapter 3 oral spelling proved to be superior. The instruction method 'problem naming', in which children were to name a part in a word that contained an ambiguous spelling (indicated through underlining; for example, in Dutch: the 'ei' in pal<u>ei</u>s'; a similar example in English would be the 'ea' in 'd<u>ea</u>r'), turned out to be equally effective, but more efficient than the instruction method 'copying': These two training procedures resulted in equals scores on the spelling test, but problem naming took less time than the training time for copying.

The study of Chapter 3, pertaining to differential effects of spelling-instruction methods, was conducted in the months May and June, about nine months after formal

reading and spelling instruction had started. An additional question in the experiments reported here was whether first graders at an earlier stage respond to the non-reading tasks the same way as the group investigated in the foregoing chapter. For that purpose, here, beginning readers and spellers will already be subjected to a spelling training in February.

Before turning to the experiments, a brief description of how spelling is instructed in Grade 1 will be given. In the reading curriculum 'Veilig leren lezen' (Caesar, 1979), spelling is secondary to reading. According to the curriculum, spelling should not be separated from reading, but reading and spelling need not keep pace, because recognising a written word is easier than reproducing it (Caesar, 1980). It is also stated that through reading an orthographic image may be formed, but that reading on its own not always suffices. The recommended procedure for learning the spelling of a word in Grade 1 is primarily to imprint its orthographic image showing the child a strip of paper on which the word is printed, whereupon the child must write the word 'in the air', reconstruct it using a 'letterbox' or, later in the year copy it on paper. The initial instruction words in the curriculum are regular words. Knowledge of the grapheme-phoneme correspondence rules is thus sufficient to read and spell these words correctly. Words that require rules in order to spell them correctly are not presented in Grade 1. An exception is the final 'd'.

Experiment 4.1

In Experiment 1 three questions were posed: 1) Are there indications that the spelling process of beginning spellers develops from a phonologic to a lexical strategy? 2) What is the effectiveness of a number of spelling-instruction methods that differ in depth of processing, and do these spelling-instruction methods lead to different spelling strategies? 3) Are these different instruction methods applicable at an early stage of formal reading and spelling instruction?

Method

Prior to the description of the experiment the general outline of the experimental design will be described. A child was asked to read out loud five words several times. Furthermore, the child had to learn the spelling of five different words. It was assumed that the child was unfamiliar with the orthography of the words. Spelling instruction occurred in four different ways (see 'Procedure' Section below). Each child learned the spelling of the words according to one of the four methods. The child's number of errors during the training and the time to complete the training were registered. After the reading and spelling training a dictation test of the training words was administered.

Materials. The word list of Kohnstamm et al.(1981) served as the source for the selection of ten semantically-familiar, but orthographically-unfamiliar words (mean length: 3.9 letters). Only stimuli that provided a spelling problem, that is, words that contained a phonologically ambiguous grapheme were selected: A Dutch example is the grapheme 'ij' in the word 'lijm', because the grapheme 'ei' is a phonologically acceptable alternative (an English example: the grapheme 'ee' in the word 'beef' has an acceptable alternative grapheme, namely, 'ea')[3]. The stimuli used in this experiment are presented in Appendix E, listed under 'easy words'.

The ten stimuli were divided in two groups of five words each. The words 1 to 5 constituted the first group and the words 6 to 10 the second group. In each group every word appeared three times, with the order of presentation such that no word appeared twice in a row. Thus in each condition 15 words were practised. To avoid that a group of stimuli was uniquely yoked to one condition, half of the subjects received one group of words in the reading condition and the other group of words in the non-reading condition, and with the other half of the subjects the order was reversed. This procedure also guaranteed that a stimulus was presented in only one of the two conditions and never in both. Each word was individually printed on a piece of paper of five by ten centimetres.

Procedure. An entire experimental session, in which each subject was tested individually, lasted approximately 15 minutes. The experiment was executed in February. Subjects were assigned to one of the following four conditions (non-reading task): 'copying', 'oral spelling', 'problem naming', or 'word composition'. Apart from the training of 15 words in a non-reading task the children also practised 15 (other) words in a reading task. The reading task involved reading out loud the words. If a word was read incorrectly the experimenter asked the child to read the word once more. Half of the children received the non-reading task in the 'stimulus-absent' condition and the other half in the 'stimulus-present' condition (see below).

In the oral-spelling condition the child was asked to read the word aloud once and then spell it out loud while the stimulus remained visible (present condition) or to spell it out loud with the word no longer being visible (absent condition). When analysing words children in Grade 1 tend to name the phonemes instead of the letters. Therefore, to make sure they correctly analysed the words they were asked to clarify each ambiguous phoneme that they named. In the copying condition the child was asked to read the word and then copy it onto a notebook. In the present condition the word remained visible while copying, whereas in the absent condition the word to be copied was covered

In the problem-naming condition the subject also had to read the word first. Subsequently, they were presented a piece of paper on which two graphemes were printed. One of these two grapheme occurred in the word. For example, with the word 'lijm' a piece of paper was shown with an 'ei' and an 'ij' (an English example: with the word 'beef' a piece of paper that contained 'ea' and 'ee' was shown). Subjects were instructed to put a circle around the correct grapheme. The location (on the left or right side) of the correct alternative on the piece of paper was pseudo randomised. Subjects in the absent condition saw the piece of paper with the word only briefly, whereas the subjects in the present condition were able to inspect the word throughout the trial.

Likewise, in the word-composition condition a word had to be read aloud. Subsequently the subject was asked to make up the word using letter tiles. Composing the words from the letter tiles was complicated by supplying not only all the correct graphemes, but also one incorrect, possible alternative grapheme for the target phoneme in the word. For instance, when the word 'zoen' was presented, not only the graphemes 'z', 'oe', and 'n' were provided, but also the grapheme 's' (an English example: when the word 'kiss' had to be composed of the letter tiles, not only the graphemes 'k', 'i' and 'ss' were presented, but also the grapheme 's'). The children assigned to the absent condition were presented the piece of paper with the word only briefly, whereas the children in the present condition were

able to inspect the word throughout the trial. In all training conditions feedback was provided.

Half of the subjects performed the reading task first, before one of the non-reading tasks; for the other half the order was reversed. The time subjects needed to complete the reading task and the non-reading task was measured with a stop-watch. Immediately after the training a dictation test was administered. For each of the 10 words a matching sentence was composed. The experimenter read out loud each sentence once and repeated the word the subject was required to write down.

Subjects. Two tests were administered to a group of 170 first grade students, a fortnight prior to the experiment: the reading decoding test (Caesar, 1975), and a spelling test developed by the first author[4]. The spelling test is presented in Appendix F.

A mean *Z*-score was computed based on the separate scores of the reading and spelling tests. From the sample of 170 subjects, 149 children participated in the experiment. The data of 136 were subjected to the analyses. The mean score of the subjects on the spelling test was 17.07 (SD = 2.7) and on the reading test it was 30.75 (SD = 11.4). The correlation between the scores on the reading and spelling test was .34 ($p < .01$)

Including a subject in the analysis required the fulfilment of two conditions, namely, an equal number of subjects in each condition (N = 17), and no significant differences between the experimental conditions on the mean *Z*-scores ($F(7, 128) = .006$, $p = 1.0$). Thus, differences between conditions could not be explained by differences in reading and spelling proficiency.

Results

The results of the training stage will be discussed before those of the test stage (the dictation test). For each word it was determined whether an error was phonologically correct or phonologically incorrect. An error was classified as phonologically correct if the pronunciation of the grapheme resulted in the intended sound (a Dutch example: in the incorrectly spelled words 'flag' - 'vlag' was intended - the 'f' is considered to be a phonologically correct error; an English example, in the incorrectly spelled words 'kis' ('kiss' was intended) the 's' is considered to be a phonologically correct error). Spelling errors, in which graphemes were deleted or the letters in a pair of graphemes were exchanged, were classified as phonologically incorrect[5]. The a priori chance of a phonologically correct error is lower than that of a phonologically incorrect error, because each word contains less ambiguous than unambiguous phonemes.

Training stage. A 4 (non-reading condition: copying vs. oral spelling vs. problem naming vs. word composition) by 2 (stimulus presence: absent vs. present) by 2 (task: reading vs. non-reading) analysis of variance (ANOVA) on training duration revealed significant main and interaction effects. The main effect of task was $F(1, 128) = 1193.64$, $p < .001$. Performing the reading task took significantly less time (mean 1.16 min) than performing the non-reading task (4.18 min). The main effect of non-reading condition was also significant, $F(3, 128) = 47.23$, $p < .001$. A Newman-Keuls post hoc analysis showed that each mean duration to complete a spelling training was significantly different from the duration in all remaining non-reading conditions ($p < .05$; oral spelling = 1.79 min, problem naming = 2.36 min, copying = 2.78 min, word composition = 3.76 min). The main effect of

stimulus presence was marginally significant, $F(1, 128) = 3.66$, $.05 < p < .10$. The spelling training in the absent condition took significantly more time than in the present condition (2.79 and 2.55 min respectively). The significant interaction effect between task and non-reading condition ($F(3, 128) = 82.04$, $p < .001$) revealed that the time to complete the reading task was not significantly different in the four non-reading conditions, but that the time to complete the four non-reading tasks did differ between conditions, $F(3, 128) = 70.59$, $p < .001$. An analysis of variance on the number of errors in the training indicated that no statistically significant differences emerged between the four non-reading conditions (the spelling-instruction methods). In the non-reading task, however, subjects made more errors (.81) than in the reading task, .54), ($F(1, 128) = 6.00$, $p < .05$.

Test stage. A 4 (non-reading condition: copying vs. oral spelling vs. problem naming vs. word composition) by 2 (stimulus presence: absent vs. present) by 2 (task: reading vs. non-reading) by 2 (error type: phonologically correct vs. phonologically incorrect) ANOVA on the mean number of errors revealed only one significant effect. The main effect of error type was $F(1, 128) = 78.83$, $p < .001$. The mean number of phonologically correct errors (.25) was significantly larger than the mean number of phonologically incorrect errors (.09). None of the main effects of non-reading condition, stimulus presence, and task reached significant levels (all F's < 1). Table 4.1 presents the mean numbers of phonologically correct and incorrect errors in all four reading instruction conditions.

Table 4.1. Mean numbers of phonologically correct and incorrect errors in Experiment 4.1 Standard deviations in parentheses.

Instruction method	Phonologically correct error	Phonologically incorrect error
Copying	.24 (.15)	.09 (.17)
Oral spelling	.27 (.16)	.09 (.12)
Word composition	.28 (.18)	.09 (.16)
Problem naming	.22 (.14)	.09 (.14)

Discussion

The results of Experiment 4.1 are <u>not</u> in accordance with those of Van Doorn-Van Eijsden (1984) and those of Chapter 3, in which it was shown that words studied in the non-reading tasks produced better results than words studied in the reading task. In Experiment 4.1 no difference in spelling performance between these two conditions was found. Furthermore, in the study of Chapter 3 differential effects were obtained for the three different spelling-instruction methods, whereas here all four methods (copying, oral spelling, problem naming, and word composition) resulted in equally good spelling performance. The spelling-instruction methods we used here differed in depth of processing. We assumed that oral spelling and word composition demanded a more thorough processing of the word than copying and problem naming. The fact that no differences occurred between conditions seems to indicate that depth of processing of a stimulus does not affect subsequent spelling performance in our present subjects. The answer to the question whether children in a very early stage of formal reading and spelling instruction benefit from the spelling-instruction

methods applied here, appears to be negative. However, a probably important difference between Experiment 4.1 and the study reported in Chapter 3 might explain why the effect failed to occur.

The stimuli in Experiment 4.1 were, as it happened, considerably less complex than those used in Chapter 3. The latter were comparable to the words used in Experiment 4.2 below. Thus, the absence of an effect could be due to near ceiling performance. This explanation is supported by two facts. First, the proportion errors in Experiment 4.1 (reading task: .17, non-reading task: .16) was relatively low as compared to that in Chapter 3 (reading task: .96, non-reading task: .57). Second, the variable stimulus presence during training did not affect performance on the spelling test. The children in the stimulus-absent condition did not make more errors than the children in the stimulus-present condition

The non-reading tasks turned out to be equally difficult (no significant differences in the number of mistakes during the spelling training), but the children did find the non-reading tasks more difficult than the reading tasks (more mistakes in the non-reading tasks than in the reading task during the spelling training).

The fact that the children made proportionally more phonologically correct than phonologically incorrect errors suggests that the children predominantly used the phonologic spelling route. The absence of interactions between the variables error type and non-reading condition, between error type and task (reading versus non-reading), and between the variables stimulus presence and error type suggests that the spelling-instruction methods investigated here do not bring about different spelling strategies. However, because it is not clear whether the stimuli we used were too simple, these conclusions are not justified yet.

The above provided reason to replicate Experiment 4.1. To check for a possible ceiling effect, words with a larger number of spelling problems were added to the original set of stimuli of Experiment 4.1, and the performance of a control group was taken into account.

The study in Chapter 3 clearly showed that oral spelling was the best method to learn the spelling of a word and that copying was the least effective one. It was, therefore, decided to exclude these methods from Experiment 4.2. The instruction method word composition was not investigated in the study of Chapter 3 and was therefore retained in Experiment 4.2. The issue whether more thorough processing of a word causes a qualitatively different spelling strategy was investigated once more. It was decided to employ the instruction tasks problem naming and word composition again. The major difference between problem naming and word composition concerns the completeness of processing. In problem naming the children only have to attend to the spelling problem, whereas in word composition the whole word has to be reconstructed. The factor stimulus presence was not manipulated in Experiment 4.2. All spelling instruction occurred in the stimulus-absent condition. Experiment 4.2 was executed in May, three months after the first experiment. The question concerning the effectiveness of the present spelling-instruction methods at an early stage can, therefore, no longer be addressed, but answers to the other two questions posed in this study (see 'Introduction' Section Experiment 4.1) can still be provided by Experiment 4.2.

Experiment 4.2

Method

Materials. The stimulus materials consisted of the ten words used in Experiment 4.1 (easy words) and ten new words (difficult words; see Appendix E). The difficult words (mean length 6.7 letters) were selected in the same way as the easy words (word list of Kohnstamm et al., 1981). The difficult words differ from the easy words in word length (they are longer) and number of potential spelling problems (they contain more). In the same manner as in Experiment 4.1 the ten new words were divided in two groups of each five words, and all words were repeated three times. This time, however, instead of printing each word on a separate piece of paper, the words were printed on a sheet of paper (A-4 format) with one word on every line. The purpose of this change was solely practical, that is, running the training sessions became easier.

Procedure. As in Experiment 1 the children were tested individually. They also needed about 15 minutes to complete the experimental session. The design and procedure were approximately the same as in Experiment 4.1. Again all children performed a reading and a non-reading task. In this second experiment only two of the four non-reading tasks were employed, namely, problem naming and word composition. Half of the children were trained with the easy words and the other half practised the difficult words. In the non-reading task the same procedure was used as in the stimulus-absent condition of Experiment 4.1. Thus, in this new design all children were to read the word aloud once, and while the spelling was practised a cover was placed on the word list.

The problem-naming and word-composition tasks were administered in the same way as in Experiment 1, except that for the children in the problem-naming condition the words and the target graphemes were now all printed on the same sheet (in Experiment 4.1 each word was printed on a separate sheet). Subjects in the control group did not participate in any of the training tasks. They only performed the spelling test.

Subjects. From a new sample of 109 Grade 1 students a group of 72 children was admitted for analyses on the basis of the same criteria as in Experiment 4.1. Each condition comprised the same number of subjects ($N = 12$). As in Experiment 4.1 the mean Z-scores of the reading and spelling-test scores were determined and care was taken that mean Z-scores did not differ significantly between conditions ($F(5, 66) = .011$, $p = 1.0$). The mean score of the subjects on the spelling test was 18.21 ($SD = 1.7$), and on the reading test it was 44.64 ($SD = 14.4$). The correlation between the scores on the reading and spelling tests was .41 ($p < .01$). Again, the reading and spelling tests were administered one week before the experiment.

Results

The results of the training stage will be discussed first, followed by those of the test stage (the dictation test). Error judgement on the dictation test was identical to that of Experiment 4.1. After discussing the results of the training and test stage of Experiment 4.2, a second

analysis will be reported, in which the dictation data of Experiment 4.1 and those of the easy word condition of Experiment 2 are combined, to test for the effect of time of testing.

Training stage. A 2 (non-reading condition: word composition vs. problem naming) by 2 (complexity level: easy vs. difficult) by 2 (task: reading vs. non-reading) ANOVA on training duration showed again a main effect of task, $F(1, 44) = 1776.48$, $p < .001$. Performing the reading task took significantly less time (.54 min) than performing the non-reading task (5.53 min). As in Experiment 1 the main effect of the non-reading condition reached significance, $F(1, 44) = 391.06$, $p < .001$. Problem naming (1.52 min) took less time than word composition (4.55 min). The significant interaction effect between non-reading condition and complexity level ($F(1, 44) = 12.93$, $p < .001$) revealed that significantly more time was required to finish the word-composition task on difficult words than on easy words, $F(1, 44) = 18.86$, $p < .001$. In the problem-naming condition no difference was found on task duration for easy and the difficult words. A significant interaction effect also emerged between task and complexity level, $F(1, 44) = 49.24$, $p < .001$. The reading task took on average .67 min in the difficult condition and .42 min in the easy condition. Completing the non-reading task in the difficult condition took on average 6.49 minutes, whereas in the easy condition it required 4.57 minutes.

An analysis of variance on the number of mistakes made during training showed no significant difference between the word-composition and the problem-naming conditions, but subjects made more mistakes during the non-reading task (1.69) than during the reading task (.56), $F(1, 44) = 29.8$, $p < .001$.

Test stage. A 3 (non-reading condition: word composition vs. problem naming vs. control group) by 2 (complexity level: easy vs. difficult) by 2 (error type: phonologically correct vs. phonologically incorrect) by 2 (task: reading vs. non-reading) ANOVA was executed on the mean number of errors. The main effect of non-reading condition was marginally significant, $F(2, 66) = 2.96$, $.05 < p < .10$. A planned contrast showed that the mean number of errors in the word-composition (.26) and problem-naming (.29) conditions did not differ significantly from each other, but subjects in both these conditions made significantly less errors than those in the control group (.39), $F(1, 66) = 5.64$, $p < .05$. The main effect of complexity showed that subjects who were trained on the easy words (.14) made significantly less errors than those trained on the difficult words (.50), $F(1, 66) = 68.39$, $p < .001$. The main effect of task revealed that more errors were made on words practised in the reading task (.36) than in the non-reading task (.28), $F(1, 66) = 19.31$, $p < .001$. As in Experiment 4.1, the proportion of phonologically correct errors was higher than that of phonologically incorrect errors (.44 and .19 respectively), $F(1, 66) = 61.19$, $p < .001$. The results are presented in Table 4.2.

Table 4.2. Mean numbers of phonologically correct and incorrect errors in Experiment 4.2 Standard deviations in parentheses.

Instruction method	Phonologically correct error	Phonologically incorrect error
Word composition	.39 (.29)	.14 (.16)
Problem naming	.38 (.33)	.21 (.22)
Control group	.55 (.39)	.23 (.23)

The significant interaction ($F(1, 66) = 11.15$, $p < .01$) between error type and complexity level qualifies the results obtained above, because the difference between the proportion of phonologically correct and incorrect errors was larger in the easy (.20 and .06 respectively; 77%) than in the difficult condition (.68 and .33 respectively; 67%). This interaction is presented in Figure 4.1. No significant interactions between error type and any of the other variables emerged. The second-order interaction between complexity level, non-reading condition and task ($F(2, 66) = 3.15$, $p < .001$), and the first-order interaction between complexity level and task ($F(2, 66) = 13.23$, $p < .01$) were also significant. The one between complexity level and non-reading condition was marginally significant, $F(1, 66) = 2.71$, $.05 < p < .10$. The interaction between complexity level and task indicated that a significant difference emerged between the reading and non-reading tasks in the difficult condition ($F(1, 66) = 32.25$, $p < .01$), but not in the easy condition ($F < 1$). The interaction between complexity level and the non-reading condition showed that control subjects made more errors on the difficult words (.68) than the experimental subjects (word composition, .38 and problem naming, .48) in the difficult condition. In the easy condition, no such differences in the easy condition between subjects of the word composition, problem naming and control group occurred (.15, .10, and .15 respectively).

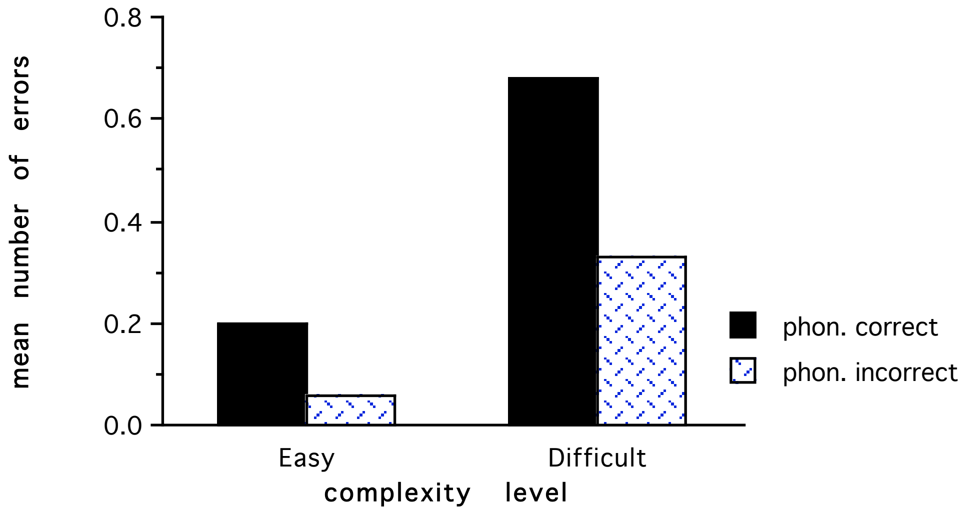

Figure 4.1. Mean numbers of phonologically correct and incorrect errors in the easy and difficult conditions of Experiment 4.2.

Results of the easy condition of Experiments 4.1 and 4.2 combined.

Training stage. A 2 (time of testing: February vs. May) by 2 (non-reading condition: word composition vs. problem naming) by 2 (task: reading vs. non-reading) ANOVA on training duration showed a significant main effect of time of testing, $F(1, 44) = 15.54$, $p < .001$.

Subjects who participated in May needed less time to complete the training (2.49 min) than those who took part in February (3.32 min). The main effect of non-reading condition was, as in the previous analyses, significant, $F(1, 44)= 93.74$, $p < .001$. The mean training duration in the problem-naming condition was significantly shorter (1.89 min) than in the word-composition condition (3.93 min). The difference in training duration between the reading (.90 min) and non-reading tasks (4.91 min) was also significant, $F(1, 44) = 1019.05$, $p < .001$. The significant interaction between time of testing and non-reading condition ($F(1, 44) = 4.52$, $p < .05$) showed that all four means differed significantly from each other (Newman-Keuls, $p < .01$), except in the word-composition condition. Subjects participating in May needed as much time to complete the word-composition task as those who took part in February.

Test stage. A 2 (time of testing: February vs. May) by 2 (non-reading condition: word composition vs. problem naming) by 2 (error type: phonologically correct vs. phonologically incorrect) by 2 (task: reading vs. non-reading) ANOVA on the mean number of errors only revealed a significant main effect of error type, $F(1, 44) = 17.79$, $p < .001$. It indicated that the mean number of phonologically correct errors (.19) exceeded the number of phonologically incorrect errors (.08). None of the main effects of time of testing, non-reading condition and task reached a significant level (all F's < 1). Mean numbers of phonologically correct and incorrect errors are presented in Table 4.3.

Table 4.3. Mean numbers of phonologically correct and incorrect errors in Experiment 4.1 (February) and the easy condition of Experiment 4.2 (May). Standard deviations in parentheses.

Time of testing	Phonologically correct	Phonologically incorrect
February	.21 (.15)	.10 (.13)
May	.18 (.17)	.07 (.13)

Discussion

The results of Experiment 4.2 are in accordance with those of Van Doorn-Van Eijsden (1984) and Chapter 3. These studies also showed that learning a word's spelling using a reading task is less effective than with the help of a non-reading task. No difference between the problem-naming and word-composition tasks was found on the spelling test. The performance of the control group was substantially worse than those of the experimental groups and, furthermore, easy words were spelled better than difficult words. It can be concluded that the absence of a differential effect of spelling-instruction methods in Experiment 4.1 was probably due to the fact that the stimuli were too easy. Indeed, the results of Experiment 4.2 revealed an interaction between the variables complexity level and non-reading condition. The children in the easy condition (analogous to the situation of Experiment 4.1) who learned to spell words using the problem-naming or word-composition method did not perform any better than the children in the control group, whereas in the difficult condition substantial differences emerged between groups.

Consistent with Experiment 4.1, proportionally more phonologically correct than phonologically incorrect errors were made. This finding supports again the assumption that,

after a spelling-instruction training, beginning spellers predominantly use the phonologic route to spelling. The absence of an interaction between the variable error type (phonologically correct vs. phonologically incorrect) and the factor non-reading condition and between type of error and task justifies the conclusion that the spelling-instruction methods investigated here did not induce qualitatively different spelling strategies. Neither the non-reading tasks word composition and problem naming nor the reading task seemed to have brought about a lexical way of spelling. The alternative explanation that the children were not yet ready for a lexical spelling strategy, because of still too low a spelling level, seems unlikely. The mean score on the spelling test was 18.21 whereas the maximum score is 20.

It was assumed earlier that word composition requires a more thorough processing of the stimulus than problem naming, and that a more thorough processing of the stimulus leads to better spelling performance. The results do not confirm this assumption because no difference occurred in mean number of spelling errors between the instruction methods word composition and problem naming. However, we would not want to draw the conclusion that depth of processing has no effect whatsoever on the completeness of the representation. Another explanation for this null-effect is that the assumption that word composition requires a more thorough processing of the stimulus than problem naming does, was wrong after all. In the word-composition task, contrary to the problem-naming task, the entire word needed to be reconstructed, but it did not require a particularly active attitude on the part of the subject, because the letters of the word were present all along. The subject had to determine the correct order of the letters composing the word, using the letter tiles. The only choice to be made concerned the target, because for this a plausible alternative was present. Therefore our operationalisation of the variable depth of processing may have been unfortunate.

The problem-naming training requires less time than the word-composition training, but they have turned out to be equally effective as spelling-instruction methods (the number of errors in the two conditions were statistically the same). The conclusion that problem naming is a more efficient instruction method than is word composition seems, therefore, warranted. Another conclusion is that not the time spent on training, but the quality of what happens in the invested time determines spelling proficiency. This conclusion is supported by the combined results of Experiments 4.1 and 4.2. The children in the problem-naming condition needed more time to accomplish both the reading and non-reading tasks in February than in May. The children in the word-composition condition also needed more time to complete the reading task in February than in May. But, the subjects who performed the word-composition task required equal amounts of time in both months on the non-reading tasks. The reason that the children who performed the word-composition task in May needed as much time as those performing this task in February probably is that the largest part of the task consists of seeking, placing, and removing the letter tiles. In other words, the attention in the word-composition training is largely non-spelling oriented.

General Discussion

The experiments reported above show that inexperienced spellers predominantly use a phonologic way to spell both easy and difficult words. A point of criticism could be that we simply did not present our subjects with the difficult words often enough to give rise to a lexical strategy. The fact, however, that also the easy words are spelled phonologically counters this argument.

Learning to spell does not seem a matter of reading a word over and over again. Like in the study of Van Doorn-Van Eijsden (1984) and the one reported in Chapter 3, it turned out that reading is the least appropriate way to learn to spell, and in fact not sufficient to learn the spelling of a word. The spelling-instruction methods copying, problem naming, and oral spelling (in Chapter 3), and word composition and problem naming (in Experiment 4.2) prove to yield better spelling results than does reading. The results of our study thus point out that Reitsma's comment (1985, p. 82) that one especially learns to spell by spelling is justified.

Reading and copying are the most frequently used didactic methods for learning to spell. Even though copying provides better results than reading this method is not particularly efficient, because it requires more time than problem naming, and yet the learning effect is the same for both. Thus, it is the quality of the learning activity that is critical and not its sheer duration. From the above it can be concluded that the relation between reading and spelling is not as tight as was assumed initially (see also the 'Introduction' Section of this chapter).

To summarise, in all situations investigated here beginning readers/spellers show a clear preference for a phonologic spelling strategy, and reading practice does not turn them into skilled spellers.

5

EVIDENCE FOR ASSEMBLED PHONOLOGY IN BEGINNING AND FLUENT READERS AS ASSESSED WITH THE FIRST-LETTER-NAMING TASK*

Summary

This study was aimed at clarifying word recognition in beginning (first graders) and fluent readers (university students). The major goal was to investigate the differences and similarities in word recognition of the two groups. The theoretical starting point was the 'dual-route' model. The research paradigm that was used was the first-letter-naming task as employed by Rossmeissl and Theios (1982). In three experiments, the factors affecting performance in this task were determined. The data suggest that the effects obtained are due to response competition. Furthermore, the data indicate that the similarities in word recognition by fluent and beginning readers far surpass the differences and, more importantly, that word recognition of both fluent and beginning readers is mediated by phonology.

This study aims at clarifying the differences and similarities of word recognition in beginning and fluent readers. More specifically, the issue under investigation is whether or not there is a developmental shift to an expert way of reading that is different from the reading by novices.

The theoretical starting point is the 'dual-route' model (Coltheart, 1978). The core of this model is the assumption that fluent readers have two independent routes at their disposal to read words. Via the 'nonlexical' route, a word is read by means of discrete grapheme-phoneme correspondence rules, and only by activating the phonology of the stimulus can word recognition come about. This process is also called 'phonologic mediation' or 'assembled phonology'. It is assumed that this nonlexical way of reading is used by readers of an alphabetic writing system only when they are learning to read (beginning readers) or when they are fluent readers, but come across infrequent words (Seidenberg, Waters, Barnes & Tanenhaus, 1984). The primary process in fluent readers is the lexical route, in which a direct match is established between the written word and its mental representation, caused by frequent presentation with a word.

Adherents of the dual-route model state that the lexical route is a prerequisite if one is to read irregular (exception) words. The dual-route model dominated the eighties and survived a proposal for a purely lexical alternative suggested by Humphreys and Evett (1985). Recent publications (e.g., Lukatela, Carello & Turvey, 1990; Perfetti & Bell, 1991; Van Orden, 1987, 1991; Van Orden, Pennington & Stone, 1990), however have made the pendulum swing back (see Morton, 1985, p. 718). These studies appear to prove that even the most fluent readers read exclusively by phonologic mediation. A critical examination of

---

* This chapter is in press (*Journal of Experimental Child Psychology*), and is co-authored by A. M. B. De Groot.

the major assumptions of the dual-route model can be found in Van Orden, Pennington and Stone (1990).

In this study we made use of the first-letter task to gain further insight into the word-recognition processes of beginning and fluent readers. The first-letter task was developed by Rossmeissl and Theios (1982), and is a variant of the Reicher-Wheeler paradigm.

Reicher (1969; see also Wheeler, 1970) presented subjects letters and letter strings tachistoscopically (35-85 ms), which were followed by a patterned mask. Subjects then had to identify a letter that had been presented in isolation or one that had appeared in a preceding letter string on a prespecified position. Identification of a letter was more accurate when it had been part of a word than when it had appeared in isolation or in an orthographically illegal letter string. This result, which is called the 'word-superiority effect' was interpreted as evidence for parallel processing of letters in words. The word-superiority effect has not only been established in adult fluent readers, but is also apparent in beginning readers as young as 98 months (Chase & Tallal, 1990).

In the first-letter task, Rossmeissl and Theios (1982) presented adult subjects with three types of stimuli: words, orthographically legal nonwords and orthographically illegal nonwords (unpronounceable letter strings, which they called anagrams). The task for the subjects was to name the first letter of the stimuli. They reasoned that if letters in words are processed in parallel and if the orthographic context provides an independent source of information, then the identification of the first letter should be facilitated by a legal letter string (words and legal nonwords), but not by an illegal one. This hypothesis was confirmed by their data. First-letter naming of words (550 ms) and legal nonwords (553 ms) was significantly faster than of illegal nonwords (562 ms). They regarded this as further evidence for a direct access model of word recognition. A model in which letters are processed serially from left to right (the assumption of the phonologic mediation route in the dual-route model) would not have predicted an effect of orthographic legality. Thus, the effects found in the first-letter paradigm seem to suggest that words can also be identified without the use of grapheme-phoneme correspondence rules or, stated more explicitly, that fluent readers recognise words lexically.

The first-letter paradigm is very suitable for experiments with beginning readers, because it requires a response that young children find easy. If beginning reading is characterised by a serial process of grapheme-phoneme conversion (see above) then young children should not show the effect established by Rossmeissl and Theios (1982), because the orthographic context can only exert an influence when letters are processed in parallel. Experiment 5.1 is a replication and extension of the Rossmeissl and Theios (1982) first-letter experiment. Not only beginners but also fluent readers served as subjects. (The rationale for Experiments 5.2 and 5.3 will be clarified following the presentation of Experiment 5.1.)

Experiment 5.1

The main question of Experiment 5.1 was whether first-letter naming in beginning readers is also, like in fluent readers, facilitated by an orthographically legal context. Therefore both reader groups participated. The experiment aims at replicating and extending the results of

Rossmeissl and Theios (1982). In the study of Rossmeissl and Theios, subjects were asked to specify the first letter of a string with its letter name. In our Experiment 5.1 (and 5.3) half of the adult subjects were asked to use the letter name to specify the first letter and the other half were asked to use the phoneme. The group of beginning readers all used phonemes to respond. This procedure was followed to justify both the comparison between the adult subject group of Rossmeissl and Theios and our adult group (letter naming), as well as the comparison between our beginning and fluent readers (phoneme naming). The procedure to have the beginning readers respond with letter names as well was considered undesirable, because, totally unfamiliar with this task, they could be expected to experience great difficulty with it. In contrast, phoneme naming, a very familiar task for Dutch first graders, is a rather unfamiliar task for fluent readers. Yet, they may be expected to manage this task.

Method

Materials. The total set of 72 stimuli consisted of 24 words, 24 legal nonwords and 24 illegal nonwords. The word stimuli were taken from the first three instruction books of the children's reading course (Caesar, 1979). All word stimuli contained four letters and had a CV_1V_2C-structure. The word stimuli served as the starting point for the construction of the nonword stimuli. By reordering the letters or introducing new letters and deleting old, but maintaining the initial one, legal nonwords (pronounceable and orthographically legal letter strings) and illegal nonwords (unpronounceable and orthographically illegal letter strings) were constructed. Appendix G shows the complete set of stimuli.

Procedure. The experiment was run on a Macintosh Classic computer. Subjects were told that stimuli would appear on the screen and that they had to name the first letter of each word and ignore the word they were part of. Beginning readers used the phoneme to specify the initial grapheme. Half of the group of adult subjects was instructed to use the letter name to specify their response (as in Rossmeissl & Theios, 1982), and the other half used phonemes to indicate their response, as the children did. Naming times were registered with a voice-key and a millisecond timer. Each response was evaluated on correctness by the experimenter by pressing a key on the computer keyboard. Before the experimental session started every subject received 10 practice trials.

Subjects. Both beginning and fluent readers served as subjects. From a population of 246 children of Grade 1 a group of 28 children was selected, which had a slightly above average score on the reading test (Caesar, 1975 mean: 35, SD = 9.2; N = 13).The group of fluent readers (N = 40) was recruited from the undergraduate population of the Department of Psychology of the University of Amsterdam.

Results

The results of the beginning and fluent readers were analysed separately but are discussed simultaneously. We are aware that normally data from both reading groups should be analysed together, but the variance in the data of the beginning readers is relatively large (a well-known problem) compared to that of the fluent ones. Consequently, the manipulations resulted in significant effects for the fluent readers only when their data were analysed separately

Before the data were analysed, responses were removed from the data set for the following reasons: naming errors (children: .9%; adults: .7%); errors due to voice-key failure (children: 3.9%; adults: 1.7%); extremely long responses (more than 3 SD above the mean; children: 1.5%; adults: .9%); extremely short responses (less than 100 ms; children: 0%; adults: .2%).
Beginning readers. Subject means were calculated for the three types of stimuli, and these means were the input for the statistical analyses. A one-way ANOVA (stimulus type: words vs. legal nonwords vs. illegal nonwords) on the means of the beginning readers revealed a significant effect, $F(2, 54) = 15.70$, $p < .001$. A post-hoc analysis (Newman-Keuls) on the subject data showed a significant difference between words and legal nonwords, and between words and illegal nonwords ($p < .01$ in both cases), but not between legal and illegal nonwords[6]. The means are presented in Table 5.1.

Table 5.1. Mean first-letter-naming latencies in ms for beginning and fluent readers in Experiment 1 (SD in brackets).

		Response type	
Reading level	Beginning readers	Fluent readers	
Stimulus type	Phoneme	Letter	Phoneme
Words	756 (134)	465 (31)	508 (64)
Legal Nonwords	790 (109)	469 (30)	518 (76)
Illegal Nonwords	808 (123)	475 (29)	527 (75)

Fluent readers. A 2 (response type: letter vs. phoneme) by 3 (stimulus type: words vs. legal nonwords vs. illegal nonwords) ANOVA on the data of the fluent readers (see Table 5.1) revealed significant main effects of both response type and stimulus type. The main effect of response type indicated that responding with the letter was faster (469 ms, $SD = 29$) than with the phoneme (518 ms, $SD = 71$), $F(1, 38)= 7.91$, p < .01. The main effect of stimulus type was: $F(2, 76)= 12.18$, $p < .001$. The interaction between stimulus type and response type was not significant ($F < 1$). The means are presented in Table 5.1.

Further investigation of the effect of stimulus type showed that all conditions were significantly different from each other (a post-hoc Newman-Keuls analysis). The difference between words (487 ms, $SD = 55$) and legal nonwords (493 ms, $SD = 62$) was significant at the 5% level; the differences between words and illegal nonwords (501 ms, $SD = 62$), and between legal and illegal nonwords at the 1% level.

A separate analysis (stimulus type: words vs. legal nonwords vs. illegal nonwords) on the data of the subjects in the letter-response condition (to compare the results with those of Rossmeissl & Theios, 1982) showed that in this group the difference between words and legal nonwords did not reach significance. Adding the data from the subjects in the phoneme-response condition increased the power of the test thus leading to a significant difference on this comparison.

Discussion

The results of the Dutch fluent readers in the letter response condition of Experiment 5.1 (see Table 5.1) replicate those of the English subjects of Rossmeissl and Theios (1982). Naming the first letter was faster when it was part of a legal letter string (words or nonwords) than when part of an illegal nonword. The differences between conditions were almost exactly the same as the corresponding differences in Rossmeissl and Theios' study. The latency difference between illegal and legal nonwords was 6 ms, between illegal nonwords and words it was 10 ms and between legal nonwords and words it was a non-significant 4 ms. The comparable data from Rossmeissl and Theios were, respectively, 9 ms, 12, and 3 ms. What is peculiar though is the fact that the Dutch subjects were on average 85 ms faster than the English subjects. The reason for this cannot have been the amount of practice, because that should have favoured the English subjects. They received 1008 trials, whereas the Dutch subjects only got 72.

The results of the total group of fluent readers, however, presented a slightly different picture. It is clear that the group who had to respond with the letter name was faster than the group that had to specify the phoneme, presumably because the latter group had to give a rather unfamiliar response. The overall data pattern shows that naming the first letter or phoneme of words is faster than of both legal and illegal nonwords, and that legal nonwords are processed faster than illegal nonwords.

That performance on legal nonwords is statistically equal to that on words in the letter condition (replicating the results of Rossmeissl & Theios, 1982), but worse in the phoneme condition might be due to a ceiling effect in the former situation. We believe that the difference between processing words and legal nonwords is real, but hard to detect in the first-letter-naming task. This assumption is not only corroborated empirically in a naming task by McCann and Besner (1987), but is also explicitly stated in most word-recognition models, for instance in Paap, Newsome, McDonald, & Schvaneveldt's activation-verification model (1982), Seidenberg and McClelland's PDP model (1989), and Van Orden, Pennington, and Stone's subsymbolic approach to word recognition (1990). Our hypothesis is that the latent, but statistically unreliable difference between legal nonwords and words became manifest in the more difficult phoneme condition.

The results of the beginning readers mimicked those of the total group of fluent readers. Naming the first letter of a word was clearly easier than of legal and illegal nonwords. Furthermore, on the additional analysis (see Note 6) it was shown that naming the first letter of legal nonwords was faster than of illegal nonwords. It thus seems that the similarities between these groups of beginning and fluent readers surpass the differences. The most important difference between the data patterns of both groups concerns their overall latencies: Beginning readers take almost 60% more time to come up with the response than fluent readers (785 ms and 494 ms, respectively).

From the above it seems safe to conclude that after eight months of formal reading instruction in the Netherlands, beginning readers have developed a process of word recognition that is rather similar to that of experienced readers. In both reader groups naming the first letter was facilitated when it was presented in an orthographically legal letter string. Furthermore, for both groups facilitation was stronger in a word context than in a legal nonword context. It is important to note that the similarity of the word-recognition processes of fluent and beginning readers is limited to the set of words the beginners have

been exposed to. It would be unjustified to generalise the results to word recognition in general. If the orthographic-context interpretation of Rossmeissl and Theios (1982) holds, then these results suggest that Dutch beginning readers fairly quickly develop the skill to recognise words via the lexical route.

However, although the interpretation of Rossmeissl and Theios (1982) sounds plausible, there is an alternative explanation for their results (and, for that matter, for our results). Their interpretation points out identification as the locus of the effect, but it is also possible that the results reflect an effect of response competition.

When a literate subject is presented with a rather familiar stimulus like a word, and, as we will assume, also with an unfamiliar letter string, he or she cannot avoid reading it. In the first-letter paradigm, however, the subject is not asked to read the stimulus, but to ignore it and to name its first letter. If it is indeed impossible to avoid processing the entire stimulus (in other words, this process is enacted automatically), it is plausible that naming the first letter is in fact hindered when it is part of a letter string. One additional assumption is needed for this explanation to apply to the data namely, that this processing has to come to an end before a subject can execute the required response. If indeed this process occurs automatically, this assumption is plausible .

This conception of mental processing in the first-letter task provides an alternative explanation of why subjects are faster in naming the first letter of a word than of an illegal nonword: automatic processing of a word terminates faster than the processing of an illegal nonword. Consequently, naming the first letter of a word can be accomplished faster than of an illegal nonword. Note that a 'horse-race' model (Coltheart, 1978; but see Paap & Noel, 1991) would not necessarily predict more interference from an illegal letter string. In terms of that model, both processes, the one involving the entire stimulus and the one forced on by the task, are started simultaneously, but the letter-naming process does not have to await the outcome of processing the whole stimulus to produce a response. In case the processes proceed independently and the first-letter naming process terminates first, no interference from the 'secondary' process has to be expected. In contrast to this view, we suggest that the subject always <u>does</u> finish processing the whole stimulus prior to attending to the required response.

Two arguments support this response-competition interpretation, one anecdotal and one empirical. Both adults and children responded occasionally with the complete stimulus instead of the letter. After completion of Experiment 5.1 all adult subjects and some beginning readers reported that they could read the stimuli. The children seemed reluctant to admit they read the stimuli as well, because the experimenter had clearly indicated prior to the experiment that they were not supposed to do so. Reicher's subjects also reported that they sometimes were able to pronounce a nonword or turn a nonword into a word (1969, p. 279).

The empirical argument is based on experiments executed by us (see Note 7) with fluent and beginning readers. In these experiments, beginning and fluent readers had to perform the same first-letter task as used here. However, in this case subjects were presented not only with words and nonwords but also with single letters. Both the beginning and fluent readers were significantly and substantially faster naming a single letter than naming any other kind of stimulus, including words.[7]

According to the response-competition interpretation, naming a single letter should be faster than naming the first letter of any kind of letter string, because in the former case there is no response competition. The orthographic-context explanation of the first-letter effect of Rossmeissl and Theios (1982), however, predicts longer naming times for single letters than for letters in words, because single letters lack a facilitating context. In other words, the single-letter effect does not support the identification interpretation of Rossmeissl and Theios (1982).

Rossmeissl and Theios (1982) explain the first-letter effect in terms of a facilitation mechanism (as we did before introducing the response-competition account). Although there is no explicit mentioning of the word 'facilitation' it is clearly implied: ".... it is assumed that orthographic knowledge provides an independent source of information helping to identify each of the letters.." (p. 448). Their experiment, however, does not really permit an interpretation in terms of facilitation because there was no neutral condition from which the direction of the effects could be assessed. The first-letter effect could thus just as well be explained in terms of less inhibition in the case of words and legal nonwords as compared to illegal nonwords. It is plausible that the single-letter condition (see note 7) is the neutral condition from which the effects should be assessed. Because it was the fastest condition, the conclusion seems warranted that in all other conditions responding is inhibited.

If the first-letter effect is indeed caused by a process of response competition, the interpretation of the effect in terms of lexical word recognition should be reevaluated. The suggestion of Rossmeissl and Theios that the first-letter effect is evidence for lexical reading can only be valid if the locus of the effect is in the identification process. If, on the other hand, the response-competition hypothesis holds, then Rossmeissl and Theios' data are not conclusive with respect to the issue of how word reading comes about.

To summarise so far, we began this investigation from the assumption that the first-letter effect might evidence lexical reading. We then argued that the common view that beginning readers typically read via a serial process of phonologic mediation would thus predict the absence of a first-letter effect for this reader group. In fact, it turned out that beginning readers behaved similarly to fluent readers in this task. But instead of drawing the conclusion these data show that both beginning and fluent readers read via the direct lexical route, we subsequently ventured the hypothesis that a process of response competition may be responsible for the effects in both reader groups. Ultimately then, the data may be quite compatible with the view that reading in both beginning and fluent readers is phonologically mediated.

In Experiments 5.2 and 5.3, the response-competition hypothesis were investigated further. In Experiment 5.2 we depart from the assumption that indeed response competition underlies the first-letter effect and focus on the nature of the competition process. More specifically, this experiment was aimed at revealing the variable or variables that cause the first-letter effect. Experiment 5.3 was designed to provide more direct evidence for the response-competition hypothesis. Along the way, support will be gathered for the view that assembling the phonology of letter strings is an imperative process whenever beginning and fluent readers encounter letter strings.

A major problem when trying to interpret the results of Rossmeissl and Theios (1982) and of Experiment 5.1 is that orthographic legality and pronounceability are generally

confounded.[8] Both our and Rossmeissl and Theios' orthographically legal letter strings were pronounceable, while the illegal ones were always unpronounceable. Later, it will be shown that orthographic legality and pronounceability can be disentangled.

Careful inspection of the experimental materials suggests that the following four stimulus characteristics may be relevant for performance in the first-letter task: 1) lexicality: is the letter string a word or not? 2) orthographic legality: does the letter string obey orthographic rules? 3) phonologic lexicality: does sounding out the letter string lead to a word? and 4) pronounceability: is the letter string pronounceable? Orthographic legality correlates almost perfectly with pronounceability, but in the Dutch language it is possible to create orthographically illegal pseudohomophones that are yet pronounceable. Table 5.2 serves as the reference point for the description of the tests that can be executed to find out which of the above stimulus characteristics is (are) critical for performance in the first-letter task.

Table 5.2. Relevant factors determining stimuli used in Experiment 5.2

	Stimulus type				
	Word	Legal PSH	Illegal PSH	Legal NW	Illegal NW
Example	zalf	zant	zalv	zulf	zfli
	(salve)	(sand)	(salve)		
Factor					
Lexicality	yes	no	no	no	no
Orthographic Legality	yes	yes	no	yes	no
Phonologic Lexicality	yes	yes	yes	no	no
Pronounceability	yes	yes	yes	yes	no

PSH = pseudohomophone; NW = nonword; legal/illegal refers to orthographic legality

If the first-letter effect were due to lexicality, a significant difference should arise between naming the first letter of words and legal pseudohomophones. A test for orthographic legality involves the comparison of legal and illegal pseudohomophones. A test for phonologic lexicality compares legal pseudohomophones and legal nonwords. In the present design it is possible to test for pronounceability only when the tests for phonologic lexicality and/or for orthographic legality fail. If phonologic lexicality turns out to be an irrelevant factor, the difference between illegal pseudohomophones and illegal nonwords is the proper test for pronounceability. If, on the other hand, the test for orthographic legality fails, the contrast of legal versus illegal nonwords is the appropriate one. To be able to test on pronounceability without having to rely on the negative outcome of any of the other factors would only have been possible if illegal pronounceable nonwords had been included in the materials.

Experiment 5.2

The main purpose of Experiment 5.2 was to establish whether the effects found by Rossmeissl and Theios (1982) and in our Experiment 5.1 are caused by the difference in orthographic legality or in pronounceability between the stimulus types used in those experiments, or perhaps by both. Experiment 5.2 was also aimed at testing the relevance of the factors lexicality and phonologic lexicality.

One remark on the procedure is in order. Prior to the experiment proper the subjects had to read the word stimuli from which the legal and illegal pseudohomophones were derived three times to make sure that they were familiar with them at the onset of the experiment. Most of these words were relatively simple for this group of beginning readers and the children may have seen them before, but they have not yet occurred in the reading curriculum so far. Earlier research (Chapter 2; Reitsma, 1983) indicates that only a few presentations with a word are necessary to familiarise a beginning reader with this word. To check whether these 'training' words had the same familiarity as well-known words (that are part of the reading curriculum), latencies on training words and well-known words were compared. In case of equal familiarity, these two stimulus types should show no significant difference on the first-letter task. The reason training words rather than well-known words were used to derive legal and illegal pseudohomophones was that only the training words lent themselves for the construction of pseudohomophones of both types.

Method

Materials. A selection of 28 one-syllable words was made from the word list of Kohnstamm et al. (1981). All stimuli could be changed into pseudohomophones. In Dutch it is possible to construct pseudohomophones that are either orthographically legal (obeying Dutch orthographical rules) or illegal (violating Dutch orthographical rules). From these 28 selected training words half were transformed into legal and the other half into illegal pseudohomophones (all pronounceable). The words also served as the starting point for the construction of nonwords. Half of the nonwords were orthographically legal and the other half were illegal (unpronounceable, also called anagrams). Finally, 14 well-known words were selected from the first three reading books of the regular curriculum of the children. This led to six types of stimuli: 28 training words, 14 legal pseudohomophones, 14 illegal pseudohomophones, 14 legal nonwords, 14 illegal nonwords and 14 well-known words (a total of 98). In every condition the letters 'b', 'd', 'f', 'g', 'h', 'k', 'l', 'm', 'n', 'p', 'r', 's', 'v', and 'z' served as the initial letter of one of the stimuli. The mean length of the stimuli in each condition was the same (4.4, *SD* = .6). Any emerging differences between conditions could thus not be due to differences across conditions in first letters or length. The stimuli used in Experiment 5.2 are listed in Appendix H. Two words appeared in the experiment that were removed later because their initial letter was a vowel, whereas the other ones were all consonants.

Procedure. The experiment consisted of two stages. In the first stage, beginning and fluent readers were asked to read the 28 training words 3 times in a random order, with the restriction that the same stimulus was never repeated without a different one interspersed.

The experiment was run on a Macintosh Classic computer. Naming times were registered with a voice-key and a millisecond timer and responses were evaluated by the experimenter.

Immediately after the training stage the subjects took part in the first-letter task. Beginning readers responded by naming the first phoneme of every stimulus, and fluent readers by naming the letter name (details of the procedure and response evaluation are described in Experiment 5.1). Prior to both parts of the experiment subjects received 10 practice trials.

Subjects. From a population of 241 children of Grade 1 a group of 20 children was selected with a reading score (Caesar, 1975) that was above average (46.6, $SD = 6.1$). The group of fluent readers ($N = 20$) was again recruited from the student population of undergraduates of the Department of Psychology of the University of Amsterdam. None of the subjects participating in Experiment 5.2 had been a subject in Experiment 5.1.

Results

Only the data of the main part of the experiment, the first-letter task, will be reported because the training phase had only one purpose: to familiarise the subjects with a set of words, the 'training' words. As in Experiment 5.1, the results of the beginning and fluent readers were analysed separately but discussed simultaneously.

The data of the first-letter task were screened. As in Experiment 5.1, responses were removed for: naming errors (children: 1.2%; adults: .15%); voice-key failure (children: 7.5%; adults: 3.4%); extremely long responses (more than 3 SD above the mean; children: 1.2%; adults: .7%); extremely short responses (less than 100 ms; children: 0%; adults: 0%).

An analysis of variance with the factor stimulus type (training word vs. legal pseudohomophone vs. illegal pseudohomophone vs. legal nonword vs. illegal nonword vs. well-known word) on the subject means revealed significant main effects of both the beginning and fluent readers (Children: $F(5, 95) = 17.17$, $p < .001$; Adults: $F(5, 95) = 7.82$, $p < .001$).

Further analyses showed identical results on post-hoc analyses (Newman-Keuls) and planned contrasts. Only the results of the post-hoc tests will be presented, all means are shown in Table 5.3. Beginning readers took significantly longer to name the first letter of an illegal nonword than that of any other stimulus ($p < .01$). Also, naming the first letter of an illegal pseudohomophone was significantly longer than it was for training words ($p < .01$) and well-known words ($p < .01$) and for legal pseudohomophones ($p < .05$). Fluent readers also named the first letter of an illegal nonword more slowly than that of any other stimulus ($p < .01$). For fluent readers no other differences between the stimulus types reached significance.

Table 5.3. Mean first-letter-naming latencies in ms for beginning and fluent readers in Experiment 2 (SD in brackets).

Stimulus type	Beginning readers	Fluent readers
Training words	748 (134)	510 (51)
Well known words	750 (128)	515 (58)
Legal pseudohomophones	770 (157)	514 (51)

Legal nonwords	791 (160)	517 (56)
Illegal pseudohomophones	811 (132)	517 (56)
Illegal nonwords	880 (180)	532 (67)

Discussion

Experiment 5.2 tested the potential relevance of four factors in the first-letter-naming paradigm, namely, lexicality, orthographic legality, phonologic lexicality and pronounceability (see Table 5.2). It appeared that the variable lexicality was not an explanatory factor for either subject group, because there was no difference in first-letter naming times between words and legal pseudohomophones. The variable phonologic lexicality was not an explanatory factor in the beginning and fluent reader groups either, because the differences in naming times between legal pseudohomophones and legal nonwords was not significant. Only in the group of beginning readers did the factor orthographic legality appear to be of critical importance: in this group, but not in the adult group, naming the first letter of an orthographically illegal pseudohomophones was slower than of an orthographically legal one. The absence of the effect in the adult group suggests that with increasing fluency, the significance of the orthographic legality factor diminishes.

When comparing the data of the beginning readers of Experiments 5.1 and 5.2, a noteworthy aspect is that only in Experiment 5.1 did processing time for words and legal nonwords differ significantly, whereas these stimulus types did not differ reliably from each other in Experiment 5.2. This could have been caused by a general greater variability in Experiment 5.2.

If our assumption is correct that also for fluent readers processing of words is faster than of legal nonwords (see 'Discussion' Section of Experiment 5.1), then it seems that a difference in processing words and legal nonwords is easier to detect with the first letter task in beginning readers than in fluent readers. We believe that differences between beginning and fluent readers mostly pertain to processing efficiency, that is, fluent readers process words and word like stimuli faster than beginners, but not so much to differences in processing.

As can be seen when comparing tables 5.1 and 5.3, latencies on words and legal nonwords in Experiments 5.1 and 5.2 are rather similar, whereas latencies on illegal nonwords are much longer in Experiment 5.2 than in Experiment 5.1. It is not clear why naming the first letter of illegal nonwords is particularly hard in Experiment 5.2. A suggestion is that the number of illegal nonwords in Experiment 5.1 (33.3 %) was higher than in Experiment 5.2 (14.3%), causing the subjects to adopt different strategies in the two experiments.

Notwithstanding the above contrasts, the results strongly suggest that pronounceability rather than orthographic legality (with which it has often been confounded), is the main factor underlying the first-letter effect. In both reader groups naming the first letter of an unpronounceable letter string was slower than for any other stimulus. Another conclusion bearing on the question how word recognition comes about, is that, both beginning and fluent readers, cannot help generating the pronunciation of letter strings they are presented with, even when these letter strings are highly irregular.

The finding that pronounceability is a relevant variable in the first-letter task is consistent with the results reported by Chastain (1981), Hawkins, Reicher, Rogers and Peterson (1976), and Mezrich (1973), all of whom used the Reicher-Wheeler paradigm (see 'Introduction' Section). Chastain showed that the word-superiority effect in the Reicher-Wheeler task depended on the pronounceability of the complete letter string (confounded with orthographic legality). Mezrich's study revealed that subjects executing the task, but instructed to pronounce the stimulus prior to performing the forced choice, did not show the common word-superiority effect, but a letter-superiority effect instead.

The most important procedural difference between the Reicher-Wheeler paradigm and the first-letter task is stimulus duration. In the Reicher-Wheeler paradigm, stimulus duration is individually set, and hardly ever exceeds 50 ms. In the first-letter paradigm, the stimulus remains on the screen until a response has been given. This procedural difference affects largely the choice of the dependent measure. In the Reicher-Wheeler paradigm the dependent variable is accuracy and in the first-letter paradigm it is response latency. Despite the differences between the Reicher-Wheeler task and the first-letter task, their results converge. With both tasks evidence was obtained that pronounceability of the stimulus determines performance.

Having shown that pronounceability is the main determinant of performance in the first-letter task and, consequently, that assembling the pronunciation of letter strings is an imperative process that cannot be prevented (the subjects actually being asked to ignore the context when naming the first letter), our next experiment provided a more direct test of the response-competition hypothesis. If response competition indeed underlies the first-letter effect, it should be possible to reduce the effect by letting the output of the automatic pronunciation process converge with the required response, thus decreasing the competition between the two processes. This was the approach taken in Experiment 5.3.

Experiment 5.3

If responding in the first-letter task is held up until the pronunciation of the stimulus as a whole is generated, similarity of the pronunciation of the first letter as part of the pronunciation of the complete stimulus and of the pronunciation of the required first-letter response should speed up responding. The reason is that, as compared to a situation in which the pronunciations of the response and the first-letter part of the complete stimulus do not match, the responses in this 'congruent' condition have been given a headstart. In Experiment 5.3 the responses in only one of the stimulus-group conditions in which a given group of subjects participates are given this headstart. To ascertain that no other differences between conditions could account for the effects, only orthographically legal pronounceable nonwords of three letters were used as stimuli.

Method

Materials. A set of 60 legal, pronounceable nonwords was created. Each stimulus consisted of one syllable containing three letters. The initial letter of stimulus was always a vowel, and was either an 'a', 'e', 'o' or 'u'. Twenty stimuli had a VCC-structure (for example "arg" [ɑrg]; 'single-cluster' stimuli), 20 had a V_1V_1C-structure (for example "aab" [ɑb]; 'double-

cluster' stimuli), and 20 had a V_1V_2C-structure (for example "auf" [ɑuf]; 'mixed-cluster' stimuli). The pronunciation of the vowel in VCC-stimuli is about the same as that of this vowel in isolation. So if phonemes have to be produced as responses in the first-letter task, the VCC-stimuli provide the congruent condition. The pronunciation of the double vowel in V_1V_1C-stimuli and that of the mixed vowel in V_1V_2C-stimuli differs from that of the isolated V_1 vowel. However, the pronunciation of the double-cluster (V_1V_1) in the double-cluster stimuli is about the same as the pronunciation of the name of the first-letter in these stimuli. So if letter names have to be produced as responses in the first-letter task, the V_1V_1C-stimuli provide the congruent condition. A frequency count of words having an initial vowel in the first three books of the reading curriculum indicated that the beginning readers had not had differential experience with these three types of initial clusters ($p >$. 35). Appendix I shows the stimuli used in Experiment 5.3.

Procedure. This experiment was run on a Macintosh Classic computer. Beginning readers were told that nonwords would appear on the screen and that they had to name the first phoneme of each stimulus. As in Experiment 5.1, half of the adult subjects were instructed to use the letter name to specify the responses and the other half was asked to use the phoneme. Naming times were registered with a voice-key and a millisecond timer. Every response was evaluated on correctness by the experimenter by pressing a key on the computer keyboard. Before the experimental session started every subject received 10 practice trials.

Subjects. Both beginning and fluent readers served as subjects. From the same population as in Experiment 5.1 two groups of beginning readers (21 medium and 20 poor readers) were selected, according to their scores on the reading test. None of these children participated in Experiments 5.1 and 5.2. The mean on the reading test (Caesar, 1975) of the medium readers was 25.3 ($SD = 4.9$), and of the poor readers it was 15.2 ($SD = 2.2$).

To ascertain that no other differences than those due to reading level could be responsible for any differences between the results of medium and poor readers that were to emerge, the results of the verbal (Bleichrodt et al., 1984) and a non-verbal intelligence tests (Raven, 1958) were taken into account. There was no significant difference between the poor and medium readers on the verbal intelligence test ($p > .10$), but the difference on the non-verbal test was significant, with the medium readers being superior (medium readers: 26.3, $SD = 9.2$; poor readers: 20.6, $SD = 4.6$), $F(1, 39) = 6.27$, $p < .05$. The effect of the non-verbal intelligence factor will later be tested in the statistical analysis.

The fluent readers taking part in this experiment were 40 students from the Department of Psychology of the University of Amsterdam. Again these subjects had not taken part in Experiments 5.1 or 5.2.

Results

The results of the beginning (children) and fluent readers (adults) will be analysed and discussed separately. As in Experiments 5.1 and 5.2, responses were removed for: naming errors (children: 2.0%; adults: 1.0%); voice-key failure (children: 4.5%; adults: .8%); extremely long responses (more than 3 SD above the mean; children: 1.3%; adults: .9%); extremely short responses (less than 100 ms, children: 0%; adults: .2%).

Beginning readers. Subject means were calculated for the three levels of the variable 'cluster': single, double and mixed. A 2 (reading level: medium vs. poor) by 3 (cluster) ANOVA on the means of the subjects showed significant main effects, but no interaction ($F < 1$). The main effect of reading level showed that medium readers (912 ms) were faster in naming the initial phoneme than poor readers (1041 ms), $F(1, 39) = 8.91$, $p < .01$. The main effect of cluster ($F(2, 78) = 26.19$, $p < .001$) revealed that the initial letter of single-cluster stimuli was named faster than of double-cluster and mixed-cluster stimuli (Newman-Keuls in both cases $p < .01$; the difference between double and mixed was not significant). In other words, responding was fastest in the congruent condition. The results are shown in Table 5.4.

A further analysis of variance was performed on the means of the subjects, with non-verbal intelligence as an additional between-subjects factor (high vs. low). The purpose of this analysis was to test for the possibility that non-verbal intelligence was responsible for the difference between poor and medium readers. The main effect of intelligence ($F < 1$) and the interaction effect did not reach significance, but the main effect of cluster again did, $F(2, 78) = 25.93$, $p < .001$. Because this effect was already discussed, it will not be commented upon further.

Table 5.4. Mean first-letter-naming latencies in ms for beginning and fluent readers in Experiment 3 (SD in brackets).

	Beginning readers		Fluent Readers	
	Reading level		Response type	
	Poor	Medium	Letter	Phoneme
Cluster type				
Single	979 (108)	863 (128)	438 (36)	468 (70)
Double	1068 (142)	927 (166)	431 (34)	481 (66)
Mixed	1076 (162)	946 (164)	451 (42)	499 (66)

Fluent readers. An error analysis on the data of the fluent readers indicated that subjects instructed to use phonemes made more errors (.95, $SD = 1.1$) than those who used letter names (.2, $SD = .41$), $F(1, 38) = 8.17$, $p < .01$.

A 2 (response type: letter vs. phoneme) by 3 (cluster: single vs. double vs. mixed) ANOVA was performed on the subjects' mean naming times. All main and interaction effects were significant. The main effect of response type indicated that subjects in the phoneme condition were significantly slower (483 ms) than subjects in the letter condition (440 ms), $F(1, 38) = 6.43$, $p < .05$. The main effect of cluster ($F(2, 76) = 22.45$, $p < .001$) revealed that mixed-cluster stimuli (475 ms; $SD = 60$) were named slower than single-cluster (453 ms; *SD* 57) and double-cluster (456 ms; $SD = 58$) stimuli, Newman-Keuls, $p < .01$. The important interaction between response type and cluster ($F(2, 76) = 5.19$, $p < .01$) showed that all means were significantly different from each other (Newman-Keuls, $p < .05$), except the one between single-cluster stimuli and double-cluster stimuli in the letter name condition. The results are presented in Table 5.4. A planned contrast between the single-cluster and double-cluster conditions in the letter condition showed a marginally

significant effect, naming the first letter of double-cluster stimuli being faster than of single-cluster stimuli, $F(1, 19) = 3.31$, $.05 < p < .10$.

Discussion

The results indicated clearly that beginning and fluent readers are faster in naming the first letter (i.e., a vowel) of a pronounceable nonword when the response is congruent with the way the vowel part of the word would normally be pronounced than when it is incongruent with its pronunciation in this context. For the children, who used phonemes to specify their responses, the single-cluster stimuli resulted in the shortest response times. The critical characteristic of these stimuli is that their first phoneme is, in fact, the response to be produced. There is no such congruence between the first phoneme of the stimulus and the required response in the case of double-cluster and mixed-cluster stimuli. The effect occurred for both poor and medium beginning readers. The only difference between these subject groups was that the overall response time was shorter for the medium readers than for the poor readers. The faster responses on the single-cluster stimuli cannot be due to differential familiarity with the three cluster types because a frequency count on the children's curriculum books had shown that the three cluster types occurred in these books equally often statistically.

With adult subjects the double-cluster stimuli had the shortest response times and smallest number of errors when their task was to produce letter names as responses. Again, what seems critical is the congruence between the response to be produced and the first phoneme of the stimulus as a whole. With letter naming as the subjects' task, the double-cluster stimuli, not the single-cluster stimuli, provide this congruence. In contrast, when the subjects' task was to produce phonemes as responses (an unnatural task for adult readers, hence the overall longer response times), the single-cluster stimuli resulted in the fastest responses, as they did with beginning readers as subjects. In other words response times were again shortest in the congruent condition. All in all, the general pattern of results is very consistent. The main determinant of performance appears to be whether or not the response is congruent with the pronunciation of the first phoneme in the stimulus.

The results of Experiment 5.3 clearly support the response-competition account of the first-letter effect. Furthermore, Experiment 5.3 (and Experiment 5.2) reveal the nature of the process that competes with the letter-name task: The competitor is a process by which the pronunciation is generated of the stimulus that is the carrier of the to-be-named letter. This pronunciation process appears to operate automatically, because the first-letter task per se does not require generating the pronunciation of the complete stimulus and the subjects are in fact asked to ignore the context of the first letter. A further relevant feature of this pronunciation process to stress is that it operates on nonwords too (all stimuli in Experiment 5.3 were of this type), that is, letter strings that are not represented lexically. The pronunciation of these stimuli can thus not be directly addressed, but has to be assembled (see also Van Orden, Johnston, & Hale, 1988). In sum, the competitor process, at least for nonword stimuli, appears to be a pronunciation assembly process that is enacted automatically.

But if the conclusion that the assembly process is automatic is correct, it will inevitably also be at work when words are encountered, either in an experimental setting or under more natural circumstances. A theoretical possibility that we will consider in more

detail in the 'General Discussion' Section of this chapter is that the pronunciation of words is both assembled and addressed directly. Anticipating the outcome of that discussion and in accordance with a number of other recent publications (see Introduction of this chapter): the data in the present study all seem compatible with the view that word pronunciations are only derived via the assembly process.

A final point to stress here is that, as in Experiments 5.1 and 5.2, beginning and fluent readers showed similar data patterns in corresponding conditions. Furthermore it was shown that differences in reading-decoding ability of beginning readers did not lead to differences in the relevant effects on the first-letter task. Poor readers were only generally slower than medium readers in naming the first letter. These results suggest that it is the same pronunciation-assembly process that underlies performance across the various subject populations.

General Discussion

One of the purposes of this research was to find out whether word recognition of beginning readers differed from that of more experienced readers. The first-letter-naming task as developed by Rossmeissl and Theios (1982) was applied to investigate the issue. The data of the adult subjects (fluent readers) in the letter-naming condition of our Experiment 5.1 replicated the results of Rossmeissl and Theios: naming the first letter of words and legal nonwords was faster than naming the first letter of illegal nonwords. However, the results were different for the fluent readers who specified the first letter by its phoneme (and also for the group of fluent readers as a whole): naming the first phoneme of words was faster than of legal nonwords which in turn was faster than naming the first phoneme of illegal nonwords. The beginning readers, who always used phonemes to specify their responses, showed the same pattern as the total group of fluent readers. Therefore we assumed that the difference between processing words and legal nonwords is a real one. We attributed the absence of the difference between processing words and legal nonwords in the adult letter-naming condition to a ceiling in performance.

Rossmeissl and Theios attributed their effect to a facilitatory influence of a word (or legal nonword) context on the identification of the stimulus' first letter. Instead, we suggested the possibility that response competition underlies the effects, and, consequently, that the effects in fact reflect inhibitory rather than facilitatory processing. The congruence effect in Experiment 3 provides particularly strong support for this interpretation, because the effects were all related to whether or not the pronunciation of the stimulus as a whole was congruent with the to-be-produced response: Response times were always shortest in the case of congruence, that is, under circumstances of the least response competition. By showing the relevance of congruence in pronunciation, Experiment 5.3 revealed the nature of the competitor process, after Experiment 5.2 had done the same with a different approach. A relevant conclusion to be drawn from both these experiments is that the competitor process entails assembling the pronunciation of letter strings. The fact that this process operates under the strict instructions to the subjects to ignore the context of the target letter leads to the further conclusion that this process comes about automatically whenever a letter string, legal or illegal, is encountered. Finally, the finding that all subject groups participating in this study, children as well as adults, child readers with poor as well

as medium skills, generally showed the same results constitutes a strong indication that this process operates independently of individual differences between readers in terms of age, experience and skill. To rephrase these conclusions in the terminology used in the introduction to this study, it appears that phonologic mediation is a mandatory process in all readers.

The results of this study thus clearly suggest that phonologic recoding is a primary process in word recognition of all readers. On the basis of this conclusion we can at least reject the suggestion of Humphreys and Evett (1985) that word recognition always comes about lexically (see introduction). It seems that we are also in a position to reject the view, held by proponents of the dual-route model, that fluent readers mainly employ the lexical route to identify words. The present data even evoke the question whether we could not do without the lexical route altogether, as proposed in a number of recent studies (e.g., Van Orden, 1987, 1991; Van Orden, Pennington, & Stone, 1990). The data of Experiment 5.2 indeed suggest we can. In this experiment response times did, not differ statistically for words and legal pseudohomophones. Had addressed phonology played a role, we might have expected shorter RTs for words than for legal pseudohomophones, because only in the case of words (being represented lexically, unlike pseudohomophones) both lexical and non-lexical reading may be operating. It may be expected that on average it would take longer for a process operating on its own to finish than it would take one out of two simultaneously ongoing processes to finish. The null-effect of lexicality in Experiment 5.2 thus seems to speak against the use of a lexical route.

However, in Experiment 1 a difference in response time for words and legal nonwords did appear, a finding which seems to suggest a role of lexical reading after all. However, this conclusion is not at all inevitable. The skill of assembling the phonology of a letter string is likely to improve with practice. The words presented in Experiment 5.1 had been encountered before both by the adults and the children (recall that these words were taken from the children's reading curriculum) whereas, of course, the legal nonwords were new to them. The difference between response times for words and nonwords may thus have come about because assembling the phonology of the words took less time than of the nonwords. Consequently, the subjects were ready to attend to the main task, first-letter naming, relatively soon in the case when a word stimulus was presented. Thus a lexicality effect does not necessarily falsify the hypothesis that phonologic mediation is the sole process via which pronunciation is assigned to words. All in all, this 'one-route' model provides the most parsimonious interpretation of the data of this study.

In short, this research shows that is probably necessary to discard the interpretation of the first-letter effect in terms of contextual facilitation and replace it with one of response competition. It also suggests that phonologic mediation is fundamental in reading. Obtaining positive evidence for phonologic coding in reading is relevant, because until now adherents of the other camp - who assume that word recognition, at least in fluent readers, comes about directly - have regarded the failure to obtain such evidence as support for their view (Van Orden, Stone, Garlington, Markson, Pinnt, Simonfy & Brichetto, 1992).

6

PHONOLOGIC MEDIATION IS FUNDAMENTAL TO READING: EVIDENCE FROM BEGINNING READERS*

Summary

Three tasks were employed to investigate the role of phonology in beginning readers. In a proofreading task the subjects had more trouble finding pseudohomophone-spelling errors (stimuli with a phonologic representation identical to a word) than spelling-control errors (stimuli that do not share a phonologic representation with a word). In a lexical-decision task they had more trouble deciding that pseudohomophones were nonwords than to do so in the case of spelling controls. Finally, in a semantic-categorisation task, subjects had more trouble rejecting pseudohomophones as a member of a designated category than to reject spelling controls. Differences between good and poor readers occurred, but appeared not to be due to differential use of phonology in word recognition. Instead, they were attributed to a difference between subject groups in efficiency of a spelling-verification mechanism. The results of the present beginning readers parallel those of the fluent readers of Van Orden et al. (1992). The main conclusion was that phonology plays an important role in word recognition in both beginning and fluent readers.

The role of phonology in reading has become an issue of extensive debate in the last few years. A certain role of phonology was never disputed because scholars believed that activation of the lexical item also activated its phonology. But what has become controversial is its role in word recognition. Adherents of the 'dual route' model assume that fluent readers come to recognise words (leastwise high-frequent) without the use of phonology. If, nevertheless, the phonology of the word becomes activated, it is supposed to have happened after recognition took place. This is known as 'addressed' phonology. A minor role for 'assembled' phonology (computing the phonology of the word in order to recognise it) is assumed when low-frequent words or pseudowords are encountered (Coltheart, Avons, Masterson, & Laxon, 1991; Jared & Seidenberg, 1991; Paap & Noel, 1991; Paap, Noel, & Johansen, 1992; Seidenberg, Waters, Barnes, & Tanenhaus, 1984).

The results of recent studies (amongst others: Daneman & Stainton, 1991; Lukatela & Turvey, 1993; Perfetti & Bell, 1991; Van Orden, 1987, 1991; Van Orden, Johnston, & Hale, 1988; Van Orden, Pennington, & Stone, 1990; Van Orden, Stone, Garlington, Markson, Pinnt, Simonfy, & Brichetto, 1992) strongly suggest however, that word recognition is <u>always</u> phonologically mediated. We will not, at this point, focus on the issue of how phonologic mediation is supposed to come about.[9]

* This chapter has been accepted for publication in *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*. It is co-authored by A. M. B. De Groot.

In the present study we direct our attention to the reading process of beginning readers. Our starting point is the word-recognition model of Van Orden (1987), in which he assumes that a visually presented stimulus, for instance, a word, automatically activates its phonology, which in turn activates its meaning (in the case of a non-homophone, e.g., 'stop') or several meanings (in case of a homophone, e.g., /roʊz/ for 'rows' or 'rose'). The meaning activation process is followed by a spelling-verification process, in which the spelling of the activated item is matched with the input spelling.

Unlike Jared and Seidenberg (1991) we do not want to suggest that Van Orden's model is simply a restatement of an early human information processing theory on reading (e.g. Rubinstein, Lewis, & Rubinstein, 1971). Van Orden clearly states (1991, p. 77) that phonologic mediation is not a way station between orthographic and lexical representations. No explicit representations are assumed and meaning is not retrieved, but created. A massively, interactive, adaptive, dynamic system is supposed to code word and nonword phonology (Van Orden, 1987; Van Orden et al. 1990, 1992).

The purpose of this study is to answer two closely related questions. The first goal is to establish whether beginning reading is also characterised by a process that proceeds from spelling to sound to meaning. Earlier research (Backman, Bruck, Hebert, & Seidenberg, 1984; Reitsma, 1983a, 1984) suggest that beginning reading is characterised by a heavy reliance on phonologic recoding which diminishes with increasing experience. Other researchers (Barron, 1986; Coltheart, Laxon, Keating & Pool, 1986; Rayner & Pollatsek, 1989), however, assume that beginning reading progresses from a purely lexical way of reading (without phonologic mediation) to one where both lexical and nonlexical (with the use of phonology) routes to the lexicon are used.

The second issue pertains to differences in word recognition between beginning and fluent readers. If word recognition in beginning readers resembles that of fluent in that both novice and expert readers show clear signs of phonologic mediation, the question remains what distinguishes a beginning reader from a fluent one.

If phonologic mediation is fundamental to reading for both beginning and fluent readers, differences between them could only occur, according to Van Orden's word-recognition model, in the verification process. Creating meaning (in terms of Van Orden) on the basis of phonology is a skill beginning readers have already developed while listening to speech in the interaction with other human beings. But beginning readers may not yet have developed an efficient spelling-check mechanism that enables them to correctly verify the orthography of a word. Therefore, we expect that differences in reading skill do occur and that they relate to more or less efficient verification processes. This hypothesis will be tested by investigating the differential effects on good and poor beginning readers in three tasks that tap word-recognition processes.

Even though all subjects in this study were attending Grade 1, relatively large differences in reading skill are apparent. In three out of four experiments (Experiment 2 being the exception) we divided the total group into a group of good and a group of poor readers. We expect a difference in overall performance between reader groups (with good readers performing better on all tasks than poor readers), but also that both reader groups show the same phonologic effects as the fluent readers of Van Orden. To be able to compare the results of beginning and fluent readers, the same experimental tasks (proofreading,

lexical decision and semantic categorisation) as used by Van Orden and his co-workers (1991, 1992) were used here.

In the proofreading task (Van Orden, 1991; Van Orden et al. 1992), subjects were instructed to read a text and mark each wrongly spelled word that appeared in it. One condition only contained pseudohomophones (a nonword phonologically identical to a base word; e.g., 'munth' for 'month') and the other only spelling controls (a nonword with similar orthography as, but without perfect phonologic correspondence to its base word; e.g., 'elbot' for 'elbow'). Similarity of orthography between pseudohomophones and spelling controls on the one hand and the base words on the other is supposed to be guaranteed by the OS-index (Orthographic Similarity) developed by Van Orden (1987). This is an estimate based on Weber's measure of graphic similarity (Weber, 1970). The OS-index ranges from 0 to 1, where 0 indicates no orthographic similarity between stimuli and 1 indicates identical orthography.

Positive evidence for the hypothesis that phonology mediates word recognition is obtained when a pseudohomophone (identical phonology to the base word) is more likely to activate the lexical representation of the base word than a spelling control (no perfect phonologic correspondence to the base word). If reading a pseudohomophone leads (erroneously) to lexical activation, it seems inevitable to conclude that computed phonology was responsible for it because pseudohomophones are not stored in the mental lexicon. Missing spelling errors in a proofreading task is taken as an indication of erroneous lexical activation.

The question that remains is: what actually goes wrong during the reading process when misspellings are not detected by subjects. According to Van Orden reading is always mediated by phonology, and also involves a spelling-verification check. In case a readers comes across a spelling-control stimulus, which has a phonologic representation that is relatively dissimilar from any real word, both the outcome of the phonology process and of the verification check indicate error. On the other had, when a reader is presented with a pseudohomophone stimulus, which shares its phonologic representation with a real word, the outcome of the phonology process suggests that the stimulus is a real word, whereas the spelling-verification mechanism indicates an error. In the latter situation only the spelling-verification mechanism can solve the contradiction. A failing spelling-verification mechanism thus seems to be responsible for missing spelling errors.

Two results obtained by Van Orden and colleagues suggested a fundamental role for phonology. First, adult subjects generally missed more pseudohomophones than spelling controls; second, pseudohomophones derived from low-frequent base words ('sleat' for "sleet") were missed more often than those derived from high-frequent base words ('grean' for "green"), whereas the mean number of correctly identified errors was the same for spelling controls derived from low-frequent base words ('speet' from "sleet") and those from high-frequent base words ('greln' from "green"). They took these effects as strong evidence for a primary role of phonology in reading. Because a frequency effect is taken as a signature of lexical access (Van Orden, 1991, p. 82), the interaction between stimulus type (pseudohomophones vs. spelling controls) and base word frequency suggests that pseudohomophones can activate the lexical representation of the base words.

The proofreading paradigm applied by Van Orden et al. (1982) was used here to study the reading process of novice readers (Experiments 6.1 and 6.2). Children with about eight

months of formal reading instruction were presented with texts they had to proofread. Apart from manipulating reading skill three other variables were manipulated, relating to three specific questions posed in this study. The first of these was whether proofreading a coherent text is different from proofreading a list of unconnected words.

Van Orden (1991) and Van Orden et al. (1992) only used stories to investigate the role of phonology in word recognition. Therefore, it is not clear whether the difference in error detection between pseudohomophones and spelling controls also emerges in a condition where the spelling errors are embedded in a list of unconnected words. Ehri and Roberts (1979) showed that careful inspection of words is more likely to happen in lists than in stories (but see also Cohen, 1980). A reason for this is that a reader proofreading a text (henceforth: story) could, for instance, because of on-going text integration processes, get more distracted from the task than when reading a list of unconnected words (henceforth: list). This distraction factor may have a more detrimental influence on the detection of pseudohomophones compared to spelling controls. The reason is that in the spelling-control condition the output of the two relevant processes are congruent, whereas in the pseudohomophone condition they are incongruent (see above). If, however, the assumed spelling-check mechanism is responsible for the difference between pseudohomophones and spelling controls this difference (possibly somewhat smaller) should be contained in the list-condition. We take the list condition as a more stringent test for the phonologic mediation hypothesis than the story condition.

The second new question of this study concerned the strategy employed by the subjects while proofreading. Van Orden et al. (1991, 1992) used a between subjects design. A subject either had to proofread a story containing solely pseudohomophones or one that only had spelling controls. In our first experiment a mixed-list condition was included: half of the target words were pseudohomophones and the other half were spelling controls. The question to be answered here is whether readers perform the proofreading task differently in case they only come across one type of error (pseudohomophones or spelling controls) as compared to a condition in which they are presented with both types. An interaction between type of stimulus and list homogeneity would indicate differences in strategy.

The third new issue concerns the influence of pronunciation during proofreading on the detection of spelling errors. Stated differently, does silent reading have a different effect upon the detection of misspellings than reading aloud? According to dual-route theory it is not necessary to activate the phonology of the words during silent reading, but it is obligatory to do so when one has to read text out loud. However, according to Van Orden's phonologic coherence hypothesis (Van Orden et al., 1990) activating the phonology of words can never be prevented. If true, we would not expect a significant interaction between task (silent reading vs. reading out loud) and stimulus type. If on the other hand silent reading does not always involve phonologic activation, the 'silent' and 'aloud' conditions should differ in that in the aloud condition more spelling controls and/or less pseudohomophones will be detected than in the silent condition.

In the lexical-decision task administered by Van Orden et al. (1992) subjects were asked to decide whether presented letter strings constituted words or not. The experimental stimuli again contained pseudohomophones and spelling controls. The subject analysis revealed a significant difference between the mean number of correctly rejected pseudohomophones and spelling controls, with the rejection score being higher for the

latter. As in the proofreading data, a significant interaction occurred between stimulus type and frequency, with only an effect of frequency for the pseudohomophones.

The occurrence of a frequency effect in lexical decision evidences lexical access in that task. However, it is uncertain whether meaning is retrieved in lexical decision. Therefore, Van Orden et al. (1992) also ran a semantic-categorisation experiment, a task that cannot be performed without the access to meaning. Subjects in that task have to decide whether a particular stimulus is a member of a designated category. The same pattern of results as in the lexical-decision experiment emerged. Experiments 6.1 and 6.2 of this study employed the proofreading task of Van Orden et al. (1992), Experiment 6.3 the lexical-decision task and Experiment 6.4 the semantic-categorisation task.

Experiment 6.1

The main goal of this experiment is to replicate and extend the proofreading results of Van Orden (1991) and Van Orden et al. (1992) with beginning readers, and to investigate the effects of context (story vs. list), homogeneity (homogeneous list vs. mixed list) and reading skill in the proofreading task. Spelling knowledge of the relevant stimuli was assessed after completion of the proofreading task.

Method

Materials. The first three reading books of the reading curriculum served as the source for the selection of the experimental stimuli. Twelve one-syllable Dutch words with a mean length of 3.6 letters (*SD* = .9) were chosen. Half of these words were high- and half low-frequent (according to their frequency count in these three books). From each selected word a pseudohomophone and a spelling-control stimulus was created. Mean OS[10] between high-frequency words and pseudohomophones was .58 (*SD* = .11); between low-frequency words and pseudohomophones it was .62 (*SD* = .11); between high-frequency words and spelling controls it was .68 (*SD* = .07) and between low-frequency words and spelling controls it was .67 (*SD* = .09). Note that the higher OS is for the spelling controls, which provides a bias against the phonologic mediation hypothesis. The words, pseudohomophones and their spelling controls are listed in Appendix J.

Two stories, "In the house" (149 words) and "In the woods" (147 words), were made up from words that had appeared in the curriculum, and were therefore known by the children. Target stimuli (selected words) were changed either into pseudohomophones (story-PsH condition) or into spelling controls (story-SpC condition).

Subsequently four lists of 60 stimuli each were created. All lists contained 48 words and 12 nonwords. List 1 consisted of the 48 words intermingled with the 12 pseudohomophones (List-PsH condition). List 2 contained the 12 spelling controls (List-SpC condition) in addition to the 48 words. Lists 3 (Mix-A condition) and 4 (Mix-B condition) were made up of the 48 words mixed with six pseudohomophones (three were derived from the original low-frequent words and three from high-frequent words) and the six spelling controls that were not their yoked ones. To control for number (in all conditions 12) and distribution of wrongly spelled words across lists, two mixed lists had to be developed. Of each list randomised versions were created.

Procedure. A child was either assigned to the story or to the list condition. Half of the subjects in the story condition was asked to proofread the two stories with only pseudohomophone substitutes; the other half was asked to read the stories with only spelling-control substitutes. In the list condition a child had to proofread one of the four stimulus lists. In both the story and the list conditions the child was asked to read the page carefully and mark the mistakes (it either being an non-existing or a wrongly spelled word).

After proofreading all children took part in a forced-choice spelling task. They received an envelope with 12 strips of paper, each containing one of the words and its derived pseudohomophone. The children were asked to mark the incorrectly spelled words.

Subjects. A sample of 120 children of Grade 1 was drawn from a population of 241. Forty children were assigned to the story-condition (20 in the pseudohomophone and 20 in the spelling-control version) and 80 children took part in the word list condition (20 subjects per list). The selected group of children represented the average reading child in Grade 1. Children with an exceptionally high or low score on the reading test were not included in the sample. Because considerable differences still remain between the ones at the higher and those at the lower end of the scale, the group was divided in a group of good (32.8; SD = 4.2) and one of poor (21.6; SD = 3.5) readers according to their scores on the reading test (Caesar, 1975), $F(1, 118) = 253.28$, $p < .001$. Much care was taken to make sure that reading level was uncontaminated with condition, ($F < 1$)[11].

Results

The results of the spelling test will be discussed first, followed by those of the proofreading task.

Spelling-test results. The results of the forced-choice task showed that the children knew the spelling of the experimental stimuli; mean proportion correct was .92. A 2 (reading level: good vs. poor) by 6 (condition: story-PsH vs. story-SpC vs. list-PsH vs. list-SpC vs. Mix-A vs. Mix-B) by 2 (frequency: high vs. low) ANOVA was executed on the subject means, indicating a main effect of reading level, $F(1, 108) = 20.04$, $p < .001$. Good readers (.96) had a higher score on the forced choice task than poor readers (.88). Neither the main effects of condition ($F < 1$) and frequency ($p > .15$) nor any interaction effect reached significant levels.

Proofreading results. The effects of stimulus type, frequency, reading level and context will be tested by means of a 2 (reading level: good vs. poor) by 2 (context: story vs. list) by 2 (stimulus type: pseudohomophone vs. spelling control) by 2 (frequency: high vs. low) ANOVA on subject means of the proportion correctly identified errors. The main effect of stimulus type was highly significant, $F(1, 72) = 223.73$, $p < .001$. The children in the spelling-control condition detected far more errors (.80) than those in the pseudohomophone condition (.27). Neither the main effect of frequency ($F < 1$), nor the interaction between frequency and any other experimental variable yielded significant results. Good readers (.60) were better in finding errors than poor readers (.46), $F(1, 72) = 14.24$, $p < .001$. A significant interaction between stimulus type and reading level ($F(1, 72) = 12.59$, $p < .001$) indicated that good readers (.40) found more pseudohomophone errors than poor readers

(.14; Newman-Keuls, $p < .01$), but in the spelling-control condition both groups found an equal number of spelling mistakes (.80).

The effect of context showed that in the list condition (.62) more errors were detected than in the story condition (.44), $F(1, 72) = 26.78$, $p < .001$. Context interacted significantly with stimulus type, $F(1, 72) = 10.84$, $p < .01$. Subjects in the list condition found more pseudohomophone errors (.42) than those in the story condition (.12; Newman-Keuls, $p < .01$), but the number of correctly identified spelling-control stimuli did not vary with context (story: .76; list: .83). In both the story and list condition more spelling controls were detected than pseudohomophones (Newman-Keuls, $p < .01$). This result needs qualification, because the second order interaction between reading level, context and stimulus was also significant ($F(1, 72) = 11.60$, $p < .01$), which is depicted in Figure 6.1. Only in the list condition a significant interaction between reading level and stimulus type occurred, $F(1, 36) = 18.21$, $p < .001$. Statistically, good readers detected as many pseudohomophones (.67) as spelling controls (.84), whereas poor readers found significantly less pseudohomophones (.16) than spelling controls (.82). The non-significant interaction between reading level and stimulus type in the story condition showed that both the good and the poor readers detected less pseudohomophones (.12 for both readers groups) than spelling controls (.76 and .77 respectively). The interaction between reading level and context ($F(1, 72) = 15.00$, $p < .01$) showed that good readers were better in finding the errors in the list condition (.76) than in the story condition (.44; Newman-Keuls, $p < .01$), but that the poor readers were equally successful in spotting the mistakes in the story and list conditions (story: .44; list: .49).

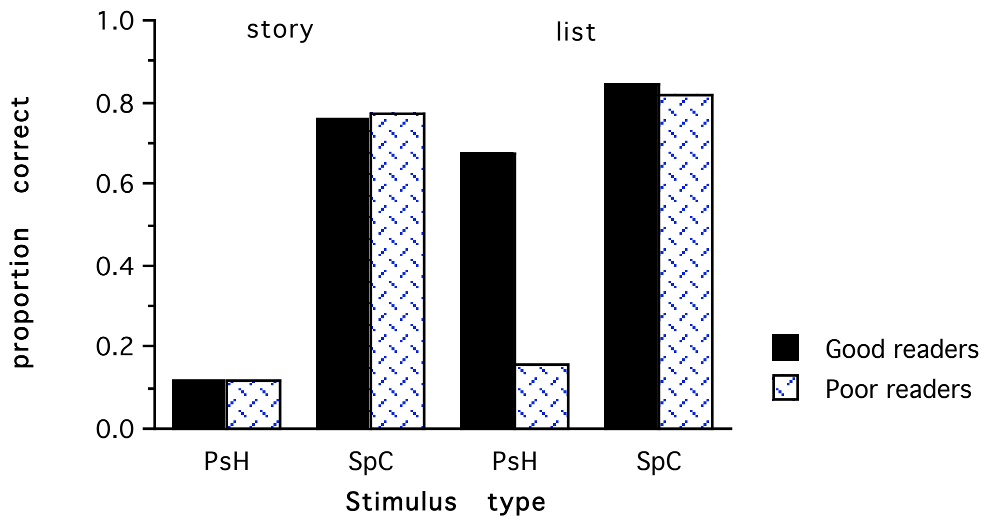

Figure 6.1. Proportion correctly identified pseudohomophones (PsH) and spelling controls (SpC) of the good and poor readers in the story and list conditions (Experiment 6.1: Proofreading).

Two separate analyses on the subject means of correctly identified pseudohomophones in the PsH-, Mix-A and Mix-B conditions and on the correctly identified spelling controls in the SpC-, Mix-A and Mix-B conditions were performed. Only the effects that pertain to the factor homogeneity will be discussed, because all other effects have been addressed before. The results showed that the mean proportion correctly identified pseudohomophones in the

homogeneous PsH-condition (.42) was the same as in the two mixed versions (Mix-A: .46; Mix-B: .35), $F < 1$. The mean proportion correctly identified spelling controls in the homogeneous SpC-condition (.83) was also equal to those in the mixed conditions (Mix-A: .79; Mix-B: .80), $F < 1$.

Discussion

The most salient result of Experiment 6.1 is that beginning readers in a proofreading task have more trouble finding pseudohomophone errors than spelling controls. This difference cannot be explained by insufficient spelling knowledge of the target stimuli, because the results on the spelling test (see 'Results' Section of Experiment 6.1) showed that both groups of readers knew the spelling of the target stimuli rather well. These findings thus strongly suggest that phonology mediates reading.

As was predicted in the introduction good readers showed superior proofreading behaviour. They were better in finding pseudohomophones than the poor readers, but good and poor readers were equally successful in finding spelling controls. Although, as shown in the forced-choice task, the spelling knowledge of the poor readers was somewhat weaker than that of the good, the difference between the groups (good: .96; poor: .86) appears too small to justify the conclusion that this difference in spelling knowledge caused the difference between the subject groups in error detection scores on the proofreading task. Instead, we conjecture the difference is due to a better developed spelling-verification mechanism in good readers.

Both the good and poor readers in the story condition and the group of poor readers in the list condition detected more spelling controls than pseudohomophones. Only the good readers in the list condition spotted as many pseudohomophones as spelling controls (see also Cohen, 1980). It seems that this latter group of good readers has developed an adequate spelling-verification mechanism that meets the requirements of this condition. However, also spelling verification in the good readers apparently has not yet reached an optimal level. This shows from the fact that performance of these readers on pseudohomophones deteriorates in the story condition, that is, in the condition where task demands are enhanced. In the story condition, good and poor readers alike, seem to have to deal with an additional factor, namely, directing attention to a second task, text integration. Assuming that both tasks require conscious control, less attention will be available for the actual one, error detection, in the story condition compared to the list condition. It seems plausible that this is causing the generally lower detection rate of pseudohomophones.

The absence of an overall effect of homogeneity and of an interaction between this variable and stimulus type indicates that the effects that do occur are unlikely to be due to the choice of particular strategic behaviours. A final noteworthy finding is that, unlike in Van Orden et al. (1992) we found no hint of a frequency effect nor an interaction between frequency and stimulus type. We will pick up this finding in Experiment 6.2 below.

In sum, this proofreading experiment strongly supports the phonologic mediation hypothesis. The effects established here are similar to those found with fluent readers in the studies of Van Orden et al (1992). Although differences between good and poor readers did show up, both groups appeared to be reading with the use of phonologic mediation. The differences between groups were explained in terms of a better developed spelling-verification process in good readers. The most striking differences between the results of the

present beginning readers and the fluent readers of Van Orden et al. (1992) are that beginning readers, unlike the fluent readers of Van Orden et al., did not show an effect of frequency nor an interaction between frequency and stimulus type.

Experiment 6.2

The main purpose of this experiment was to investigate the role of overt pronunciation in proofreading. A second goal was to try to replicate the results of Experiment 6.1, and to enforce the interaction that was surprisingly absent in Experiment 6.1, namely the one between frequency and stimulus type. By means of a spelling training prior to the proofreading task it was possible to control for differential presentation with target stimuli. The frequency factor was thus experimentally manipulated.

Method

Materials. The stories used in Experiment 6.1 were employed in this experiment. Details on the materials can be found in the 'Materials' Section of Experiment 6.1.

Procedure. Prior to the proofreading task, all children took part in a spelling training. Recall that all children had encountered the target words before in their reading curriculum. In groups of 2, 3 or 4 the children were asked to copy a list of 30 words. Six of the words (three of which occurred frequently in the curriculum and three infrequently) had to be copied four times (High-frequent condition) and the other six (again three of which occurred frequently in the curriculum and three infrequently) just once (Low-frequent condition). The assignment of words to the frequency condition was counterbalanced. After the subjects finished the copy task their knowledge of the words' spelling was checked with the help of a forced-choice spelling test, identical to the one in Experiment 6.1. They were asked to mark which of the two spellings (the word itself or its pseudohomophone) was the wrong one. After the spelling training they went back to their classroom. After a break (minimum interval was half an hour, maximum was three hours) half of the children had to proofread the stories silently. The list condition of Experiment 6.1 was excluded from the present experiment. They were tested again in groups of 2, 3 or 4, and were asked to read the two stories and mark each mistake they encountered (it either being a non-existing or a misspelled word). The other half of the children participated in the reading out loud condition. They were tested individually, and had to read the story to the experimenter and were also asked to mark every mistake they came across.

Subjects. A sample of 40 children of Grade 1 from a population of 192 participated. None of the subjects in this experiment took part in Experiment 6.1. As in Experiment 6.1 only children with an average score on the reading test (27.1; *SD* = 3.7) were included Reading level was not manipulated in this experiment.

Results

The data of the spelling test will be described first, followed by those of the proofreading task. Finally the effect of the preparatory spelling training on the results of the proofreading task will be tested by comparing the results of Experiments 6.1 and 6.2.

Spelling-test results. An analysis of variance on the mean correct choices showed that the children knew the spelling of the words they practised only once (low-frequent words: .95 correct) equally well as the ones they had practised four times (high-frequent words: .96 correct), $F < 1$.

Proofreading results. A 2 (task: silent vs. aloud) by 2 (stimulus type: pseudohomophone vs. spelling control) by 2 (frequency: high vs. low) ANOVA was performed on the subjects' mean number correctly identified errors. The effect of stimulus type was significant, $F(1, 36) = 17.08$, $p < .001$. Subjects found significantly more spelling controls (.77) than pseudohomophones (.46). None of the other experimental variables interacted significantly with stimulus type. Performing the proofreading task silently (.56) or aloud (.67) did not have a differential effect ($p > .10$) on the mean number of correctly identified errors, nor was there a significant interaction between task and stimulus type. Experimental stimuli that corresponded to high-frequent words (practised four times in the spelling-training task) were equally often ($F < 1$) correctly identified in the proofreading task as their low-frequent counterparts (practised once in the spelling-training task), .60 and .63 respectively. The factor frequency did not interact with any of the other experimental variables. The results are shown in Figure 6.2

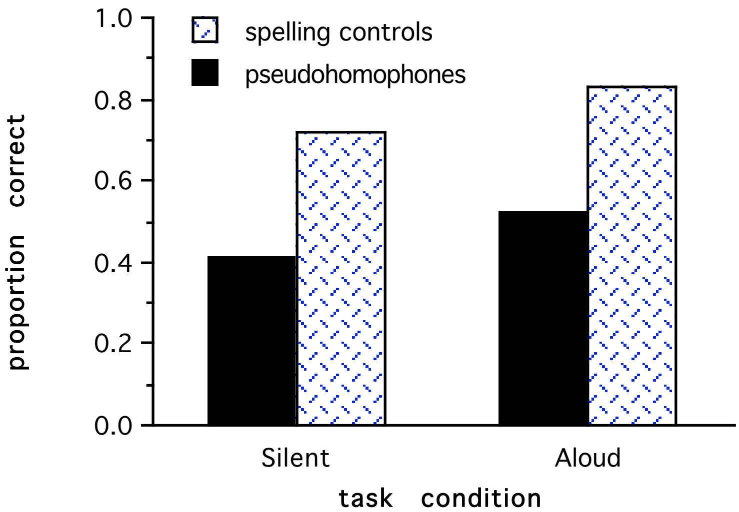

Figure 6.2 Proportion of correct no-responses to pseudohomophones and spelling controls in the silent and aloud conditions of Experiment 6.2

Effect of a preparatory spelling training. A 2 (spelling training: absent vs. present) by 2 (stimulus type: pseudohomophone vs. spelling control) ANOVA was performed on subjects' mean number correctly identified errors of the subjects of the story condition of Experiment 6.1 (absent condition) and of the total group of subjects of Experiment 6.2 (present condition). Both the main effects and the interaction appeared to be significant. More spelling controls (.77) were found than pseudohomophones (.29), $F(1, 76) = 125.25$, $p < .001$. Subjects who received a spelling training (.62) spotted more errors than those who did

not (.44), $F(1, 76) = 17.44$, $p < .001$. The interaction between stimulus type and spelling training ($F(1, 76) = 15.44$, $p < .001$) revealed that subjects who received a spelling training (.46) did find more pseudohomophones than those who did not (.12), but both groups detected an equal number of spelling controls (.77 and .76 respectively).

The analysis on the data of the spelling test showed a marginally significant main effect of spelling training ($F(1, 78) = 3.50$, $.05 < p < .10$), suggesting that the group who received the spelling training (.96) knew the spelling better than those who did not (.92).

Discussion

The results of Experiment 6.2 mimic those of Experiment 6.1. All subjects obtained a high score on the spelling test (more than 95% correct). This indicates that they knew the spelling of the target stimuli very well. Again less pseudohomophone errors were detected than spelling controls, and no sign of a frequency effect nor an interaction between frequency and stimulus type emerged, despite the preparatory spelling training. Furthermore, it appears that reading aloud does not change the reading process qualitatively, because proofreading the stories silently had the same effect on the error detection task as reading them aloud. We take this as further support for our assumption that the reading process of this group of beginning readers is phonologically mediated.

As shown, spelling training did improve the error detection score of the pseudohomophones in the proofreading task (but not those of the spelling controls). The reason why spelling training improved performance may be that subjects who received such a training were more wary of possible pitfalls. Nevertheless, the effect of stimulus type occurred in these subjects, albeit reduced.

Once more no frequency effect materialised. We oppose the view that assessment of the levels of the frequency variable was such that it could not render effects. It may seem too optimistic to assume that copying a word four times constitutes a high-frequent condition, whereas copying it once makes the word low-frequent. The results of Chapters 2 and 3 show that frequency effects are established with a rather limited number of presentations (0 vs. 1, and 1 vs. 2). To establish whether the null effect of frequency is stable over different types of tasks, the frequency factor will be investigated again in the experiments below.

Experiment 6.3

This experiment aimed at replicating and extending the lexical-decision results of Van Orden et al. (1992) with beginning good and poor readers. The same preparatory spelling training as in Experiment 6.2 was conducted to try again to enforce an interaction between stimulus type and frequency.

Method

Materials. The word list of Kohnstamm et al. (1981) was consulted for the selection of ten highly familiar one-syllable stimuli. Two different lists (A and B) were made up that contained all stimuli. In list A five words appeared four times (high-frequent condition) and the other five appeared only once (low-frequent condition). In list B the five high-frequent

words of list A were assigned to the low-frequent condition and the five low-frequent words became the high-frequent condition. The total number of words per list was 25.

The original ten words were changed into pseudohomophones (OS = .62; *SD* = .10) and spelling controls (OS = .66; *SD* = .05) with statistically equal OS, to serve in the lexical-decision task. The first three reading-instruction books of the reading curriculum offered 24 suitable word stimuli (12 high-frequent and 12 low-frequent). Finally twelve nonwords were created to serve as fillers. The actual list of stimuli (56) thus consisted of pseudohomophones (10), spelling controls (10), high-frequent (12) and low-frequent (12) well known words, and nonwords (12). The stimuli of the lexical-decision task are listed in Appendix K.

Procedure. The experiment consisted of three parts. Subjects started with the spelling training, identical to the one applied in Experiment 6.2. Children were asked, in groups of 2, 3 or 4, to copy in a note-book all 25 words of one of the lists. The spelling training was followed by a spelling test. The subjects received 10 strips of paper, each containing the correct spelling and an incorrect (pseudohomophone) spelling of the original word, and were asked to mark the incorrectly spelled words. The final part of the experiment was the lexical-decision task. After an interval of minimally half an hour and maximally three hours, the subjects were asked to carry out individually the lexical-decision task on a Macintosh Classic computer. Stimuli were presented on the computer screen, and the subjects were asked to decide as quickly as they could whether the stimulus was a word or not. Two buttons ('yes' and 'no') were connected to two separate serial ports of the computer, and subjects were asked to use their dominant hand for the yes-button. Latencies and responses were registered by the computer and the experimenter evaluated the validity of the responses and controlled the presentation of the next trial. Prior to the experiment proper the subjects received 8 practice trials to become familiar with the task.

Subjects. From a population of 192 children 20 new subjects were selected according to their scores on the reading test. Ten children were considered good readers (mean 31.7; *SD* = 4.8) and ten poor (mean 19.6; *SD* = 2.1).

Results

The results of on the spelling test will be described first, followed by those of the lexical-decision task.

Spelling-test results. A 2 (reading level: good vs. poor) by 2 (frequency: high vs. low) ANOVA on the subject means showed that the children generally knew the spelling of the target stimuli (mean number correct .95). Neither the main effect of reading level nor the interaction between reading level and frequency were significant. But there was a main effect of frequency, showing that both groups of readers had a significantly higher score on the high-frequent words (.98) than on the low-frequent words (.91), $F(1, 18) = 5.49$, $p < .05$.

Lexical-decision results. Proportions correctly identified errors, that is, correct nonword decisions, were calculated for high-frequent and low-frequent pseudohomophones and high and low frequent spelling controls.

A 2 (reading level: good vs. poor) by 2 (stimulus type: pseudohomophone vs. spelling control) by 2 (frequency: high vs. low) analysis of variance on correctly identified errors

was performed on the subject means. The main effect of stimulus type was significant, $F(1, 18) = 62.79$, $p < .001$. Spelling controls (.91) were more often correctly classified as nonwords than pseudohomophones (.54). Stimulus type and reading level interacted significantly, $F(1, 18) = 14.13$ $p < .01$, showing a significant difference between good (.70) and poor (.39) readers in the pseudohomophone condition, $F(1, 18) = 14.56$ $p < .001$, but not in the spelling-control condition (.89 and .93 respectively). Figure 6.3 presents the interaction. No main effect of frequency nor a significant interaction between frequency and any of the remaining variables appeared.

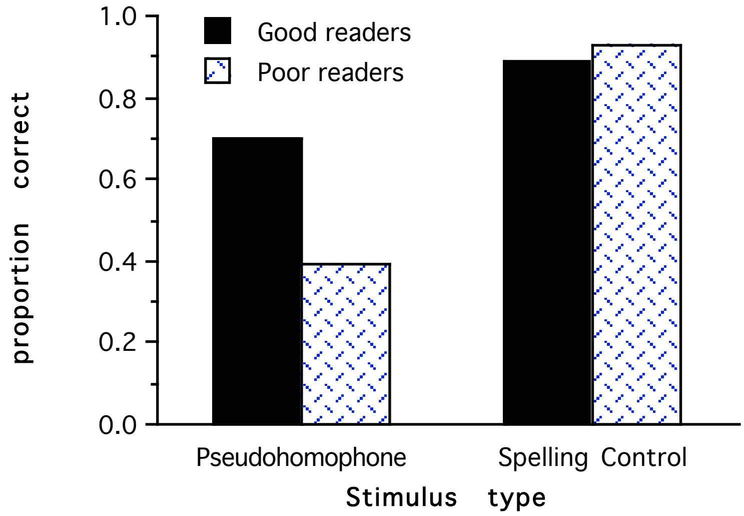

Figure 6.3. Proportion correct no-responses to pseudohomophones and spelling controls of the good and poor readers collapsed across the high and low frequency conditions (Experiment 6.3: Lexical decision).

Mean decision latencies of correct no-responses to pseudohomophones (54%), correct no-responses to spelling controls (91%), correct yes-responses to words (94%), and incorrect yes-responses to pseudohomophones (46%) were also computed.

Mean decision latencies of correct no-responses to pseudohomophones (3734 ms, $SD = 1000$) were statistically equal to those of correct no-responses to spelling controls (3886 ms, $SD = 988$), $F < 1$. Mean decision latencies of correct yes-responses to words (2177 ms, $SD = 622$), however, were significantly shorter than those of incorrect yes-responses to pseudohomophones (3627 ms, $SD = 1016$), $F(1, 18) = 33.74$ $p < .001$. Finally, the mean decision latencies of correct no-responses to pseudohomophones (3734 ms) were equal to those of incorrect yes-responses to pseudohomophones (3627 ms), $F < 1$.

Discussion

The results of the lexical-decision task[12] parallel those of the two foregoing proofreading tasks. More spelling controls were classified as nonwords than pseudohomophones, and again no effect of frequency nor an interaction between frequency and stimulus type emerged. It was also clear that all children knew the spelling of the target stimuli rather well (an average of 95% correct). As in the proofreading task good readers were better in finding

(here, correctly classifying as nonwords) the pseudohomophones than the poor readers, but the two groups were equally good in finding spelling controls.

Conclusions based on the latency data, particularly the Yes-response latencies, have to be approached with caution, first, because the condition means are based on a rather limited number of trials and, secondly, because it is highly debatable whether correct and incorrect 'yes'-responses reflect similar processes (see also Jared and Seidenberg, 1991, p. 366). Like the fluent readers in the lexical-decision experiment of Van Orden et al. (1992) our beginning readers also needed equal processing time to decide correctly that pseudohomophones and spelling controls were not words. On the other hand, however, our beginning readers needed considerably more time to decide incorrectly that pseudohomophones were words than to decide correctly that words were indeed words. Van Orden et al. (1988) got the same result in a semantic-categorisation task with experienced readers, but attributed this to outliers. Indeed, when they truncated the latency distributions the difference in response times between pseudohomophones and words disappeared. We did not apply this statistical method to the data of our beginning readers, because we are not convinced that a correct yes-response to a word is based on the same process as the incorrect yes-response to a pseudohomophone. One statistical fact supports the idea that this is indeed not the case: the standard deviation of the incorrect yes-responses was rather large (1016) as compared to that of the correct yes-responses (622).

What processes could underlie incorrect yes-decisions to pseudohomophones? Such a decision is clearly not easy, given the fact that it took the children nearly 1500 ms longer than to decide correctly that a word was a word. Subjects in this task appear to be getting "stuck" between the outcome of the phonologic mediation process, accepting because it sounds right, and the outcome of the spelling-verification process, rejecting because it does not look right (cf. Van Orden, 1987, p. 191). The fact that the children, despite long viewing times, were still not able to reject the pseudohomophones eventually, is taken as a further sign of a not fully developed verification process. Whereas the yes-response latencies may show the difficulty these children had accepting pseudohomophones, the no-latencies show their problems with rejecting them.

It was as difficult to reject pseudohomophones as it was to reject nonword spelling controls. Rejection of nonwords is difficult for these subjects because they tend to accept most word like stimuli as real words. One has to bear in mind that these beginning readers learn new words every day. It also took the subjects equally long to reject (3734 ms) as to accept (3627 ms) a pseudohomophone. These results also support the 'getting stuck' explanation.

To complete this study a semantic-categorisation experiment was executed. Unlike in lexical-decision, meaning evaluation is obligatory in semantic categorisation.

Experiment 6.4

The main goal of this experiment is to replicate and extend the results, obtained with fluent readers, of the semantic-categorisation experiment of Van Orden et al. (1992) with good and poor beginning readers. Prior to the semantic-categorisation task subjects received a preparatory spelling training (identical to the one of Experiment 6.3). In the present

categorisation task a subject is presented with a category name such as flower, followed by a stimulus, the target, that is either an exemplar of the category (e.g., 'rose') or is not an exemplar (e.g., 'dog'). The subjects are required to press a 'yes'-button if the target is an exemplar of the pre-designated category and a 'no'-button if it is not.

Method

Materials. The same lists (A and B) as developed for the spelling training in the lexical-decision task (Experiment 6.3) were used in this experiment to serve as materials in the spelling-training phase (see 'Materials' Section of Experiment 6.3). The stimulus sets of pseudohomophones and spelling controls were identical to those of Experiment 6.3. Twenty highly familiar words were chosen from the word list of Kohnstamm et al. (1981) to serve as fillers. Ten words were actual members of the category (exemplar condition) and ten were not (non-exemplar condition). The total list of stimuli consisted of 40 target stimuli (10 pseudohomophones, 10 spelling controls, 10 exemplar, and 10 non-exemplar words). The words were instances of ten different semantic categories. The stimuli are listed in Appendix J.

Procedure. This experiment also consisted of three stages. The spelling training and the spelling test were identical to those of 6.3 (see 'Procedure' Section). After an interval of minimally half an hour and maximally three hours, the subjects carried out the semantic-categorisation task individually on a Macintosh Classic computer. The procedure was a bit different from that used in experiments with fluent adult readers (Meyer & Gutschera and Meyer & Ruddy in Van Orden, 1987; Van Orden, 1987). In our experiment the experimenter named the semantic category to the child and immediately presented the target stimulus on the screen, by pushing a key on the keyboard. The child was asked to decide as quickly as possible whether the word on the screen was a member of the category just named by the experimenter. There were two reasons for conducting the experiment this way. The first was that reading the semantic category was thought to be too complicated in some cases, because these words are often difficult to read for beginners. The second reason was that the children might get confused having to change the response every other event (reading and remembering the category of the first stimulus, followed by reading and deciding on the membership of the second). The new procedure appeared to be satisfactory. Note that both pseudohomophones and spelling controls required a 'no' response. Prior to the experimental session subjects received extensive instruction and practice in the semantic-categorisation task. The experimenter was convinced that all subjects understood the task correctly.

Subjects. The beginning readers (N = 28) who participated in this experiment were all from the same classroom. They were also instructed according to the curriculum that was used by the children of all three foregoing experiments. However, the group that took part in Experiment 6.4 had changed to a revised edition. This meant that the initial instruction words differed from those in the old edition, but its basic principles remained unchanged. Again the group was divided in a group of good (29.3, SD = 4.3, N = 13) and poor (18.9, SD = 2.5, N = 15) readers based on the scores of the reading test (Caesar, 1975).

Results

The results of the analyses on the spelling test will be described first, followed by those of the semantic-categorisation task. Finally the lexical-decision data of Experiment 6.3 and the semantic-categorisation data of Experiment 6.4 will be analysed jointly to test for task effects on the error detection scores.

Spelling-test results. A 2 (reading level: good vs. poor) by 2 (frequency: high vs. low) ANOVA on the subject means showed that the subjects knew the spelling of the target stimuli (mean number correct .95). The main effect of reading level was significant ($F(1, 26) = 5.37$, $p < .05$) with the good readers (.98) being better on the spelling test than the poor readers (.92). The main effect of frequency and the interaction between reading level and frequency were not significant (both F's < 1).

Semantic-categorisation results. Analyses similar to those of Experiment 6.3 were performed on the data of this experiment. A 2 (reading level: good vs. poor) by 2 (stimulus type: pseudohomophone vs. spelling control) by 2 (frequency: high vs. low) ANOVA on the subjects' mean number of correct no-responses revealed only a significant effect of stimulus type, $F(1, 26) = 354.99$, $p < .001$. It indicated that more spelling controls (.98) were identified as non-exemplars than pseudohomophones (.32). Neither the main effect of reading level nor the interaction between stimulus type and reading level reached significant levels (both F's < 1). Figure 6.4 shows the results. The effect of frequency ($F < 1$) was also not significant, and no interaction occurred between frequency and any of the remaining experimental variables.

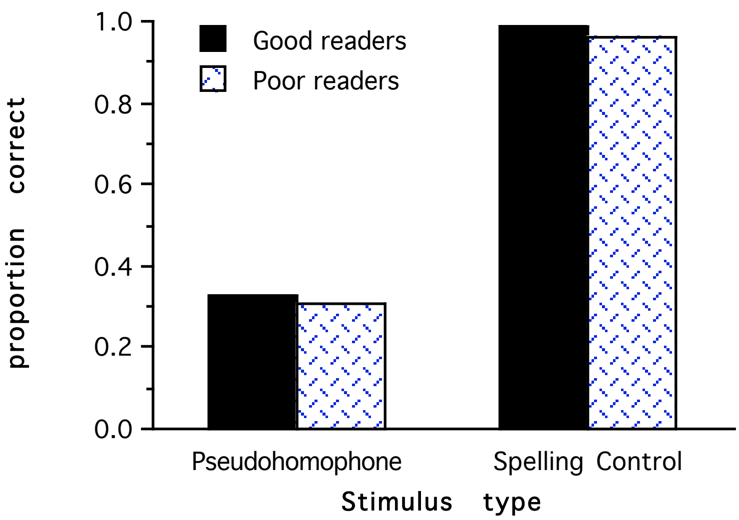

Figure 6.4. Proportion correct no-responses to pseudohomophones and spelling controls of the good and poor readers collapsed across the high and low frequency conditions (Experiment 6.4: Semantic categorisation).

Mean decision latencies of correct no-responses to pseudohomophones (31%), correct no-responses to spelling controls (97%), correct yes-responses to exemplar words (93%), and incorrect yes-responses to pseudohomophones (68%) were also computed.

Mean decision latencies of correct no-responses to pseudohomophones (3812 ms, *SD* = 1667) were statistically equal to those of correct no-responses to spelling controls (3937

ms, *SD* = 1340), $F < 1$. Mean decision latencies of correct yes-responses to exemplar words (2339 ms, *SD* = 672), however, were significantly shorter than those of incorrect yes-responses to pseudohomophones (3227 ms, *SD* = 1569), $F(1, 26) = 9.31$ $p < .01$. Finally, the mean decision latencies of correct no-responses to pseudohomophones (3812 ms) were equal to those of incorrect yes-responses to pseudohomophones (3227 ms), $p > .15$.

Data analysis of Experiments 6.3 and 6.4 combined. An analysis on the data of the subjects from the lexical-decision task and those of a random selection of 20 (from 28) subjects from the semantic-categorisation task was executed to test for an effect of task. A 2 (reading level: good vs. poor) by 2 (task: lexical decision vs. semantic categorisation) by 2 (stimulus type: pseudohomophones vs. spelling controls) ANOVA was executed on the subject means. The discussion of the results will be limited to those that pertain to the variable task. The main effect of task was significant, $F(1, 36) = 4.30$, $p < .05$. Subjects in the lexical-decision task (.73) rejected more pseudohomophones and spelling controls than those in the semantic-categorisation task (.66) The interaction between reading level and task did not reach significance, but the one between task and stimulus type did, $F(1, 36) = 19.53$, $p < .001$. The subjects in the lexical-decision task rejected significantly ($F(1, 36) = 11.99$, $p < .001$) more pseudohomophones (.54), but significantly ($F(1, 36) = 9.42$, $p < .01$) less spelling controls (.91) than those in the semantic-categorisation task (.33 and .98 respectively).

The second order interaction between reading level, task and stimulus type was also significant, $F(1, 36) = 5.42$, $p < .05$, showing that the interaction between reading level and stimulus type was significant in the lexical-decision task, but not in the semantic-categorisation task. This interaction reflects the presence of an interaction between reading level and stimulus type in Experiment 6.3 and its absence in Experiment 6.4.

To make sure that the subjects who participated in the lexical-decision task did not differ on the variable reading level from those in the semantic-categorisation task an analysis on their reading scores was performed. Neither a significant interaction between task and reading level ($F < 1$), nor a significant main effect of task occurred.

Discussion

The overall picture of the results of this semantic-categorisation experiment is almost identical to that of the lexical-decision experiment. The subjects participating in the semantic- categorisation task knew the spelling of the target stimuli equally well as those in the lexical-decision task (an average of 95% correct). The only difference is that in the semantic-categorisation experiment no differences occurred between good and poor readers on detecting pseudohomophones. Both groups had severe trouble detecting pseudohomophone misspellings, whereas identifying spelling controls was relatively easy for both groups. Again no sign of a frequency effect, nor an interaction between frequency and stimulus type showed up. Categorisation latencies on no-responses to pseudohomophones and spelling controls were identical, whereas latencies on incorrect yes-responses to pseudohomophones were much longer than on correct yes-responses to words. This latter finding is in line with the results of the semantic-categorisation task of Van Orden et al. (1988). An interpretation of this result was given in the 'Discussion Section of Experiment 6.3.

The analysis comparing the lexical-decision task with the semantic-categorisation task indicated that the subjects performing the lexical-decision task were superior in detecting pseudohomophones, but inferior in correctly identifying spelling controls (cf. Van Orden et al., 1992, p. 263-265; also for more detailed interpretation). It is interesting to note again (cf. the story versus list condition of Experiment 6.1), that when good readers are forced to perform a task that is again rather diverting, here semantic categorisation, their pseudohomophone detection score drops to the level of that of poor readers. This suggest that the spelling-check mechanism in this group of good readers is not yet fully developed, so that a slight increase in task complexity leads them astray.

General discussion

The general pattern of our results is clear-cut. In all tasks both good and poor readers had more trouble finding pseudohomophones than spelling controls. Generally, good readers were better in finding pseudohomophones than poor readers and both groups were equally successful in finding spelling controls. Latency data on both the lexical-decision and the semantic-categorisation task revealed the superior reading skills of the good readers; they were always faster. Latency data also showed the reluctance of the beginning readers in deciding about pseudohomophones; correct no-responses were as fast (or as slow, as the case may be) as incorrect yes-responses.

The present proofreading studies showed a rather robust pseudohomophone effect. None of the three new variables that were investigated here affected the task seriously. While it was clearly easier to detect pseudohomophones in a random list of words than in a coherent text, in both conditions the pseudohomophone effect emerged. Proofreading a text aloud or silently did not have a differential effect either on the error-detection scores; in both cases a pseudohomophone effect was apparent. Finally, it turned out that the beginning readers presented with a list that contained both types of errors (pseudohomophones and spelling controls) had the same processing strategy as those who read a list that had only one type of spelling mistake.

The above results of our beginning readers were almost identical to those of the fluent readers of Van Orden et al. (1992). For comparison, Figure 6.5 presents the main findings of our beginning readers and the corresponding results of the fluent readers of Van Orden et al. (1992).

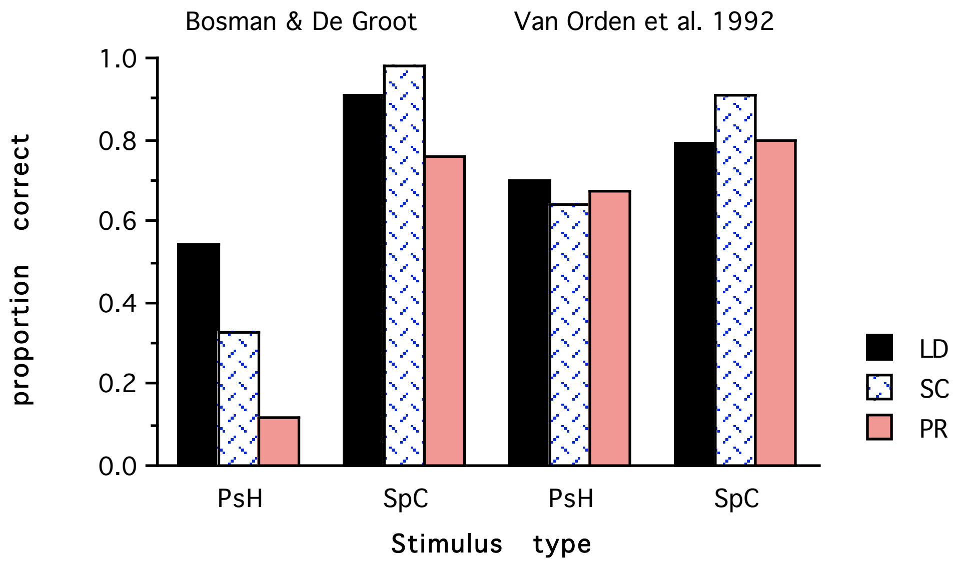

Figure 6.5. Proportion correct no-responses in lexical decision (LD) and semantic categorisation (SC) and proportion correctly identified errors in proofreading a story (PR). PsH = pseudohomophones; SpC = spelling control.

As can be seen, the pseudohomophone effect occurs in all three tasks and both for beginning and fluent readers. If the pseudohomophone effect as established here legitimates the inference that word recognition is phonologically mediated, then the conclusion seems warranted that phonologic mediation is mandatory in word recognition by all readers (see also Chapter 5). This conclusion deviates from that in much of the earlier research on word recognition in beginning and fluent readers (See 'Introduction' Section of this chapter, and below), where developmental differences in the occurrence of phonologic mediation is typically assumed.

However, we do not want to suggest that no differences whatsoever exist between the word-recognition processes of beginning and fluent readers. One difference between these reader groups appears to be that phonologic mediation comes about faster in fluent readers than in beginning readers. This shows from the fact that word recognition in beginning readers is generally slower than in fluent readers. A second difference is that spelling verification, following meaning activation, appears to be less well developed in beginning readers. This is suggested by the finding that beginning readers have more trouble finding pseudohomophone errors than fluent readers, whereas spelling-control errors are detected as often in beginning as in fluent readers (see Figure 6.5). This same pattern occurs when comparing the performance of good and poor beginning readers. Again, the less fluent group (the poor readers), shows worse performance on pseudohomophone errors, but similar performance (except in Experiment 6.4) on spelling-control errors (see Figures 6.1, 6.2a, and 6.3a). Note that these differences between reading groups are unlikely to be due to lack of spelling knowledge concerning the test words: The spelling test that was administered in all four experiments showed good spelling knowledge about these words in all subjects groups. In sum, it appears that spelling verification develops with increasing experience in reading. A more specific assumption is that spelling verification gradually automatises with reading experience; in beginning readers it still requires attention.

This assumption that spelling verification requires attention in beginning readers, is supported by the differential results across the three tasks used in this study. Attentional

processes are generally thought to be susceptible to disturbance (e.g., Schneider, Dumais & Shiffrin, 1984). Increasing tasks demands should therefore negatively affect attentional processing, here, spelling verification. This should become reflected in lower pseudohomophone detection rates under the most demanding experimental circumstances. In accordance with this prediction, pseudohomophone detection rates were higher in lexical decision than in semantic categorisation. The latter is plausibly the more demanding of the two, because as compared to lexical decision it requires an additional processing stage (obligatory access of meaning on every trial). Also in accordance with the prediction, pseudohomophone detection rates were better in the list condition than in the story condition of the proofreading task (Experiment 6.1). In the story condition text integration processes will consume part of the subjects' attentional resources. Thus, less attention will be available for the subjects' actual task, error detection. Therefore, spelling verification will suffer. As can be seen in Figure 6.6, pseudohomophone detection by the fluent readers of Van Orden et al. (1992) is about equally good in all three tasks. This indicates that spelling verification in these readers is much more automatised than in beginning readers.

As has been mentioned before, our data mimic those of Van Orden et al. (1992) to a large degree. The only consistent difference between their and our results is that we failed to find a frequency effect for pseudohomophones whereas in all three tasks with fluent readers (proofreading, lexical decision, and semantic categorisation) Van Orden et al. did obtain this effect. Van Orden (1991) takes the frequency effect as a signature of lexical access. From the failure to find such an effect in our data one could be tempted to conclude that no lexical access occurred in our beginning readers. But Monsell's view (1991) on the nature and locus of frequency effects suggests an alternative interpretation.

Monsell showed that the common generalisation that frequency effects are weaker in naming than in lexical decision (Balota & Chumbley, 1984; De Groot, 1989) needs qualification. The size of the frequency effect when naming stimuli that are exceptions to general linguistic rules (for instance, exception words and stress-final disyllables) approaches that of the effect in lexical decision. Only when stimuli are obeying linguistic rules, a weaker effect of frequency in naming turns up. His conclusion is that "the more that naming depends on transcoding at a lexical level, and the less on sublexical transcoding, between spelling and sound, the larger the frequency effect" (Monsell, 1991, p. 177). Phrased differently, the most likely locus of frequency effects is between identification of the orthographic pattern and the subsequent activation of the phonologic representation, inasmuch as one can discriminate between these two processes.

We do haste to state that we are not returning here to a symbolic interpretation in which a lexical and nonlexical route to word recognition are postulated, but rather interpret Monsell's view within Van Orden et al.'s (1990) subsymbolic model of word recognition. Assuming that beginning readers have a slower transcoding process between spelling and sound than fluent readers, one could regard the mediation process of a beginner as one that operates on the sublexical level, whereas the more efficient mediation process of fluent readers may be seen as one operating on the lexical level. In this framework, assuming the existence of explicit lexical representations would not be necessary. Monsell's analysis thus provides an explanation of our failure to find a frequency effect on pseudohomophones: Beginning reading depends more on sublexical coding and is therefore less susceptible or not susceptible at all to the frequency factor.

The results of this study add to the growing body of evidence that phonologic mediation is obligatory in word recognition. Not only is this new piece of proof obtained with beginning readers rather than fluent, but it is also found in a language other than English, namely, Dutch. As was said in the introduction, the controversy on how reading develops still exists. At one extreme we find the advocates of the idea that initially reading is characterised by a heavy reliance on phonologic recoding and that with increased experience with written material the role of phonology diminishes (Backman et al., 1984). Reitsma (1983b) points out that phonologic recoding becomes less evident in more experienced beginners. In the other camp are proponents of the viewpoint that the developmental shift is from accessing the mental lexicon directly (without phonologic mediation) to processing words directly as well as indirectly (see for a review Rayner & Pollatsek, 1989, pp. 364-382). Our results show that none of these positions is fully tenable. It seems that proponents of the first view are right in their assumption that beginning reading is dominated by phonology, but contrary to their belief the role of phonology does not vanish, nor does it seem to diminish. Support in the present data for the latter claim is that our good (beginning) readers showed the same phonologic effects as their less skilled classmates. Additional support comes from the similarities between our beginning readers and the fluent readers of Van Orden et al. (1992).

As was said in the General Introduction (Chapter 1), the Dutch language has a fairly shallow orthography. One could infer from this that the present findings were expected, because a phonologic strategy suits a shallow orthography. A role for phonology is indeed more obvious in a language like Dutch than in a language with a deeper orthography, but nevertheless a survey of studies on different languages and writing systems suggests that our results with beginning readers are not trivial. Studies of both shallow alphabetic writing systems like Serbo-Croatian (Lukatela, Carello, & Turvey, 1990) and deeper alphabetic orthographies like French (Ferrand & Grainger, 1992), English (among others: Lesch & Pollatsek, 1993; Perfetti & Bell, 1991), and Hebrew (Frost & Kampf, 1993; Navon & Shimron, 1981) all indicate a primary role for phonology (but see also for null effects, e.g., Baluch & Besner, 1991 for Persian; Tabossi & Laghi 1992, for Italian).

The evidence for phonology as a primary constraint on word recognition is even more compelling when it can be obtained from non-alphabetic writing systems like Chinese and Japanese. Indeed, Perfetti and Zhang (1991) showed that phonologic activation in Chinese readers occurs, and automatically so, and the work of Wydall, Patterson, and Humphreys (1993) suggests that reading logographic Japanese Kanji is achieved by parallel access to meaning from both orthography and phonology. They state that their results can be incorporated in a modified version of Van Orden's model on word recognition (1987).

Given the ubiquity of phonologic effects, irrespective of the type of orthography under investigation and, as stressed in this study, the reader groups tested and the tasks used, our answer to the question of Carello, Turvey, and Lukatela (1992): "Can theories of word recognition remain stubbornly nonphonological?" is a convinced : 'No'.

7

THE ROLE OF ORTHOGRAPHIC KNOWLEDGE IN READING AND THE NATURE OF THE SPELLING PROCESS*

Summary

In this study two predictions derived from the 'dual-route' model and the 'phonologic coherence hypothesis' were tested. The first hypothesis pertained to the reading process. Dual-route theory predicts that words of which subjects have orthographic knowledge are read faster than words of which subjects have no, wrong or incomplete orthographic knowledge. The phonologic coherence hypothesis, however, predicts equal reading times for both groups of words. The second hypothesis involved the spelling process. According to the dual-route model, subjects should produce similar proportions of phonologically correct and phonologically incorrect errors when spelling words they know the spelling of, but when spelling words they do not know the spelling of the proportion of phonologically correct errors is expected to surpass the proportion of phonologically incorrect errors. The phonologic coherence hypothesis, however, does not predict a differential effect of orthographic knowledge on the distribution of error types. The pattern of data obtained in a set of four experiments designed to test these hypotheses supported the phonologic coherence hypothesis and disconfirmed the dual-route model. It was concluded that a single-route model, that is, the phonologic coherence hypothesis, suffices to explain the word-recognition and spelling-generation processes investigated here.

Within the research domain of visual word recognition two classes of models can be distinguished, the single-route models and the dual-route models. Proponents of dual-route models postulate two functionally independent routes to word recognition, a lexical and a nonlexical route (Coltheart, 1978; Coltheart, Curtis, Atkins, & Haller, 1993; Paap, McDonald, Schvaneveldt, & Noel, 1987; Paap, Noel, & Johansen, 1992). Experienced readers are supposed to use both routes to read words. Reading via the nonlexical route implies that a subject has to compute the phonology of the word by application of grapheme-phoneme correspondence rules before recognition can occur. Reading by means of the lexical route, however, occurs without phonologic computation, because a direct match is made between a written word and its representation in the mental lexicon. The nonlexical route is also referred to as the indirect route, because phonology mediates the word-recognition process. The lexical route, in which no such mediation takes place, is also called the direct route.

Dual-route theory dominated the eighties, but the tide appears to be turning. In contrast to dual-route models, single route models assume the existence of just one route to

* This chapter has been submitted for publication with A. M. B. De Groot and M. Van Leerdam as second and third author respectively.

word recognition, either the direct route or the indirect one (Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990). Both the single-route model of Van Orden et al. (1990) and that of Seidenberg and McClelland (1989) claim that empirical facts about word recognition can be explained without assuming two functionally independent routes. If the claims made by single-route models are correct, both for theoretical and empirical reasons, the Van Orden et al. model seems preferable to that of Seidenberg and McClelland.

Theoretically, it is difficult to escape the impression that the Seidenberg and McClelland model (1989) is actually the old dual-route model in a new bottle. In fact they admit that their model is a dual-route model, because they distinguish between a direct and an indirect way of generating the pronunciations of words. Yet they argue that the difference between the traditional dual-route model and their parallel distributed processing-approach (PDP) is more than notational, because these two functionally independent processes are implemented within one and the same architecture, whereas the traditional dual-route models use different types of knowledge representations (pp. 558-559).

Empirically, ample evidence is available that phonologic mediation plays a central role in reading (amongst others: Daneman & Stainton, 1991; Lukatela & Turvey, 1993; Perfetti & Bell, 1990; Van Orden et al., 1990; Van Orden, Stone, Garlington, Markson, Pinnt, Simonfy, & Brichetto, 1992). Although both single-route models assume that phonology plays a role in word recognition, it is not fundamental in the Seidenberg and McClelland model (1989), whereas it is in the model of Van Orden et al. (1990). In the framework of dual-route theory one could say that of the two single-route models the model of Seidenberg and McClelland retains the direct route, whereas the model of Van Orden et al. retains the indirect route.

Presently, the validity of the dual-route model receives much attention. This study intends to contribute to that discussion. More specifically, it aims at investigating diverging predictions that can be derived from dual-route theory and from the single-route subsymbolic approach of Van Orden et al. (1990). Before introducing these predictions, the dual-route model and Van Orden's model will be described in some more detail.

In the dual-route model explicit representations of words in a mental lexicon are assumed. That is, the mental lexicon not only contains a representation of the phonology and meaning of words, but also an orthographic representation. Grammatical class of the word is also often assumed to be represented, but for present purposes that fact is irrelevant. If an orthographic representation of the word is present in the mental lexicon, a direct match between the written word and this representation suffices to recognise the word and activate its meaning. Whether subsequently the phonology gets activated (post-lexically) depends on the task at hand. In naming phonologic activation is obligatory, whereas in lexical decision it is not. If, on the other hand, no orthographic representation exists the reading process has to resort to the nonlexical route. That is, before recognition and meaning activation can take place the written word has to be recoded phonologically. One important axiom of dual-route theory that has to be made explicit is the assumption that the indirect route lags behind the direct route. That is, recognising a word through the indirect route takes longer than recognising it through the direct route, because assembling the phonology is supposed to be relatively time consuming (Van Orden et al. 1990; but see also Paap, Noel, & Johansen, 1992).

Advocates of the dual-route model have remained noticeably silent on the issue of the spelling process. Simon and Simon (1973), Kreiner (1992), and Kreiner and Gough (1990) propose that essentially the same routes as assumed in reading, are used in spelling. The lexical route in spelling implies that, analogous to the lexical route in reading, the spelling of a word is retrieved directly from the lexicon. Thus, the word's orthographic representation is accessed, and this representation is read out during spelling. This lexical process will only succeed when such a representation is indeed present. However, if an orthographic representation is not represented in the mental lexicon a subject needs to map the individual phonemes of the word to be spelled onto graphemes. This is the nonlexical route of spelling and is analogous to the indirect route of reading.

Two hypotheses were derived from the assumptions of the dual-route model. The first concerns the reading process. According to the dual-route model subjects should process words of which they have orthographic knowledge faster than words of which they have no, wrong, or incomplete orthographic knowledge. The ground for this prediction is that words of which an orthographic representation exists in the mental lexicon can be addressed via the fast direct route, whereas words of which no such orthographic representation is present necessarily have to be read through the slower indirect route. The second hypothesis pertains to the spelling process. Given the existence of an orthographic representation in the mental lexicon, subjects are assumed to read out the spelling pattern (also referred to as 'orthographic image'; Ehri, 1980). Since reading out the spelling of a word from the mental lexicon does not a priori guarantee that the phonetic form of the word will be preserved, the proportion of phonologically correct spelling errors should not exceed the proportion of phonologically incorrect errors. If no orthographic representation is present, the orthography has to be computed from its phonology, and it is predicted that in the spelling of words the proportion of phonologically correct errors exceeds the proportion of phonologically incorrect errors. It goes without saying that generally less spelling errors are expected on words of which subjects have an orthographic representation, than of words of which no such representation exists. Thus the second prediction from dual-route theory that will be tested is that the distribution of errors (phonologically correct vs. phonologically incorrect) is different within a set words of which the orthography is known as compared to a set of words of which the orthography is unknown.

The model of Van Orden et al. (1990), the 'phonologic coherence hypothesis', is not a full account of the visual word identification process, but a detailed description of only the initial stimulus encoding process. In contrast to dual-route theory no explicit representations (i.e., symbols) are assumed in the Van Orden et al. model. The functionally invariant units (nodes) in this model are graphemes and phonemes (subsymbols), however, they are only maintained when they are active, and their properties emerge from system dynamics. It is assumed that in a network connections exist between orthographic subsymbols (graphemes) and phonologic subsymbols (phonemes). After being exposed to a great variety of words, the network learns that covariations exist between orthographic and phonologic subsymbols. A more detailed description of the mechanism that underlies the initial stage of word recognition was presented in the General Introduction (Chapter 1).

As was said before, in Van Orden et al.'s model only the initial stimulus encoding process is described in more detail. The present study, however, necessitates a description of the spelling process. Therefore, we allow ourselves to speculate about the spelling process

along the lines of Van Orden et al.'s subsymbolic approach. In order to be able to spell, as in the reading process connections have to established between phonologic subsymbols and orthographic subsymbols, but in the reverse direction. It has to be noted that the connections established in a network are usually uni-directional, that is, a connection between a subsymbol X and a subsymbol Y does not automatically imply the existence of a connection between subsymbol Y and subsymbol X (but see Frost & Katz, 1989). Establishing this latter connection requires a separate learning process. It seems natural to assume that the same kind of dynamics as apply to the system involved in the reading process also apply to the system that takes care of the spelling process. Thus, for instance, for a subject to spell a word like /ME/ correctly, connections between the phonologic subsymbols [m] and [i:] and the orthographic subsymbols 'M' and 'E' have to established. The phonologic subsymbol [m] is shared consistently across contexts with the orthographic subsymbol 'M', but the phonologic subsymbol [i:] is inconsistently shared across contexts; in /ENTRY/ it is connected to the orthographic subsymbol 'Y', in /KEY/ to 'EY', in /LEAF/ to 'EA', in /CHIEF/ to 'IE', in /BE/ to 'E' and in /BEEF/ to 'EE'. Again, only when the word as a whole contributes to the orthographic coding will the proper spelling be activated. Thus the phonologic subsymbol [m] not only has to be connected to the orthographic subsymbol 'M', but also to orthographic subsymbol 'E', and the phonologic subsymbol [i:] has to be connected to the orthographic subsymbols 'E' and 'M'. This description of establishing connections between phonologic subsymbols and orthographic subsymbols (i.e., the spelling process) mirrors the one between orthographic and phonologic subsymbols (i.e., the reading process) as posited by Van Orden et al. (1990).

The foregoing inevitably leads to the conclusion that the reading process is independent from the spelling process. This is not to suggest that these skills do not correlate - we know they do (Malmquist, 1958, in Frith, 1980). But we believe that this correlation is not due to the two processes being related. Instead, we think that subjects who establish connections easily between orthography and phonology are more likely to do so as well between phonology and orthography. In other words, it may be that a latent competence to establish connections between linguistic (sub)symbols underlies the apparent correlation between reading and spelling skills (see also Mommers, Van Leeuwe, Oud, & Janssens, 1986, and Van Bon & Bouwens, 1990).

Spelling performance usually lags behind reading performance in beginning reading, and even the most experienced language users read far more words properly than they spell these words correctly. According to the phonologic coherence hypothesis, the observed asymmetry between reading and spelling skills has to be attributed to differential learning of reading and spelling. Besides the fact that most people spend far more time on reading than on spelling, it is also the case that while they are spelling they are visually presented with the stimulus and are thus also involved in reading. This implies that in the spelling process not only the matrix of connection weights between phonologic and orthographic subsymbols increases, but also the one between orthographic and phonologic subsymbols. In contrast, when subjects read they are generally not involved in spelling. This suggests that during reading only the matrix of connection weights between orthographic and phonologic subsymbols increases, but not the reversed matrix. Thus, reading may simply be easier than spelling, because the amount of practice spent on reading is much larger than that on spelling.

As from the dual-route model, two hypotheses were derived from the phonologic coherence hypothesis (Van Orden et al., 1990). Again, the first pertains to the reading process. Given the assumption that reading and spelling are essentially independent processes, it was predicted that words of which the spelling is known are not necessarily read faster than words of which the spelling is unknown. The ground for this prediction is that having established the connection between phonology and orthography does not imply that the connection between orthography and phonology will be established as well. As before, the second hypothesis concerns the spelling process. Spelling a word can only occur through the coding of phonemes into graphemes. No second option as posited in the dual-route model is available. Like the reading process the spelling process is phonologically constrained. Therefore, we expect that in spelling the proportion of phonologically correct errors will always surpass the proportion of phonologically incorrect errors, irrespective of the strength of connections. However, the stronger the connections between phonologic subsymbols and orthographic subsymbols the less spelling errors occur. The strength of these connections will be increased through practice.

To summarise: The dual-route theory predicts that subjects read words of which they have orthographic knowledge faster than words of which they have no orthographic knowledge, whereas the phonologic coherence hypothesis predicts that knowing the spelling of a word does not in itself guarantee faster reading. Furthermore, dual-route theory predicts that the distribution of spelling errors is different in words of which subjects have spelling knowledge compared to words of which subjects have no or incomplete spelling knowledge. The phonologic coherence hypothesis, on the other hand, predicts that, irrespective of the subjects' knowledge on the orthography of the word, proportionally more phonologically correct than phonologically incorrect errors will occur. Both models predict that subjects will make more errors on words they have little orthographic knowledge of than on words they have full orthographic knowledge of.

Four experiments were conducted. In Experiments 7.1 and 7.2 both the hypotheses on reading and spelling were tested. Furthermore, a related issue was investigated: Is the spelling training administered in Experiments 7.1 and 7.2 effective for both later spelling and later reading performance? Earlier studies showed that a reading training had a rather limited effect on subsequent spelling performance of subjects (Chapters, 2, 3, and 4 Van Doorn-Van Eijsden, 1984). Studies that investigated whether learning to spell does help readers learn to read not only yielded conflicting results, but also produced results that were difficult to interpret. Uhry and Shepherd (1993) executed a training study in which beginning readers from Grade 1 were instructed in segmentation and spelling skills. Their findings seemed to indicate that the training was beneficial for reading. The training, however, left the subjects ample time to practice reading as well, because the words remained visible during the training. This complicates the interpretation as to what caused the superior reading performance. The same holds for the results of a study by Ehri and Wilce (1987).

Roberts and Ehri (1983), on the other hand, did not find a facilitatory effect of a spelling training on a subsequent naming task. Subjects from the experimental group remembered the words' spellings better than the control subjects, but the reading results showed that both groups' reading accuracy was almost perfect and they were equally fast on a list of experimental pseudowords. This null-finding is again difficult to interpret, because

all subjects received extensive practice that also involved being visually presented with the words.

Thus, to get an unambiguous answer as to the role of spelling training in subsequent spelling and reading it is necessary to disentangle the two processes experimentally, that is, no reading should occur during the spelling training. Therefore, we chose to teach the subjects the spelling of words they had never seen before and we never visually presented them with the stimuli during the training. With beginning literates as subjects, it is relatively easy to create a set of words they are semantically familiar with but have not seen before, and of which they do not have any orthographic knowledge. This allowed us to use words as stimuli when investigating the effect of spelling on reading in beginning literates. The degree of literacy of fluent readers, on the other hand, has reached a level where it becomes virtually impossible to find words they are semantically familiar with, but to which they have not been exposed in print before. Therefore, when they served as subjects we had to resort to pseudowords as stimuli.

In Experiments 7.3 and 7.4 only the first question was tested again, namely whether or not knowledge of a word's orthography speeds up the reading process. Unlike in Experiments 7.1 and 7.2, in Experiments 7.3 and 7.4 the subjects were presented with words they were both semantically and orthographically familiar with. This way it was possible to test for differential effects caused by the absence or presence of prior visual presentation with the experimental stimuli.

The subject groups consisted of both beginning and fluent readers. The studies reported in Chapters 2, 5 and 6 suggest that the reading and spelling behaviour of both groups merely differ quantitatively, not qualitatively. It would be interesting to see whether this result extends to the current set of experiments, using different tasks.

Experiment 7.1

The three main questions of Experiment 7.1 were a) whether beginning readers and spellers (the rather lengthy designation 'readers and spellers' will henceforth be summarised under the terms 'literates') read words they learned the spelling of faster than words they did not learn the spelling of; b) whether the distribution of types of spelling errors is different in words subjects have orthographic knowledge of than in words they do not have orthographic knowledge of; and c) whether a spelling training is beneficial for both spelling and reading. Reading speed was measured through naming and lexical decision[13]. Both tasks require reading the stimulus, but lexical access is not mandatory to carry out the naming task successfully, whereas when subjects perform lexical decisions under the appropriate circumstances (that is, when legal pseudowords constitute the non-word condition; De Groot, 1987) lexical access appears to be guaranteed (but see Balota, 1990). A fourth, more subsidiary issue pertains to differences between skilled and less skilled literates. In Grade 1 differences between these groups emerge soon after the beginning of formal reading and spelling instruction. In the current experiment we will investigate whether literacy level has a differential effect on task conditions.

In Experiment 7.1 beginning literates were instructed in the spelling of words they had not seen in print before. During the spelling training no visual presentation of the word was

provided. Because we were interested in the speed with which subjects read words they know the spelling of as compared to the reading speed for words of which they lack spelling knowledge, in the test stage subjects were presented with the training words as well as with a set of words they were not presented with during the training. Note that a difference in reading time between training and new words might occur simply as a result of priming (Clarke & Morton, 1983; Roediger & Blaxton, 1987). To be able to evaluate the data unambiguously, a third condition was therefore added, which was to serve as baseline: In the training stage a set of words (the 'repetition words') was orally trained by having the subjects repeat the words, but without providing the opportunity of learning the spelling. The repetition condition thus serves as the proper baseline. Whether the spelling training was successful, that is, whether subjects indeed acquired orthographic knowledge of the spelling words, but not of the repetition words, will be checked by means of a spelling-recognition test and a spelling-production test. On each trial in the spelling-recognition test the subjects are presented with both the correct spelling and a plausible alternative and they have to choose the correct spelling from the forms presented. In the spelling-production test the subjects have to produce the spelling themselves. These two different spelling tests were administered, because we found (Chapter 3) that beginning literates often recognise the correct spelling, but become rather hesitant when they have to produce it.

Method

Materials. The word list of Kohnstamm et al. (1981) served as the source for the selection of 18 semantically-familiar, but orthographically-unfamiliar stimuli. Each word contained one obvious spelling problem (the target), because one of the phonemes could in principle be represented by two different graphemes (as in the example of 'speak' vs. 'speek'). Pronouncing the stimulus does not disambiguate the word's spelling. The mean length of the stimuli was 5.5 letters (SD = .71). Furthermore a set of 18 pseudowords (mean length 5.5 letters, SD = .71) was created to serve as stimuli in the lexical-decision task. These pseudowords were derived from the set of experimental words, and were all orthographically legal and pronounceable. The experimental stimuli are presented in Appendix L.

Procedure. The experiment started with a training phase. Subjects had to learn the spelling of six stimuli (henceforth: spelling words), and were to repeat orally after the experimenter six other stimuli (henceforth: repetition words). The remaining six stimuli (henceforth: new words) were not presented during the training. The subjects were instructed in the spelling of six spelling words according to the 'oral-spelling' method: They were named a word and subsequently had to spell the whole word aloud. Appropriate feedback was given by the experimenter. Half of the subjects started the training with the spelling condition followed by the repetition condition, and the other half received the reversed order. Both the spelling and the repetition stimuli were practised four times. Note that a subject was never visually presented with the word.

The training stage was followed by a test stage. All subjects started with a reading task, which involved naming or lexical decision. Half of the subjects participated in the naming task (18 experimental words), which was executed on a Macintosh Classic computer. Subjects were asked to read as quickly and as accurately as possible all stimuli

(training and new words), that were presented on the screen in a random order. Latencies were registered with a voice-key and a millisecond timer, and responses were evaluated by the experimenter. The other half of the subjects performed the lexical-decision task on all 18 experimental words plus the 18 pseudowords. Lexical-decision latencies were registered with two buttons connected to two separate serial ports of a Macintosh Classic and a millisecond timer.

After the naming or lexical-decision task the subjects performed a spelling-recognition test and a spelling-production test on all 18 stimuli. The recognition task was a forced-choice spelling test, which was also presented on the screen of a Macintosh Classic computer. The subject was presented with two spellings of a word, the correct and an incorrect one. The incorrect alternative, in fact a pseudohomophone spelling of the proper word, was always the other plausible spelling of the word (Dutch example: vriend vs. vrient; English example: suit vs. sute). Subjects were asked to indicate with the mouse which of the two they thought was the correct spelling. The spelling-production task was a writing down to dictation test. All 18 stimuli were presented in both spelling tests.

Subjects. From a population of 192 Dutch-speaking children attending Grade 1 of regular primary schools, 32 were selected for participation in the experiment. Half of these subjects were high literates (good readers and spellers) and the other half were low literates (poor readers and spellers), but none had severe reading and spelling problems. Reading and spelling scores were highly, but not perfectly (.77) correlated in this group. The scores on the spelling test (Mommers & Van Dongen, 1986) were used for the assignment to the groups of high and low literates. The group of subjects in the naming (16) and lexical-decision (16) task each contained eight good spellers (29.4; SD = .46 and 29.3; SD = .52, respectively) and eight poor spellers (21.8; SD = 1.9 and 21.8; SD = 1.3, respectively). The difference between good and poor spellers on the reading test (Caesar, 1975) was significant $F(1, 31) = 46.31$, $p < .001$. The group of good spellers (42.4; SD = 10.7) had a higher mean score on the reading test than the poor spellers (21.4; SD = 6.1).

Results

The results of the naming and lexical-decision tasks will be discussed first, followed by those of the spelling-recognition and the spelling-production tasks[14].

Naming results. Naming errors (8.6%), errors due to voice-key failure (12.5%), extremely long (more than 3 SD above the mean, .3%), and extremely short (less than 200 ms; 0%) responses were removed from the data set. A 2 (literacy level: high vs. low) by 3 (stimulus type: spelling word vs. repetition word vs. new word) analysis of variance was performed on the subjects' mean latencies. The main effect of literacy level was the only significant result, $F(1,14) = 15.74$, $p < .01$. High literates (1099 ms) were significantly faster in naming the stimuli than the low literates (2953 ms). Naming times of spelling words (2105 ms) were statistically equal to those of repetition (1927 ms) and new words (2046 ms), $F < 1$. The interaction between literacy level and stimulus type was not significant either ($p > .30$). Figure 7.1a presents the results.

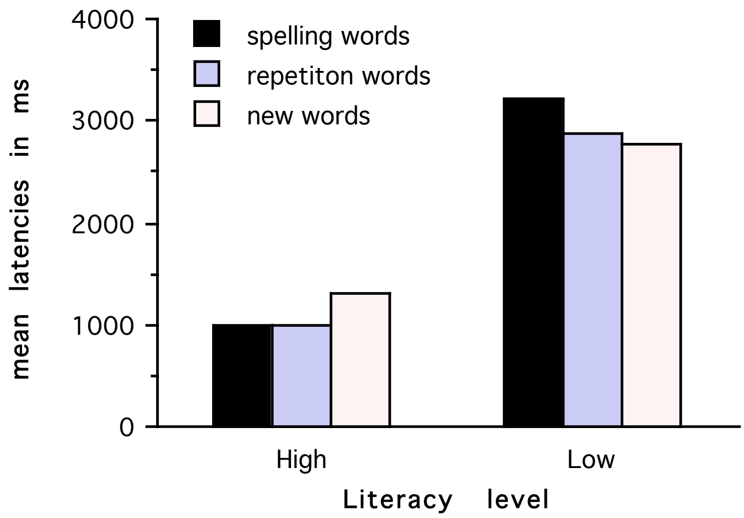

Figure 7.1a. Mean naming latencies of the high and low literates in the naming task of Experiment 7.1.

Lexical-decision results. Errors (7.8%), extremely long (more than 3 SD above the mean, 0%), and extremely short (less than 200 ms; 0%) responses were discarded from the data set before the data were subjected to analysis. A 2 (literacy level: high vs. low) by 3 (stimulus type: spelling word vs. repetition word vs. new word) analysis of variance was performed on the subjects' mean decision latencies. Again only the main effect of literacy level yielded a significant result, $F(1, 14) = 10.44$, $p < .01$, showing that high literates (1610 ms) had faster decision times than low literates (3861 ms). The non-significant main effect of stimulus type indicates that decision times on spelling words (2676 ms), repetition words (2630 ms) and new words (2901 ms) were statistically equal ($F < 1$). The interaction between literacy level and stimulus type was not significant either, $F < 1$. The means are depicted in Figure 7.1b.

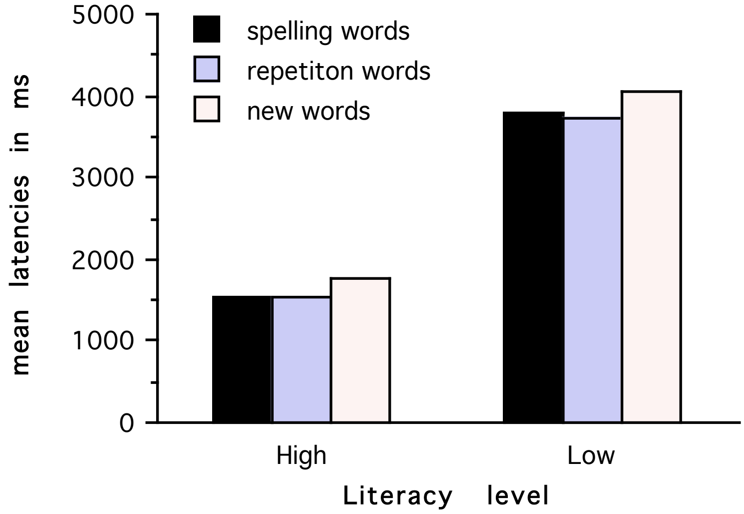

Figure 7.1b. Mean decision latencies of the high and low literates in the lexical-decision task of Experiment 7.1.

Spelling-recognition results. The analysis of the data of the forced choice test was executed on subjects' mean number of errors. A 2 (literacy level: high vs. low) by 2 (task: naming vs. lexical decision) by 3 (stimulus type: spelling word vs. repetition word vs. new word) analysis of variance was performed. Only the main effects of literacy level ($F(1, 28) = 18.58$, $p < .001$) and of stimulus type ($F(2, 56) = 3.32$, $p < .05$) yielded significant results. High literates (.17) made less errors than low literates (.38). The number of errors on the spelling words (.22) was significantly smaller than on the repetition words (.33) and new words (.28), $F(1, 28) = 4.53$, $p < .05$. None of the interactions reached significance, (all *F*'s < 1).

Spelling-production results. The analysis on the data of the dictation test were performed on the mean number of phonologically correct errors and the mean number of phonologically incorrect errors. An error was considered to be phonologically correct when application of grapheme-phoneme correspondence rules yielded the same pronunciation as of the word in question, and incorrect when application of grapheme-phoneme rules gave rise to a pronunciation different from the intended one. For instance, in the incorrectly spelled stimulus 'strant' instead of 'strand' (Dutch for 'beach') is the misspelled final 't' a phonologically correct error, whereas the final 'k' in 'strank' for 'strand' was classified as a phonologically incorrect error.

A 2 (literacy level: high vs. low) by 2 (task: naming vs. lexical decision) by 3 (stimulus type: spelling word vs. repetition word vs. new word) by 2 (error type: phonologic vs. non-phonological) analysis of variance on subjects' mean number of errors was performed. The main effect of task was not significant ($F < 1$). Subjects who participated in the naming condition (.32) produced an equally large number of errors as those who took part in the lexical-decision task (.32). High literates (.18) made significantly less errors than low literates (.47), $F(1, 28) = 45.38$, $p < .001$. The main effect of error type was also significant, $F(1, 28) = 51.41$, $p < .001$. The mean number of phonologically correct errors (.52) was much larger than the mean number of phonologically incorrect errors (.13). The significant main effect of stimulus type ($F(2, 56) = 6.03$, $p < .01$) showed that less errors were made on spelling words (.25) than on repetition (.36) and new (.36) words, Newman-Keuls, $p < .01$. The interaction effect between error type and literacy level ($F(1, 28) = 6.51$, $p < .05$) showed a smaller, though still significant, difference between phonologically correct errors (.31) and phonologically incorrect errors (.05) in the group of high literates than in the group of low literates (.73 and .20, respectively). Both groups made significantly more phonologically correct (high literates: 86%; low literates: 78%) than incorrect errors. Separate analyses were performed on the differences between the mean numbers of phonologically correct and phonologically incorrect errors in each of the three stimulus type conditions. It appeared that in all three conditions the proportion of phonologically correct errors exceeded the proportion of phonologically incorrect errors. In the spelling-word condition it was 82% ($F(1, 28) = 26.80$, $p < .001$), in the repetition-word condition it was 86% ($F(1, 28) = 46.23$, $p < .001$), and in the new-word condition it was 75%, $F(1, 28) = 25.09$, $p < .001$. This holds for both the high and low literates, because the three-way interaction between literacy level, stimulus type and error type did not reach significance, $F < 1$.

Discussion

In the 'Introduction' Section of this chapter two contrasting hypotheses were derived from the dual-route model and the phonologic coherence hypothesis. The first prediction pertained to the nature of the reading process. According to dual-route theory, words of which an orthographic representation is present in the mental lexicon will be read via the fast direct route, whereas words of which such a representation is absent will have to be read through the slower indirect route. The results, however, show that neither in the naming task nor in the lexical-decision task did subjects read spelling words faster than repetition words (and new words), despite their superior spelling knowledge of the spelling words (see the results of the spelling tests).

The prediction from the phonologic coherence hypothesis was that knowing the spelling of a word does not guarantee in itself that these words will be read faster than words of which the subject lacks orthographic knowledge, because spelling is not simply the reverse of reading. The present null-finding, that is, equal naming and lexical-decision latencies of spelling words, repetition words and new words, corroborates this prediction. It suggests that all three stimulus types were read through one and the same route. This has to be the nonlexical route, because for repetition and new words this is the only route available. We will return to the equal reading times for spelling, repetition and new words shortly.

The second hypothesis pertained to the nature of the spelling process. Both the dual-route model and the phonologic coherence hypothesis predict that less spelling errors will be made on words subjects have orthographic knowledge of than on words of which they do not have orthographic knowledge. In both spelling tests the number of spelling errors on repetition and new words surpassed those on spelling words, thus corroborating the shared prediction of the two models. The assumptions of dual-route theory, however, led to the conjecture that the distribution of errors on words subjects have orthographic knowledge of (in this case the spelling words) should be different from that of words subjects do not have orthographical knowledge. If the subject has an orthographic representation it will enable her or him to read out the orthographic image. The a priori chance that during the 'reading out' process a phonologically correct error will occur is not higher than that an orthographically incorrect error will occur. If, on the other hand, a subject lacks orthographic knowledge (as is assumed in the case of the repetition and new words) the proportion of phonologically correct errors will exceed the proportion of phonologically incorrect errors, because the only route available is the phonologic spelling route. Dual-route theory thus predicts a larger proportion of phonologically correct errors in the repetition-word and new-word condition than in the spelling-word condition. Contrary to this prediction, no difference in the distribution of spelling errors emerged between the spelling-word condition, on the one hand, and the repetition-word and new-word conditions, on the other. The spelling-production test revealed that in all three conditions proportionally more phonologically correct than incorrect errors occurred. This is exactly what was predicted by the phonologic coherence hypothesis; subjects use the phonology to spell words. It appears that the spelling process is indeed phonologically constrained, as is the reading process.

A further issue investigated in Experiment 7.1 was whether both spelling and reading performance benefited from the spelling training. The results are quite straightforward. On

the spelling tests, administered after the training, subjects showed superior spelling skills on the spelling words. They did not, however, show superior reading skill on the spelling words. This outcome parallels the finding that a reading training is beneficial for reading, but that it does not provide the beginning literate with the required knowledge to spell words correctly (Chapters 2, 3, and 4; Van Doorn-Van Eijsden, 1984). No transfer whatsoever from spelling to reading seems to have occurred, because the training words (spelling words and repetition words) were read equally fast as the new words, that is, words the subjects were not presented with before. We will return to this issue in the 'General Discussion' of this chapter.

A final topic that was investigated was whether qualitative differences between high and low literates emerged on the spelling and reading tasks. In all cases high literates outperformed low literates. High literates made fewer errors on both spelling tests than low literates, and they were faster on the naming and lexical-decision tasks than their less skilled classmates. Literacy level did not interact with any of the experimental variables, except with error type. It appeared that both high and low literates produced far more phonologically correct than incorrect errors, but the proportion of phonologically correct errors was somewhat larger in the high literate group than in the low literate group. This interaction, however, does not necessarily indicate that good and poor readers differ in the way they spell words, because both groups made proportionally more phonologically correct errors than phonologically incorrect errors. This suggests that both groups used a phonologic strategy to spell words, but that the high literates were more successful in applying the phoneme-to-grapheme correspondence rules (see also Waters, Bruck, & Seidenberg, 1985). Thus, it seems safe to conclude that in general high and low literates read and spell words the same way, indicating that they merely differ quantitatively and not so much qualitatively. This result replicates the findings of Chapter 2.

In short: Experiment 7.1 has provided support for the phonologic coherence hypothesis, and against the dual-route model. Experiment 7.2 will be run to substantiate this conclusion, but unlike in Experiment 1 the subjects will be adults, and hence, presumably, fluent literates.

Experiment 7.2

As in Experiment 7.1 with beginning literates, advanced literates will be trained in the spelling of words. Because of the superior orthographic knowledge of fluent readers and spellers, words could not serve as stimuli. We, therefore had to resort to pseudowords. In creating the pseudowords a new orthographic property was manipulated, namely, spelling ambiguity. Half the pseudowords contained at least one ambiguous phoneme, that is, the correct spelling of these phonemes can only be produced when additional information is given (as is the case for the 'schwa' phoneme). The other half of the pseudowords were unambiguous, that is, correct application of phoneme-grapheme correspondence rules always provides the speller with the correct spelling of the word. Stated differently, to spell ambiguous words correctly word-specific knowledge is required, whereas no such knowledge is necessary to spell unambiguous words.

According to dual-route theory, if orthographic knowledge affects naming speed, an interaction may be expected between spelling ambiguity and type of training (spelling vs.

repetition). No effect from type of training should emerge for the unambiguous words, because no specific spelling knowledge has to be represented. On the other hand, an effect of training should come about when subjects read ambiguous words, because the spelling training of ambiguous words is supposed to cause a fully established orthographic representation, whereas ambiguous words which subjects were only familiarised with in the repetition training should not cause an orthographically complete representation.

Again, according to the phonologic coherence hypothesis, no effect of orthographic knowledge on the naming times of stimuli should show up, nor is it expected that ambiguity has a differential effect on the naming times of the spelling words. Naming times of both ambiguous and unambiguous stimuli, whether subjected to a spelling or to a repetition training, should be equal.

The question whether the spelling process, like the reading process, is phonologically mediated or not was tested again in Experiment 7.2, as well as the question whether reading performance and spelling performance would both benefit from a spelling training.

Method

Materials. Forty pseudowords were created to serve in the experiment. Half of the pseudowords were regular, that is, they can be spelled by application of phoneme-grapheme conversion rules. These stimuli will henceforth be called 'unambiguous'. The other half all contained at least one ambiguous phoneme. Only by supplying additional information is it possible for a subject to correctly spell the created pseudoword according to the spelling the experimenters had decided on. These pseudowords will be called 'ambiguous'. Thirty pseudowords were experimental stimuli and ten served as fillers. The mean length of both the ambiguous and unambiguous pseudowords was 7.7 letters (*SD* = .61; range 7 to 9 letters). The experimental stimuli are presented in Appendix M.

Procedure. The experiment consisted of a training and a test stage. In the training stage subjects had to learn the spelling of ten pseudowords (five ambiguous and five unambiguous). As in Experiment 7.1 subjects were instructed according to the oral-spelling method. Recall that no subject was ever presented visually with the stimuli. Furthermore they were asked to repeat ten pseudowords of which again five were ambiguous and five unambiguous. To make sure that enough attention was paid to the repetition stimuli, they were also asked after each repeat either how many syllables the pseudoword they just named had, or what the first and last letter was. Questions about the pseudowords were randomised to prevent anticipation. We decided on these two elaboration procedures (first and last letter question and number of syllable question) because we thought they would not lead to learning the spelling of these pseudowords. Whether or not we succeeded will be shown by the spelling test. No knowledge of the spelling should be apparent for the repetition words. Both the spelling and the repetition words were practised three times. Ten pseudowords were not practised and served as new words in the test stage. By means of proper counterbalancing all thirty pseudowords appeared in all conditions, thus avoiding a confounding between condition and stimulus set.

After the training subjects took part in a naming task, and they were tested on their spelling knowledge with a spelling-recognition and a spelling-production test. The naming task was performed on all trained stimuli (ten spelling and ten repetition pseudowords), on

the set of ten new pseudowords and on ten fillers (also new). The naming task was executed on a Macintosh LCII computer. The procedure of latency registration was identical to that of Experiment 7.1. The procedures of the spelling-recognition and spelling-production tests were also identical to those of Experiment 7.1 (details in the 'Procedure' Section of Experiment 7.1).

Subjects. Twenty Dutch-speaking undergraduates from the Psychology Department of the University of Amsterdam participated to fulfil course requirements.

Results

As in Experiment 7.1, the results of the naming task will be discussed first, followed by those of the spelling-recognition and spelling-production tests.

Naming results. Before latencies were subjected to analysis, naming errors (6.8%), responses based on voice-key failure (2.4%), extremely long (more than 3 SD; .9%) above the mean), and extremely short latencies (less than 200 ms; 0%) were discarded from the data set. A 3 (stimulus type: spelling vs. repetition vs. new) by 2 (spelling ambiguity: ambiguous vs. unambiguous) analysis of variance was executed on the subjects' means. Both main effects were significant, but the interaction was not. The main effect of stimulus type ($F(2, 38) = 12.58$, $p < .001$) indicated that naming times of spelling pseudowords (580 ms) and repetition pseudowords (592 ms) did not differ significantly from each other, but both these types of stimuli were named faster than new pseudowords (626 ms; Newman-Keuls, $p < .01$). The main effect of spelling ambiguity revealed that ambiguous pseudowords (631 ms) were named slower than unambiguous pseudowords (568 ms), $F(1, 19) = 48.41$, $p < .001$. The results are presented in Figure 7.2.

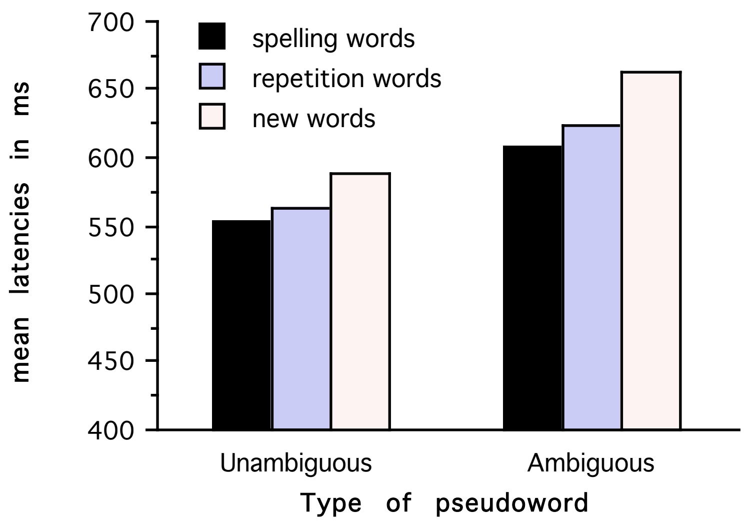

Figure 7.2. Mean naming latencies of the fluent literates in the naming task of Experiment 7.2.

An additional analysis was executed to check whether a dissociation emerged between subjects who learned the orthography of the spelling pseudowords and those who were less successful. Spelling performance was assessed by computing the mean number of errors in

the spelling condition on the spelling-production test. A group of successful (12) and a group of unsuccessful (8) spellers could be formed with means of .08 and .53 respectively, $F(1, 19) = 62.04$, $p < .001$. The difference between these groups on the naming latencies was tested by means of a 2 (spelling performance: successful vs. unsuccessful) by 2 (stimulus type: spelling vs. repetition) analysis of variance. The main effects of spelling performance ($F < 1$) and the main effect of stimulus type ($p > .20$) were not significant, and neither was the interaction effect between spelling performance and stimulus type, $F < 1$.

Spelling-recognition test. The analysis of the data of the forced choice test was executed on the spelling and repetition pseudowords only. Asking which of a pair of new pseudowords is the correct one is obviously an absurd question. A 2 (stimulus type: spelling pseudowords vs. repetition pseudowords) by 2 (spelling ambiguity: ambiguous vs. unambiguous) analysis of variance was performed on the subjects' mean error score. The main effects of stimulus type ($F(1, 19) = 17.77$, $p < .001$) and ambiguity ($F(1, 19) = 37.64$, $p < .001$) were significant. The effect of stimulus type showed that less errors were made in the spelling (.07) than in the repetition (.21) condition, and the effect of spelling ambiguity revealed that more errors were made on ambiguous pseudowords (.23) than on unambiguous pseudowords (.05). The interaction between stimulus type and spelling ambiguity was also significant, $F(1, 19) = 18.70$, $p < .001$. A post-hoc analysis revealed that the number of errors on the unambiguous pseudowords in the spelling (.04) and repetition (.06) conditions were the same, whereas the number of errors on ambiguous pseudowords in the repetition condition (.37) was significantly higher than in the spelling condition (.10; Newman-Keuls, $p < .01$).

Spelling-production test. As in 7.1, the mean numbers of phonologically correct and phonologically incorrect errors were calculated. A 3 (stimulus type: spelling pseudowords vs. repetition pseudowords vs. new pseudowords) by 2 (error type: phonologic vs. non-phonological) analysis of variance was performed on the subjects' mean error scores. Both the main effects of stimulus type ($F(2, 38) = 60.28$, $p < .001$) and of error type ($F(1, 19) = 158.29$, $p < .001$) were significant. The main effect of stimulus type showed that the mean number of errors in the spelling condition (.07) was significantly less than those in the repetition (.27) and new (.31) conditions (Newman-Keuls, $p < .01$). No significant difference emerged between the repetition and new conditions. The effect of error type revealed that subjects made more phonologically correct (.35) than incorrect errors (.08). Separate analyses were performed on the differences between the mean numbers of phonologically correct and phonologically incorrect errors in each of the three stimulus type conditions. It appeared that in all three conditions the proportion of phonologically correct errors exceeded the proportion of phonologically incorrect errors. In the spelling-word condition it was 71% ($F(1, 19) = 10.69$, $p < .01$), in the repetition-word condition it was 87% ($F(1, 19) = 94.19$, $p < .001$), and in the new-word condition it was 78% ($F(1, 19) = 53.51$, $p < .001$).

Discussion

We will begin our discussion with the naming data. To recall, dual-route predicts that stimuli of which subjects have spelling knowledge will be named faster than stimuli of which such knowledge is absent. In contrast, the phonologic coherence hypothesis predicts that no difference in naming times should occur between the two types of stimuli.

Furthermore, dual-route theory predicts an interaction (see the introduction of the 'Method Section' of Experiment 7.2) between stimulus type and ambiguity. That is, stimuli containing an ambiguous spelling cluster, but not those containing an unambiguous spelling cluster, should be named faster when preceded by a spelling training than when preceded by a repetition training, (recall that the ambiguity variable pertains to the spelling process and not to the reading process). The phonologic coherence hypothesis does not predict such an interaction.

The phonologic coherence hypothesis seems again to be in the right. No interaction between stimulus type and ambiguity emerged, and no difference in naming times between spelling stimuli and repetition stimuli was apparent. Yet, it is clear from the results of the spelling tests that subjects had superior spelling knowledge on the spelling stimuli. The additional analysis showed that both the subjects who had actually learned the orthography of the spelling stimuli and those who were less successful named the spelling stimuli equally as fast as the repetition stimuli. Note that the difference between the subjects who completed the spelling training successfully and those who were less successful was rather large. The latter group produced more than 5 times as many errors as the former. The fact that even the good spellers showed no effect of stimulus type (spelling stimuli vs. repetition stimuli) is another indication that spelling knowledge does not necessarily contribute to the naming process. This finding thus also counters the prediction derived from the dual-route model and corroborates the phonologic coherence hypothesis.

One finding seems to discredit the phonologic coherence hypothesis. It was predicted that naming times of ambiguous stimuli measures up to naming times of unambiguous stimuli. However, ambiguous stimuli were named slower than unambiguous stimuli. A control experiment[15] was executed to investigate the possibility that the ambiguous pseudowords were inadvertently more complex than the unambiguous pseudowords. Rather than spelling ambiguity per se a difference between stimulus sets in terms of orthographic complexity could have caused the present effect on naming latencies. The data of that control study suggested that the longer latencies of the ambiguous words were indeed caused by their larger orthographic complexity. The spelling ambiguity factor turned out to be confounded with orthographic complexity.

The results of the spelling tests revealed that the common prediction of the dual-route model and the phonologic coherence hypothesis was correct: Less spelling errors occurred on stimuli that were subjected to a spelling training than on stimuli that were subjected to a repetition training or which were not trained at all (new stimuli). The results of the analysis that included the factor error type, however, distinguished between the predictions made by the two models under investigation. As discussed before, according to the dual-route model, stimuli of which subjects know the spelling should a different distribution of phonologically correct and phonologically incorrect errors than of words they do not have orthographical knowledge of. Contrary to this prediction, the pattern of errors in the relevant condition (the spelling condition) was identical to that in the remaining conditions (the repetition and new condition): In all three conditions, and in agreement with the prediction of the phonologic coherence hypothesis, proportionally more phonologically correct than phonologically incorrect errors occurred. If spelling is phonologically constrained, as the phonologic coherence hypothesis asserts, the distribution of errors should indeed not differ between stimuli subjects have orthographic knowledge of and those they have no such knowledge of.

The results of this experiment also show that the spelling training was beneficial for subsequent spelling performance, but not for reading performance. This finding parallels the outcome of Experiment 7.1. However, a difference between the experiments is that, unlike the beginning literates of Experiment 7.1, the adult subjects of Experiment 7.2 named the training stimuli (spelling and repetition stimuli) faster than the stimuli that were not presented during training (new stimuli). Some transfer from training to test seems to have occurred. Note that this transfer is unrelated to acquired spelling knowledge, because the reading of the repetition stimuli profited as much from the prior training, which did not involve reading, as the reading of the spelling stimuli.

In short: results of the spelling tests of Experiment 7.2 also provided evidence for the validity of the phonologic coherence hypothesis, and seem to falsify the dual-route model. In Experiments 7.3 and 7.4 we focused on the question whether orthographic knowledge contributes to the reading process. In addition, we investigated the issue of differences in performance between good and poor literates.

Experiment 7.3

Method

It was tested whether beginning literates are faster reading words they have come across in their reading curriculum and of which they know the spelling, than words they also have seen in their curriculum, but do not know the correct spelling of. To control for a correlation between orthographic knowledge and frequency, word frequency was taken into account.

Materials. The books of the subjects' reading curriculum (revised edition of "Veilig leren lezen", Caesar, 1991) served as the source for the stimuli used in Experiment 7.3. Stimuli (N = 50) were chosen such that they caused one potential spelling problem for beginners. Over half of the words (29) contained a so-called ambiguous phoneme, that is, sounding out or hearing the phoneme does not disambiguate the word's spelling [spi:k] could either be 'speek or 'speak'). Consequently, the proper grapheme has to be memorised. The remaining 21 contained various other types of spelling problems, one of which was letter order. The correct order of two letters in a diphthong can be confusing (for instance, the Dutch [ʊ:] is 'oe' and not 'eo'). Almost half of the words were low-frequent (24) and the remaining 26 were high-frequent, according to a frequency count of the reading books. The stimuli are presented in Appendix N.

Procedure. Subjects were tested on their spelling knowledge of all 50 experimental stimuli by administering three dictation tests over a period of three weeks. The experimenter read a sentence containing one of the experimental (target) words. The subjects had to write down the target word. Because every word was tested three times the spelling knowledge of the children could be assessed rather reliably. The spelling of a word was considered to be known when subjects had at least two out of three spellings correct, and unknown when none or one of the attempts was successful. No feedback on the subjects' performance was supplied, since they were not supposed to consciously learn the spellings.

At the end of the third week, three days after the last dictation test, the subjects were required to name all 50 words on visual presentation. The same procedure as used in Experiment 7.1 was applied here (see 'Procedure' Section of Experiment 7.1). We chose to administer the dictation tests before the naming task because subjects find the dictation tests rather tedious, and this way we could present the computer part of the experiment as a reward for their hard labour.

Subjects. All subjects were first-grade readers and spellers from one classroom (30 children). The data of 25 of them were used in the analysis. One child did not take part in the experiment because her reading and spelling skills were too limited. The data of four other children were excluded from the data set because they were absent during one of the spelling tests or the naming task. As in 7.1, the subjects also received reading instruction according to the curriculum "Veilig leren lezen", but they used the revised edition. The initial teaching words of this edition differ from the ones the subjects in Experiment 7.1 received, but the essentials of the method remained unchanged.

As in Experiment 7.1 reading and spelling scores were fairly highly, but not perfectly, correlated (.58) in this group. The scores on the spelling test (Mommers & Van Dongen, 1986) were used for the assignment to the groups of high literates and low literates. The good spellers had a mean score of 27.8 (SD = 1.0) and the poor spellers had a mean score of 23.2 (SD = 3.5). The difference between good and poor spellers on the reading test (Caesar, 1975) was significant, $F(1, 21) = 6.75$, $p < .05$. The good spellers (26.0; SD = 5.8) had a higher score on the reading test than the poor spellers (20.2; SD = 4.8).

Results

Dictation tests. A 2 (literacy level: high vs. low) by 2 (frequency: high vs. low) analysis of variance on the subjects' mean number of correct spellings on the three dictation tests showed that high literates (1.90) spelled better than low literates (1.60), $F(1, 23) = 4.34$, $p < .05$. High-frequent (2.26) words were spelled better than low-frequent words (1.24), $F(1, 23) = 317.88$, $p < .001$. The interaction effect between literacy level and frequency did not reach a significant level, $F < 1$. Item error means from the groups of good and poor readers correlated .94, indicating that the words posed no differential difficulties for these groups.

Naming task. Subjects' mean latencies were calculated for high-frequent words of which they knew the spelling, high-frequent words of which they did not know the spelling, and the same conditions for the low-frequent words. Before the means were subjected to analysis naming errors (6.9%), responses based on voice-key failure (8.3%), responses based on stimuli that were missed during one of the spelling tests (3.3%), extremely long (more than 3 SD above the mean; 1.2%), and extremely short latencies (less than 200 ms; 0%) were discarded from the data set.

A 2 (literacy level: high vs. low) by 2 (spelling knowledge: known vs. unknown) by 2 (frequency: high vs. low) analysis of variance indicated that the high literates (1042 ms) were faster in naming the stimuli than the low literates (1672 ms), $F(1, 23) = 4.13$, $p = .05$. The main effect of frequency was significant, revealing that high-frequent words (1231 ms) were named faster than low-frequent words (1483 ms), $F(1, 23) = 23.25$, $p < .001$. The main effect of spelling knowledge was marginally significant, $F(1, 23) = 3.91$, $.05 < p < .10$. This borderline effect suggests that known words (1295 ms) were read faster than unknown

words (1419 ms). The marginally significant interaction between literacy level and spelling knowledge, however, qualifies this result, $F(1, 23) = 3.02$, $.05 < p < .10$. Only the group of low literates showed a marginally significant difference between known (1555 ms) and unknown (1789 ms) stimuli, $F(1, 23) = 3.32$, $.05 < p < .10$, whereas no such difference is apparent in the group of high literates ($F < 1$; 1035 and 1050 ms respectively). The effect is presented in Figure 7.3.

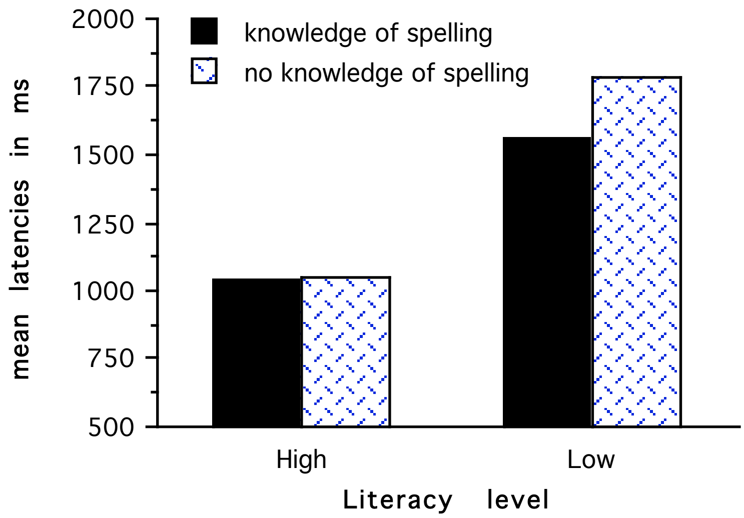

Figure 7.3. Mean naming latencies of the high and low literates in the naming task of Experiment 7.3.

The interaction effect between literacy level and frequency was also marginally significant, $F(1, 23) = 3.87$, $.05 < p < .10$. High literates did not name high-frequent (968 ms) stimuli significantly faster than low-frequent stimuli (1117 ms), whereas a significant difference ($F(1, 23) = 11.08$, $p < .01$) occurred between high-frequent and low-frequent words in the group of low literates (1494 and 1850 ms respectively). No further interaction effects reached significant levels.

Discussion

A number of findings are in line with results found earlier. One of them is that high literates, as assessed with the spelling test (Mommers & Van Dongen, 1986), had a higher mean score on the dictation tests than the group of low literates, and that high literates were faster in the naming task than the low literates (see Experiment 2.1). The outcome that high frequency words were named faster than low frequency words is in agreement with an abundance of studies on word recognition. Furthermore, the fact that the same experimental words were difficult for both high and low literates is again an indication that the spelling skills of the subject groups merely differ quantitatively.

The first indication in this study that readers might read words they know the spelling of faster than words of which they have no, incomplete or even incorrect representations, is provided by a group of subjects of which it was least expected namely, the low literates. In the framework of the dual-route model this finding would suggest that the low literates have developed a direct route to the lexicon, whereas the high literates have not. This conclusion,

however, would be at odds with most of the literature on beginning reading, including that on dual-route theory (Chapter 2; Gibson & Levin, 1975; Reitsma, 1983, 1984; Rozin & Gleitman, 1977). Most of the researchers on beginning reading who adhere to the dual-route model agree that initial reading is characterised by a heavy reliance on phonologic mediation. Only after considerable experience will beginning readers be able to make use of the direct route. Good readers are supposed to develop this direct route more quickly, not more slowly, than poor readers.

However, Rayner and Pollatsek (1989) disagree with the idea that beginning readers show a shift from initially accessing the lexicon indirectly to using both indirect and direct routes. They posit an opposite developmental shift from the exclusive use of the direct access route to the use of both the direct and indirect routes. But even if Rayner and Pollatsek are right, it would be impossible to reconcile the present data with dual-route theory, because also the high literates should have shown some effect of spelling knowledge on naming time. The reason is that during the course of reading development the direct route is not replaced by the indirect route. Instead, the indirect route is added to the direct as a second process through which lexical access can be effectuated, in parallel with the direct route. In sum, the finding that no effect of orthographic knowledge whatsoever occurred in high literates seems to invalidate an interpretation of this effect in low literates in terms of dual-route theory. In contrast, at least the data of the high literates are in agreement with the phonologic coherence hypothesis and converge with the majority of the remaining results of this study. Therefore we are inclined to believe that not the high literates, but the low literates showed the deviant result. To substantiate our view that the high literates showed the expected pattern of results a final experiment, conceptually similar to Experiment 7.3, was performed with adult literates as subjects.

Experiment 7.4

Method

Materials and Procedure. The words (128 in all) of a two-choice spelling test for adult readers and spellers (Langereis & Elshout, 1985) constituted the materials used in this experiment. Subjects started with a naming task. They were asked to name the words from the spelling test in their correct spelling as quickly and as accurately as possible. Latencies were registered with a voice-key and a millisecond timer, and responses were evaluated by the experimenter. After completion of the naming task the subjects performed a distraction task (repeating and naming pseudowords). Finally, the two-choice spelling test was administered by means of a computer. The correct spelling and a plausible alternative (identical to the one in the original test) were presented on the screen of a Macintosh Plus computer. Subjects had to decide which of the two spellings was the correct one (a comparable example in English would be: ECSTASY / ECSTACY), and also whether they had confidence in their choice. The mouse was used to indicate choice and confidence on a so-called click and touch area. Responses were registered by a computer program. The reliability of the spelling test had been assessed before and appeared to be satisfactory, r = .81 (N = 387). The two-choice spelling test is presented in Appendix O.

Subjects. In this study 16 Dutch-speaking undergraduates from the Department of Psychology of the University of Amsterdam participated to fulfil course requirements.

Results

Before the analysis was performed, latencies that were based on naming errors (.7%), errors due to voice-key failure (3.4%), extremely fast responses (less than 200 ms, .3%), and extremely slow responses (more than 3 SD above the mean, 1.5%) were removed from the data set.

For every subject the mean naming latencies of the following four word groups were calculated: words they spelled correctly (that is, which they chose the correct spelling of) and were confident about (correct and certain); words they spelled incorrectly and were confident about (incorrect and certain); words they spelled correctly and were not confident about (correct and uncertain); and finally, words they spelled incorrectly and were not confident about (incorrect and uncertain). A 2 (spelling knowledge: correct vs. incorrect) by 2 (confidence: certain vs. uncertain) analysis of variance was executed on the means. Neither the main effects (effect of spelling knowledge: $p > .15$; effect of confidence; $F < 1$) nor the interaction between them ($p > .15$) reached significant levels. Mean naming latencies were 493 ms ($SD = 64$) in the 'correct and certain' condition, 494 ms ($SD = 74$) in the 'incorrect and certain' condition, 501 ms ($SD = 69$) in the 'correct and uncertain' condition, and finally 487 ms ($SD = 78$) in the 'incorrect and uncertain' condition. The results are depicted in Figure 7.4. One additional analysis was executed. The mean naming times on correct and certain stimuli (493 ms, $SD = 64$) were compared with those of the collapsed conditions of 'incorrect and certain', 'correct and uncertain', and incorrect and uncertain' stimuli (498 ms, $SD = 69$). Again, no statistical difference occurred between the types of stimuli ($p > .25$).

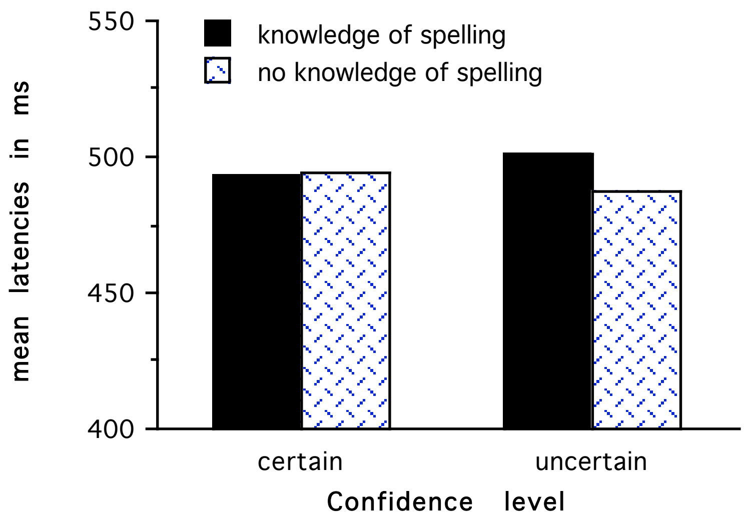

Figure 7.4. Mean naming latencies of the fluent literates in the naming task of Experiment 7.4.

Discussion

The results are quite straightforward. Fluent readers do not name words of which their orthographic knowledge is correct any faster than words of which they have incorrect knowledge. They named words they thought they knew the spelling of (certain condition) equally as fast as words of which they stated not to be certain about the correct spelling (uncertain condition). Furthermore, naming times of words for which subjects chose the correct spelling were not influenced by whether or not they had confidence in their choice, nor were naming times of wrong choices (incorrect condition) affected by the confidence variable. In other words, even the existence of incorrect orthographic representations, suggested by the incorrect-certain condition, did not slow down the subjects' word recognition. Thus, the prediction of dual-route theory that words of which subjects have an orthographic representation will be read through the fast direct route and that words without such representation will be read through the slower indirect route did not prove correct. What is more, the results again support the phonologic coherence hypothesis. Furthermore, the present finding indeed support the interpretation (see the 'Discussion" Section of Experiment 7.3) that the low literates of Experiment 7.3, not the high literates, showed the unexpected result.

General Discussion

Dual-route theory predicts that words of which an orthographic representations is present in the mental lexicon will be read faster than words of which such a representation is absent, incomplete or wrong, because the former can be read by the fast direct route, whereas the latter have to be read through the slower indirect route. The phonologic coherence hypothesis, on the other hand, does not predict faster reading processes per se for words subjects have an orthographic representation of, because having established the connections between phonologic subsymbols and orthographic subsymbols (i.e., the spelling process) does not imply that the connections between orthographic subsymbols and phonologic subsymbols (i.e., the reading process) have been formed as well. The results of all four experiments repudiate the dual-route model and corroborate the phonologic coherence hypothesis, because beginning and advanced literates do not show a naming time difference between words of which they have correct orthographic knowledge and words of which they have limited or incorrect orthographic knowledge. However, one group of subjects deviates from this general pattern, namely, the low literates in Experiment 7.3. It remains unclear why specifically these beginning literates read high-frequent words they know how to spell faster than high-frequent words they do not know the spelling of.

In Experiments 7.1 and 7.2 subjects were to read words they had never seen before, neither prior to the experiment, nor during the training session of the experiment. One could argue that an orthographic representation to be used in reading also involve a visual aspect, as indeed is suggested by the term 'orthographic image' (Ehri , 1980). That Experiments 7.1 and 7.2 do not show an effect of the orthographic representation formed during training on subsequent naming of the corresponding words might then be held due to the fact that the developed representations are simply not appropriate for reading. The concept orthographic image, however, does <u>not</u> refer to a memorised geometric figure (Ehri, 1980), but is a sequence of letters bearing a systematic relationship to phonologic properties of the word.

What this notion of the orthographic representation entails, is exactly what has been established during the training in Experiments 7.1 and 7.2, as was shown by the results on the spelling tests.

Another possible objection to the procedure used in the Experiments 7.1 and 7.2 is that the number of presentations (four in Experiment 7.1 with beginning literates, and three in Experiment 7.2 with fluent literates) with a word during training was too limited to really secure a useful orthographic representation. This, however, is countered by the findings of Experiment 7.3 and 7.4, in which the presence of such orthographic representations has to be assumed. The results of these experiments converge with those of Experiments 7.1 and 7.2, so clearly the most parsimonious interpretation of Experiments 7.1 through 7.4 combined is that orthographic representations, in the sense used in dual-route theory, are not employed during reading. Furthermore, earlier studies showed that a rather limited number of presentations of a word (one or two) is enough to familiarise a beginning reader with the orthography of the word (Chapter 2; Reitsma, 1983).

The second major topic of this study was the nature of the spelling process. Dual-route theory predicts that if an orthographic representation is present in the mental lexicon the reader simply reads out the spelling from memory (the direct spelling route). If true, it may be expected that the proportion of phonologically correct errors does not exceed the proportion of phonologically incorrect errors in case a subject knows the spelling of a word, because there is no apparent reason why the reading out process would cause the one to be higher than the other. Dual-route theory also predicts that if no orthographic representation is present in the lexicon a subject has to recode the phonemes into graphemes to produce the orthography of the word (the indirect spelling route). In this case a higher proportion of phonologically correct than phonologically incorrect errors should occur. In contrast, the phonologic coherence hypothesis assumes that there is only one way to compute the orthography of a word, namely, from its phonology, and therefore predicts that in all cases the proportion of phonologically correct errors exceeds that of phonologically incorrect errors. Again the effects seemed to favour the phonologic coherence hypothesis, and proved the dual-route model to be wrong.

The overriding importance of the role of phonology in spelling seems more accepted than its role in reading (Frith; 1979; Hoffman & Norris, 1989; but see also Kreiner, 1992; Kreiner & Gough, 1990). Studies with dyslexics (Bruck, 1988; Hatfield & Patterson, 1983), with normal good and poor spellers (Frith, 1980; Holligan & Johnston, 1991; Perin, 1982), and with adults (Fischer, Schankweiler, & Liberman, 1985) all show that the majority of spelling errors are phonologically plausible. It is, however, the case that the difference between the number of phonologically correct and phonologically incorrect errors is smaller in subjects with specific reading and spelling difficulties compared to normals (Newman, Fields, & Wright, 1993), in poor beginning spellers as compared to good beginning spellers (Waters, Bruck, & Seidenberg, 1985), and in poor advanced spellers as compared to good advanced spellers (Holmes & Ng, 1993).

Another issue of this study was whether a spelling training is beneficial for reading. The answer to this question appears to be unequivocally negative. On the naming task, beginning literates in Experiment 7.1 and advanced literates in Experiment 7.2 did not seem to profit from the newly acquired spelling knowledge. The fact that subjects were not involved in reading during the spelling training made it possible to assess purely the

influence of a spelling training on reading. According to 'transfer-appropriate-processing' theory (TAP-theory, originated by Morris, Bransford and Franks, 1977), it depends on the degree of overlap between processing operations during study (training) and test, whether spelling is beneficial for reading, and for that same matter reading beneficial for spelling (see for excellent summaries: Roediger, 1990, Roediger, Weldon, & Challis, 1989). The larger this overlap the better the performance in the test stage will be. Thus, a large overlap between spelling processes during training and reading processes in the test stage, and between reading processes during training and spelling processes in the test stage should produce beneficial effects. A limited overlap, on the other hand, should not cause transfer from spelling to reading or from reading to spelling. It has been shown before that a reading training is not a very effective way to learn the spelling of a word, whereas it is very effective to learn to read (Chapter 2). Here we showed that a spelling training is not a very effective way to learn to read, but it is for learning to spell. If the assumption posited by TAP-theory is correct, then it shows from a different perspective what is also shown by this study, namely, that limited or no overlap exists between reading and spelling.

A final result we like to comment on briefly is the failure to find substantial qualitative differences amongst beginning literates, and between beginning and fluent literates. Given the overall similarities across all subject groups on both the reading and the spelling tasks, it is unnecessary to assume a developmental shift (Rayner & Pollatsek, 1989). Obviously, fluent literates were faster on the reading tasks than the beginners and they made less spelling errors on the spelling tasks. However, both groups exhibited heavy reliance on phonology when spelling. The absence of a qualitative difference between literacy levels is in accordance with the results of two earlier studies that allowed a comparison between beginning and fluent readers on several reading tasks (Chapters 5 and 6).

8

GENERAL DISCUSSION

"..and always, for most readers, we can say ...that "to read is, in effect, to translate writing into speech." And while this inner speech is but an abbreviated and reduced form of the speech of everyday life, a shadow copy as it were, it nevertheless retains the essential characteristics of the original." (E. B. Huey, 1908, p. 123).

The purpose of this chapter is to demonstrate that the findings of the studies reported in this dissertation can be accounted for with a single-route model as developed by Van Orden and his co-workers. As was said in the General Introduction, when I started my research on beginning reading and spelling not the model of Van Orden, but the dual-route model was taken as its theoretical starting point.

Orthographic images and the phonologic coherence hypothesis

A study by Reitsma and Vinke (1986) suggested that Dutch-speaking beginning readers of Grade 1 can bypass the nonlexical reading route after a rather limited number of presentations of a word. They found that beginning readers, after having read words zero or three times, show equal naming times for words and their corresponding pseudohomophones, whereas after nine presentations of a word its naming time appeared to be faster than that of its yoked pseudohomophone. This result suggested, in terms of dual-route theory, that word-specific knowledge (an orthographic representation) had been established between three and nine presentations of a word, and that from then on beginning readers use the direct reading route. In Chapter 2, I tried to determine more precisely the number of visual presentations of a word that would secure an orthographic representation in the mental lexicon, enabling beginning readers to circumvent the phonologic route. The same paradigm was used as applied by Reitsma and Vinke (1986). However, with a group of readers similar to that of Reitsma and Vinke, I failed to find the interaction they report. Thus, the question after how many presentations with a word a beginning reader shifts from an indirect to a direct way of reading, could not be answered. It was clear from the data, however, that this group of beginning readers did develop an orthographic representation useful in the reading process, but its nature remained unknown. The failure to find the expected interaction was explained in terms of transfer: When readers are presented with

words they will gain knowledge of their orthography, as is shown from the finding that naming times of words decrease as a function of the number of presentations. The orthography of pseudohomophones, however, resembles that of their corresponding words to a greater or lesser extent. Thus, it was hypothesised that the orthographic knowledge acquired during the reading of words was used in the reading of pseudohomophones.

The transfer explanation seems to get support from two studies by Reitsma (1988, 1989). He showed that a reading training on a set of neighbour words (for instance, 'vier', 'hier', 'bier') with subjects similar to those participating in the study of Chapter 2 resulted in shorter naming times of a neighbour word not presented in the prior reading training (for instance, 'dier'). Theloosen and Van Bon (1993), and Van Daal, Bakker, and Reitsma (1986), did not find a transfer effect in a similar experiment. Their subjects, however, were extremely poor readers, and probably have to be classified as developmental dyslexics. It is plausible that dyslexics need far more presentations of a word to enable them to profit from the orthographic structure of the word (see also Reitsma & Vinke, 1986).

Before describing the findings of the remainder of the experiments of this dissertation, it should be noted that I do not have a satisfactory explanation for the diverging results of Chapter 2 and those of Reitsma and Vinke (1986). Furthermore, the absence of an interaction between stimulus type (words vs. pseudohomophones) and presentation frequency (high vs. low) in Chapter 2 remains puzzling, because not only the dual-route model, but also the single-route model of Van Orden et al. (1990) would predict that this interaction occurs at some point in the development of reading skill, albeit for different reasons. In the General Introduction I explained how the phonologic coherence hypothesis accounts for the frequency by regularity interaction as established by Seidenberg et al. (1984); the same logic applies here. Reading a word involves establishing coherence between orthographic and phonologic subsymbols. The more often a reader is presented with a word the faster will resonance occur, because the connections of these words are updated (their strength is increased) relatively often (this mechanism provides a simple account of the frequency effect). No reading time differences are expected between words and the corresponding pseudohomophones when readers have only limited experience with the words (low-frequent condition), but after a certain number of word presentations resonance of the high-frequent words should be faster than that of the corresponding pseudohomophone.

In terms of Van Orden et al.'s single-route model, the initial question phrased in terms of dual-route theory - how many presentations of a word are necessary to establish an orthographic image? - would run as follows: How many presentations of a word are needed to establish fast resonance? The answer to this question presumably depends on several factors. First, the regularity of the writing system: the more regular the system the faster the emergence of coherence. Secondly, the amount of experience the reader has with written language. Thirdly, the type of words the reader has been presented with. If, for instance, a reader is familiar with a set of neighbour words ('hill, 'mill', and 'till'), resonance will probably be faster when a new neighbour word ('still') is presented than when a new non-neighbour word (e.g., 'attractor') has to be read. Deze zin hoort hier niet thuis. Furthermore, not only the relation between orthography and phonology, but also the one between phonology and meaning, and orthography and semantics, will determine speed of resonance. Given that the precise relation between each of these factors and the speed of resonance is

(still) unknown, it seems that the question how many presentations of a word are needed to establish fast resonance was slightly premature, and has to remain unanswered for the time being. I will now move on to discussing the remaining experiments of my research.

Response competition underlies the first-letter effect

As was explained in the General Introduction, in Chapter 5 another attempt was made to investigate the shift from indirect to direct reading. There, an alternative paradigm was used, namely, the first-letter-naming task. Recall that Rossmeissl and Theios (1982) found that fluent readers named the first letter of words and legal nonwords faster than of illegal, unpronounceable nonwords. They reasoned that fluent readers made use of an orthographically legal context to determine the first letter of a letter string, and took the first-letter effect as another indication for direct access. I reasoned that, given the correctness of their interpretation beginning readers who still read via the phonologic route should not show the effect. However, both beginning (poor, medium and good readers) and fluent readers showed it. Moreover, the first-letter effect in beginning readers was established with both highly familiar words and fairly unfamiliar ones. Instead of accepting the rather implausible conclusion that Dutch-speaking beginning readers of Grade 1 already read via the lexical route, the interpretation of the first-letter effect in terms of contextual facilitation was called in question. The alternative explanation that response competition underlies the first-letter effect seemed to provide a better account of the results. I reasoned that readers presented with word-like stimuli automatically start to process these letter strings, and that the outcome of this process competes with the outcome of the task at hand, naming the first letter.

Explaining the first-letter effect in terms of response competition does not in itself reveal the nature of the reading process, that is, whether the involuntary reading of the word occurred lexically or nonlexically. In fact, the results of the Experiment 5.1 left unresolved whether phonology did play a role at all. Experiment 5.2, however, did answer this question. It was concluded that the most important factor in the first-letter task was pronounceability, indicating that phonology was fundamental. Experiment 5.3 was a more direct test of the role of phonology in the first-letter task. Assuming that response competition underlies the first-letter effect, the conclusion seemed inevitable that phonologic recoding of the legal nonword stimuli was responsible for the robust congruency effect that emerged there. In Experiment 5.2 words and legal nonwords, including pseudohomophones, caused statistically equal levels of interference, but less than orthographically illegal, unpronounceable, nonwords, i.e., unpronounceable stimuli. In Experiment 5.1, however, words caused less interference than orthographically legal nonwords, which in turn caused less interference than orthographically illegal nonwords. The phonologic coherence hypothesis would predict precisely this order, because resonance in words emerges faster than in legal nonwords, which in turn is established faster than in orthographically illegal nonwords. Thus, the conclusion seems justified that the effects established with the first-letter task can be explained without having to assume that words are processed lexically.

To substantiate the response-competition interpretation I will supply additional evidence for the hypothesis. The evidence was provided by Goutbeek and Van Leerdam

(Master's Theses in preparation). Goutbeek found large congruency effects with words in both adults and in beginning readers with only 3 months of formal reading instruction. Naming the first letter [ei] of 'amen', is congruent with pronouncing the word [eimen], but naming the first letter of 'abuse' [əbju:s] or 'abscess [œbses] is not. Naming the first phoneme of congruent words was faster than of incongruent words. Two results with fluent readers in the work by Van Leerdam also support the response competition interpretation. The first is that naming a complete word took 30 ms shorter than naming its first letter. The fact that naming a single isolated letter is in itself a faster process than naming a word (Rossmeissl & Theios, 1982; but see Cattell, 1886), suggests that naming the first letter of a word is delayed. A likely reason is that the word serves as an inhibitor in the first-letter-naming condition. Because no competitor is operative in the word naming task, latencies in that task are relatively short. The second finding of Van Leerdam that I take as support for the response competition explanation is that naming the first letter of short orthographically legal pseudowords (3 letters) is faster than of long orthographically legal pseudowords (5 letters). Presumably the cause of this effect is that automatic processing of the shorter word-like stimuli finishes relatively early, so that the subject in that condition can attend to the actual task earlier than when longer words are presented.

Explaining the first-letter effect in terms of response competition does also immediately raise the question whether in fact the first-letter task is a variant of the Stroop task (Stroop, 1935). Results obtained with the picture-word interference task indeed show notable similarities between this modified Stroop task and the first-letter results found in Chapter 5. In the picture-word interference task subjects are presented with simple line drawings on which words are superimposed. The relevant condition in that task, for the issue here, is picture naming. Several studies (Ehri & Wilce, 1979; Posnansky & Rayner, 1978) show that labelling the picture is facilitated when the word superimposed is congruent with the picture (e.g., the word 'cat' in a drawing of a "cat") as compared to naming a picture without a word superimposed. Naming a picture with a word that is incongruent with the picture (the word 'pig' in a drawing of a "cat") inhibits picture naming as compared to both the congruent and the drawing only conditions (Lupker, 1982; Rayner & Posnansky, 1978; Schadler & Thissen, 1981). This effect has been interpreted in terms of response competition. While processing the picture, a literate subject cannot avoid also reading the superimposed word. Analogously, in the first-letter task, while processing the first letter the subject cannot ignore the word it is part of. The response-competition hypothesis as an explanation of the Stroop effect has been rejected by MacLeod (1991; p. 188; also for an excellent and extensive review of the Stroop task). The response-competition explanation, however, does not only (MacLeod, 1991, p. 188) account for the majority of fairly well established Stroop phenomena, but in the alternative PDP-model of the Stroop effect (Cohen, Dunbar, & McClelland, 1990; see also Phaf, Van der Heijden, & Hudson, 1990) it returns as a major property, emerging from the system. The difference between traditional explanations and that of the PDP-models is that in the PDP-models competition is not restricted to the response stage, but can occur anywhere in the process. Evidence for multilevel causes of interference is supplied by several empirical studies in the picture-word interference task. Golinkoff and Rosinski (1976) showed that interference is at least partly semantically based. Rayner and Posnansky (1978) proved that word shape also played a role, and Lupker (1982) showed that picture naming suffered when superimposed words

were orthographically and phonologically similar to the picture label. Finally, and most interestingly, Bakan and Alperson (1967) were able to establish the role of pronounceability in a classic Stroop task. In the classic Stroop task incongruence between a colour name (for instance, the word 'red') and the name of the colour of the ink in which it is printed (for instance 'blue' ink) slows down colour naming times. Bakan and Alperson (1967) used stimuli that differed in degree of pronounceability. They found that colour naming was easier when the nonword stimulus was highly unpronounceable (e.g., FJQ) as compared to a pronounceable nonword (e.g., DAP). Competition in the picture-word interference task thus seems to occur at all levels of stimulus processing.

The question remains whether competition in the first-letter task is also due to multi-level interference. I do not oppose such a hypothesis, but as yet there are reasons to believe that in the first-letter task pronunciation is the main cause of the interference. For interference (or competition) to occur, the two competing processes have to have something in common. It seems plausible that in the first-letter task the major competitive process is pronunciation, because one process is based on a word or word-like stimulus, whereas the other is always based on a letter. Interference at the semantic level is unlikely, because the semantic content of a single isolated letter is rather limited as compared to that of a word. The same holds for orthography. Indeed, as was said before, the most influential variable in first-letter naming appears to be pronounceability.

A rather robust finding in the picture-word interference task is that the Stroop effect emerges within all reader groups from the very beginning literate to the fluent adult reader (Golinkoff & Rosinski, 1976; Posnansky & Rayner, 1977; Rosinski, Golinkoff, & Kukish, 1975; Seegers, Feenstra, & Hooft, 1985). Schadler and Thissen (1981) used the classic Stroop colour-naming task. They showed that colour naming interference effects started off at the beginning of literacy (beginning readers), peaked between Grades 2 and 4, and decreased again with increasing reading experience (Grade 6). In all groups, however, the Stroop effect occurred. The present first-letter effect is also apparent in both beginning and fluent readers, and the size of the effect is also much larger in beginning than in fluent readers. It appears that with age an ability to suppress irrelevant information develops, but complete suppression is never acquired (see also Gernsbacher, 1993; Ridderinkhof, 1993; Simpson & Foster, 1986). To end this discussion, it seems justified to conclude that the first-letter task is a modified Stroop task.

Spelling verification and reading level

The major conclusion of Chapter 5 was that the reading process of beginning and fluent readers was most parsimoniously explained in terms of a single-route model that retained the phonologic route. If phonologic mediation is fundamental to reading, differences between beginners and fluent readers should not occur at the phonology level. To check whether that was true, I used the same experimental paradigms with beginning readers as Van Orden (1991) and Van Orden et al. (1992) to investigate the reading behaviour of fluent readers. The results of these experiments were reported in Chapter 6, and the effects were remarkably similar to those found with the fluent readers of Van Orden et al. (1992). Both fluent and beginning readers had more trouble detecting pseudohomophone errors than

errors that were based on stimuli orthographically similar, but phonologically dissimilar from the target words (spelling-control stimuli). Thus, both beginning and fluent readers showed phonologic effects. But there was also an important difference between the two groups. The fluent and beginning readers alike managed to find the majority of the spelling-control errors, but beginning readers were much less capable of detecting pseudohomophones (despite their knowledge of the words' spelling) than were fluent readers. What could have caused the difference between reader groups?

According to Van Orden (1987) and Van Orden et al. (1990) phonologic activation causes semantic activation, which in turn is followed by a verification process. This verification process is in fact a spelling-check mechanism. It is likely that missing a spelling error is caused by a failing spelling-verification mechanism. The reason why the majority of the spelling controls is detected, is that the outcome of the phonologic process is congruent with the outcome of the spelling-verification process: they both tell the system that something unfamiliar came about. On the other hand, when a reader comes across a pseudohomophone the outcome of the phonologic process is no cause for alarm, but the outcome of the spelling-verification process can tell the reader that an error has occurred. In case a reader is presented with a pseudohomophone a well-developed spelling-verification mechanism is therefore indispensable. The results of Chapter 6 seem to indicate that the spelling-verification process of beginning readers is not fully developed yet. Spelling-verification differences between reader groups are already apparent early in the development of literacy, as shown by the higher pseudohomophone error detection rate in good as compared to poor readers in Chapter 6 (see also Ormrod, 1985 and Supramaniam, 1983).

The use of nonwords in word-recognition tasks

The phonologic effects in the proofreading, lexical-decision and semantic-categorisation tasks (Chapter 6) were obtained with nonwords. This necessitates comment. Van Orden et al. (1992) and I used nonwords (pseudohomophones and spelling controls), because phonologic effects obtained with nonwords have to be attributed to phonologic mediation, whereas according to adherents of the dual-route model, phonologic effects caused by the use of words could be the result of pre-lexical or post-lexical processing. However, a problem arising when using nonwords is that no logical imperative guarantees that the same (phonologic) processes underlie the reading of words and nonwords. This provides proponents of dual-route theory (Coltheart, Laxon, Rickard, & Elton, 1988; Patterson & Coltheart, 1987) with the argument that phonologic effects found with nonwords do not prove that phonologic mediation is fundamental to normal word reading.

However, Van Orden, Johnston, and Hale (1988) argue that their results provide unequivocal evidence for assembled phonology in normal word reading, because their subjects mistakenly took an equal number of words (i.e., homophones; for instance, 'hare' as a member of the category 'part of the human body') and nonwords (i.e., pseudohomophones; for instance, 'sute' as a member of the category 'article of clothing') as the instance of the target items 'hair' and 'suit' respectively. The conclusion that similar processes underlie reading words and nonwords is also supported by the finding that, the response latencies of

homophones and pseudohomophones are remarkably close (but see Coltheart, Avons, Masterson, & Laxon, 1991).

As I said above, adherents of the dual-route model do not need to accept proof of a phonologic effect established with nonwords as evidence for nonlexical normal reading. When proponents of two opposing models do not accept each other's empirical results as evidence for their respective hypotheses a deadlock has occurred. The solution to this problem therefore, should not be sought at the empirical, but at the theoretical level. An important axiom to which all sciences subjugate is the principle of economy. If two rivalling models can both explain the current empirical facts, the simplest model is preferred. In case of the choice between a single-route versus a dual-route model the decision seems to be clear. Van Orden et al.'s (1990) single-route model, the phonologic coherence hypothesis, appears to be a simpler, but at least as powerful a model.

The relationship between reading and spelling

In Chapter 7, the phonologic coherence hypothesis was contrasted with the dual-route model. The main focus of this study concerned the relationship between reading and spelling. To deny a relationship of any sort between reading and spelling means ignoring the outcome of several studies that show moderate to high correlations between reading and spelling skills (Malmquist, 1958, in Frith, 1980; Mommers, Van Leeuwe, Oud, & Janssens, 1985; Van Bon & Bouwens, 1990). On the other hand, the imperfect correlations also show that reading is not simply the reverse of spelling. Spelling skill usually not only lags behind reading skill (Seymour & Porpodas, 1980), but it also seems more difficult to master. This appears from the fact that it is relatively easy to find children whose reading skill is at a normal level and whose spelling performance is below the average, whereas it is much harder to find the opposite (Frith, 1980). In the discussion about the assumed relationship between reading and spelling two major parties can be distinguished. There are researchers who emphasise the similarities between the skills, and claim that the processes (Ehri, 1980; Gough, Juel, & Griffith, 1992) or representations (Perfetti, 1992) underlying reading and spelling are in fact the same. Other researchers stress the differences between reading and spelling processes (Bryant & Bradley, 1980; Frith & Frith, 1980). Frith (1979) concluded that reading occurred by 'eye' and spelling by 'ear', which suggests that the spelling process is phonologically mediated, but the reading process is not. According to the phonologic coherence hypothesis both reading and spelling are phonologically constrained. But, unlike predictions from dual-route theory, this does not automatically imply that the reading process is simply the opposite of the spelling process, and that being able to read a word entails being able to spell it, and vice versa (see Chapter 7). According to the phonologic coherence hypothesis, learning to read words is based on the connections between orthographic and phonologic subsymbols, whereas the spelling process is based on the opposite connections, that is between phonologic and orthographic subsymbols. Thus, establishing the connections in one direction does not automatically imply the emergence of the ones in the reverse direction. In Chapter 7, I showed that a spelling training which did not involve reading was not beneficial for learning to read, although it was useful for learning to spell. Another indication that reading and spelling are separate processes was

that knowledge of the spelling of a semantically-familiar word does not affect the reading process.

The outcome that reading does not benefit from spelling finds it counterpart in the result that spelling does not profit from a reading training (Chapters 2, 3, and 4). Beginning readers who were to read words did not establish a representation that was useful in the spelling process, not even after 30 presentations of each word (see Chapter 2). The suggestion that the words used in Chapter 2 were too difficult to learn for beginning readers and spellers seems inappropriate, because in Chapter 3 words of equal orthographic complexity as in Chapter 2 were used, and the subjects in that study acquired superior spelling knowledge after only two presentations with a word in a spelling training. In Chapter 3, three spelling-instruction methods were applied: copying, problem naming, and oral spelling. The most effective instruction method for both good and poor spellers appeared to be oral spelling. The relationship between phonology and orthography is most explicitly trained in the oral-spelling method. Moreover, in the oral-spelling method the subject had to try to recover the sequential ordering of the letters by heart, whereas the copying and problem-naming methods enabled the subjects to inspect the words visually during the training. In this respect, the oral-spelling method mimics normal spelling more closely than problem naming and copying do. In Chapter 4, no differential effects occurred from the spelling-instruction methods problem naming and word composition, but they induced better spelling performance than a reading training (Experiment 4.2). Both these spelling-training methods allowed the subject to inspect the stimulus word visually during the spelling training. It appears that the larger the degree of overlap in processes between (spelling) training and (spelling) test, the better the performance on the test (Roediger, 1990; see also the 'General Discussion' Section of Chapter 7).

The prediction from the phonologic coherence hypothesis that the spelling process is phonologically mediated does receive unequivocal support from the findings in this dissertation. Evaluation of spelling errors shows that both beginning and fluent readers make proportionally more phonologically correct than phonologically incorrect errors, irrespective of the level of orthographic knowledge one has of the word to be spelled (Experiments 7.3 and 7.4; Barron, 1980). The difference between good and poor spellers, and between normals and dyslexics is reflected in the poor spellers' overall weaker knowledge of standardised sound-to-spelling correspondences, and does not reside in differences in spelling processes (Bruck & Treiman, 1990; Frith, 1978; Waters, Bruck, & Malus-Abramowitz, 1988).

Phonologic mediation is fundamental to spelling

The first to point out that phonology plays an important role in even very young children was Read (1971, 1975 in Goswami & Bryant, 1990; pp. 49-57). He showed that pre-readers' invented spellings are not random. Highly skilled literates would probably classify the misspelling 'chrap' for 'trap' as a phonologically incorrect error (an analogous example in Dutch is 'bir' or biir' for 'beer'). However, both Treiman (1985) and Read in (1986 in Goswami & Bryant, 1990) show that it is a phonologically incorrect error only in the eyes of the fluent reader who knows about the standardised representation of speech sounds. Pre-

literates' judgement of speech sounds differs from that of fluent readers, probably because their perception of speech is not affected by orthographic knowledge. The perception of sounds and the production of spellings of pre-literates appear to represent the phonetic aspects of words more closely than those of fluent readers.

The fundamental role of phonology in spelling has been less disputed by adherents of the dual-route model than its role in reading (amongst others Bruck, 1988; Frith, 1979; Kreiner, 1992; Marsh, Friedman, Welch, & Desberg, 1980). Ehri (1980; see also Farnham-Diggory & Simon, 1975) is probably the least convinced that phonology is fundamental to reading and spelling once the reader has acquired a certain level of literacy. One finding of Ehri and Wilce (1982) corroborates the assumption that phonology is fundamental to reading: In a task in which beginning readers had to judge whether a letter was present in a word that was named, they appeared better in judging that a pronounced letter (for instance, 't' in "straw") was present than that a silent letter was present (for instance, 'w' in "straw"). However, the other result in their study seemed to indicate that with increased reading and spelling skill the visual (orthographic) aspect becomes more salient. They based this conclusion on the counterintuitive finding that silent letters were better cues for the recall of words previously judged on the presence of letters, than pronounced letters. This finding supports the idea of Ehri, which she formulated in her study of 1980 (p. 338) that, "...orthographic images can be scanned like real words seen in print...", and "...they include all of the letters in a word's spelling...". Careful analysis of the experimental tasks, however, revealed a serious confounding. In the experiments in which letters served as the cues for recalling words, it appeared that in the condition where silent letters were cues, significantly fewer or no other words than the target word appeared to have the cue letter, whereas in the pronounced letter condition the cue letter was shared by on average three to four words. The conclusion of Ehri and Wilce (1982), that when silent letters gain entrance to memory they acquire a special status, might still be true, but not because of their finding that silent letters are better cues, since that effect can be explained by the fact that in the silent letter conditions there was less confusion.

Van Rijnsoever (1988) argues that the visual-orthographic structure, as assessed by positional bigram frequency (PBF), exerts an influence early in the development of reading and spelling. He, however, seriously neglects the role of phonology. Van Rijnsoever (1988) used a letter cancellation task, in which subjects had to look for the letter 'e' in lists of single words. His results showed that missing the letter 'e' in words with a high PBF was higher than missing the 'e' in words with a low PBF, but only when the target letter appeared in the second syllable of the word. This effect was apparent in beginning readers who only had had four months of formal reading and spelling instruction. Two problems with his set of experiments were noticed by Reitsma and Van Daal (1989). First, parts of the experimental materials were probably not within the subjects' passive vocabulary. Secondly, the index of the visual-orthographic structure, the PBF, was based on an adult positional letter frequency count (Rolf & Van Rijnsoever, 1984). Reitsma and Van Daal (1989) replicated Experiment 3 of Van Rijnsoever with beginning readers of Grade 1. When they used materials based on the adult PBF they found effects similar to those of Van Rijnsoever, but when the materials were based on a PBF of a corpus drawn from reading materials of the subject group investigated, the effect of PBF vanished completely. Reitsma (1990) provided yet another possible explanation for the effect reported by Van Rijnsoever, namely, that it might be a

position effect. The target letter 'e' was more often missed when its position was more to the right in the word. It appeared that, particularly in the second syllable, the target letter was more to the right in the high-PBF condition than in the low-PBF condition.

Again, when I had a closer look at the materials a major confounding in Van Rijnsoever's study was apparent: PBF was almost perfectly correlated with the pronunciation of the letter 'e' in the two-syllable word stimuli. The letter 'e' in syllables with a high PBF was always a schwa, whereas in syllables with a low PBF it had the standard [ɛ] pronunciation in more than 90% of the cases. The PBF-effect could thus also be explained by the assumption that phonologic recoding of the words occurred while looking for the target letter 'e', and that this process negatively affected performance when the target letter did not obey the standard pronunciation rule, as is the case with the schwa-pronunciation. Consistent with this assumption, Corcoran (1966, 1967; see also Drewnowski & Healy, 1982, MacKay, 1968) found phonologic (Corcoran used the term 'acoustic') effects in a letter cancellation and in a proofreading task: a silent 'e' was less likely to be detected than a pronounced 'e'. Thus, the results of both the studies of Ehri and Wilce (1982) and of Van Rijnsoever (1988) can be explained upholding the view that phonology is fundamental to spelling.

Before presenting my final conclusion I will indicate briefly the fundamental role of phonology in three domains which have not been a topic of investigation of this dissertation. The majority of researchers investigating the reading process of reading-disabled children and adult illiterates have stressed the importance of the development of phonologic awareness for reading, and have referred to the reciprocal relationship between phonemic segmentation skill and reading (amongst others, Liberman, Shankweiler, Fischer, & Carter, 1974; Morais, Cary, Alegria, & Bertelson, 1979; Read, Zhang, Nie, & Ding, 1986; see also Brady & Shankweiler 1991; Goswami & Bryant, 1990). A second domain in which phonology is now thought to play a more critical role than was previously assumed is in the reading of alphabetical languages with deep orthographies other than English, as in French (Ferrand & Grainger, 1992), and Hebrew (Frost, in press; Frost & Kampf, 1993; Navon & Shimron, 1980), and in logographic writing systems, as in Chinese (Lam, Perfetti, & Bell, 1991; Perfetti & Zhang, 1991), and Japanese Kanji (Wydell, Patterson, & Humphreys, 1993). The third field in which phonology rather unexpectedly appears to be of major importance is in the reading of congenitally or prelingually profound deaf people. Hanson and Fowler (1987) found phonologic effects in the word reading of deaf adults, using a lexical-decision task. Leybaert and Alegria (1993) showed an effect of phonology in adolescents using a Stroop task, and Dodd (1980) found similar phonologic effects in deaf and hearing adolescents performing a spelling task (see also Hanson, 1991; Hanson, Liberman, & Shankweiler, 1984).

In sum, not only the empirical results of this dissertation, but also, an overwhelming number of studies in different fields show that a single-route model, in which phonology constrains word recognition at an early stage, suffices to explain reading and spelling behaviour observed in beginning and fluent readers.

9

NOTES

1. An additional analysis was performed on a subset of the materials of Experiment 1 to check whether these findings were an artefact caused by the choice of materials (some of the one-syllable pseudohomophones contained a letter cluster 'sg', homophonic to 'sch', that does not occur within a single syllable in Dutch orthography). An analysis of variance on the mean naming latencies for the training words and pseudohomophones after removing responses to 'sg' and 'sch' stimuli was performed. A 2 (stimulus type: training words vs. pseudohomophones) by 5 (presentation frequency: 3, 5, 7, 9, 18) ANOVA again showed a significant main effect of stimulus type, $F(1, 44) = 7.12$, $p < .05$. However, the effect of presentation frequency dropped to a marginally significant level, $F(4, 176) = 2.17$, $.05 < p < .10$. Exclusion of 33% of the data is probably responsible for this drop. Again no significant interaction effect emerged ($p > .10$).

2. As in Experiment 2.1 (see note 1), an analysis was performed on a subset of the responses (naming times for words containing a 'sch' and pseudohomophones containing an 'sg' were discarded). The analysis, in which stimulus type (training words vs. pseudohomophones) and presentation frequency (0, 1, 2, 3, 6, 9) were tested, showed a significant main effect of presentation frequency ($F(5, 160) = 5.81$, $p < .001$), and a marginally significant effect of stimulus type, $F(1, 32) = 3.81$, $.05 < p < .10$. The interaction effect between stimulus type and presentation frequency was not significant ($F < 1$). Only in Frequency condition 9 did a significant difference occur between training words and pseudohomophones, $F(1, 32) = 5.8$, $p < .05$, indicating that naming times of training words were shorter than naming times of pseudohomophones.

3. Experiment 4.1 was conducted at a time in their formal reading and spelling instruction that the children had only been presented with the ambiguous [ɛi] phoneme; the [ɑu] phoneme and the final 'd' as [t] sound have not been taught. Therefore, in addition to the [ɛi] phoneme the less ambiguous [s]-[z] and [v]-[f] were selected as spelling problems. It is well known that beginning spellers (particularly in the western part of the Netherlands) sometimes have problems with the [s]-[z] and [v]-[f].

4. The spelling test that was used is fairly reliable (Cronbach's alpha = .88; $N = 218$) and correlates $r = .85$ ($N = 207$) with the spelling test of Mommers and Van Dongen (1986).

5. It is impossible to assess the spelling errors on phonologic correctness in an entirely systematic way. The errors were categorised on the basis of a number of arbitrarily chosen rules. In doing so internal consistency was emphasised.

6. A separate analysis on the legal vs. illegal nonwords showed that the difference was significant, $F(1, 27)= 4.37$, $p < .05$.

7. A first-letter experiment was run with two groups of beginning readers (Grade 1; good and medium readers) drawn from a population similar to the one used in the experiments reported here. Stimuli used in the experiment were: letters, words, pronounceable nonwords (legal and illegal) and unpronounceable nonwords. The main effect of stimulus type was significant, $F(3, 111) = 94.81$, $p < .001$. For both groups of readers naming an isolated letter (842 ms) was significantly faster than naming the first letter of any other stimulus type. The isolated-letter condition was in fact as much as 183 ms faster than the fastest letter-string condition. A similar experiment was executed with fluent readers. Stimuli in this condition were: letters, words, pseudohomophones, and illegal unpronounceable nonwords. Again the main effect of stimulus type was significant, $F(3,57) = 32.11$, $p < .001$. Naming an isolated letter (508 ms) was significantly faster than naming the first letter of any other stimulus type. The isolated-letter condition was 30 ms faster than the fastest letter-string condition.

8. Three of the twelve examples ('cra', 'fra' and 'tra') of the orthographically illegal nonword stimuli used by Rossmeissl and Theios (1982, p. 446) are in fact pronounceable, although they stated in their 'Method' Section that: "The resulting letter strings were highly unpronounceable and did not follow common English orthography" (p. 448). It is not clear whether the proportion of pronounceable example stimuli (25%) reflects the proportion of pronounceable experimental stimuli in their orthographically illegal nonword condition.

9. Van Orden et al. (1990) distinguish between a rule-governed coding process in which discrete rules operate between graphemes and phonemes, and a rule-like coding process in which a statistical regularity holds between orthography and phonology.

10. Van Orden's (1987) measure of orthographic similarity (OS) is an adaptation of Weber's (1970) measure of graphic similarity (GS). The computation of the formula will be illustrated by means of a Dutch example, namely, the pseudohomophone-word pair **ZAUT** and **ZOUT**. Foil here refers to either a pseudohomophone or a spelling control.

$$OS = (\text{GS of foil and target}) / (\text{GS of target and itself})$$

$$GS = 10\{[(50F + 30V + 10C)/A] + 5T + 27B + 18E\}.$$

The first value refers to the pair **ZAUT-ZOUT** and the second to the pair **ZOUT-ZOUT**

F = number of pairs of adjacent letters in the same order shared by word pairs:	F=1, F=3
V = number of pairs of adjacent letters in reverse order shared by word pairs:	V=0, V=0
C = number of single letters shared by words:	C=3, C=4
A = average number of letters in the two words:	A=4, A=4
T = ratio of number of letters in the shorter word to the number in the longer	T=1, T=1
B = 1 if the first letter in the two words is the same; otherwise B = 0	B=1, B=1
E = 1 if the last letter in the two words is the same; otherwise E = 0	E=1, E=1

GS of foil and target: **ZAUT** and **ZOUT** is

10{[(50(1) + 30(0) + 10(3))/4} + 5(1) + 27(1) +18(1)} = 700

GS of target and itself: **ZOUT** and **ZOUT** is

10{[(50(3) + 30(0) + 10(4))/4} + 5(1) + 27(1) +18(1)} = 975

Thus, the orthographic similarity between ZAUT and ZOUT is
OS = (GS of zaut-zout) / (GS of zout-zout) = 700 / 975 = .72.

11. Scores on the non-verbal (Raven, 1958) and verbal (vocabulary test of the RAKIT; Bleichrodt et al., 1984) intelligence tests were also obtained. In general no significant differences occurred between good and poor readers on these variables. Scores on a spelling test (Mommers & Van Dongen, 1986) were also available, and the differences between good and poor readers were generally significant on the spelling test. This is mainly due to a significant correlation (r = .50) that exists between reading and spelling in first grade.

12. The same lexical-decision task with reading training instead of spelling training as a preparatory task was performed with a similar group of beginning readers. The results of that study were nearly identical to those of the present investigation.

13. The idea to include the lexical-decision task was kindly provided by Ken Paap (personal communication).

14. An experiment similar to the present Experiment 7.1 has been executed before, but did not include the lexical-decision task and lacked proper counterbalancing of the stimuli. The results, however, were completely in accordance with those of Experiment 7.1.

15. To check possible confounding by orthographic complexity, a manipulation on the materials was performed. From each ambiguous stimulus an unambiguous stimulus was created by replacing the ambiguous phoneme by an unambiguous one, and the other way round for the unambiguous stimuli. Subjects from the same population as those who took part in Experiment 7.2 had to perform a naming task on one half of the original materials supplemented with the other half of the new materials. If spelling ambiguity was responsible for the effect then a main effect of spelling ambiguity should again show up, whereas if orthographic complexity underlies the spelling ambiguity effect an interaction between materials and spelling ambiguity should emerge. A 2 (materials: old vs. new) by 2 (spelling ambiguity: ambiguous vs. unambiguous) analysis of variance on the responses revealed no significant main effects but a highly significant interaction effect between materials and spelling ambiguity, $F(1, 17) = 55.59$, $p < .001$. In the old materials condition the ambiguous stimuli (610 ms) were named significantly slower than the unambiguous stimuli (578 ms) as was the case in Experiment 7.2, but in the new materials condition the unambiguous (605 ms) stimuli were named significantly slower than the ambiguous ones (586 ms). It thus seems that the spelling ambiguity effect in Experiment 7.2 was caused by the higher complexity of the ambiguous stimuli and not because of the fact that these stimuli contained an ambiguous spelling problem.

10

REFERENCES

Assink, E. M. H. (1981). Schrijfstrategieën van intuïtieve spellers bij moeilijke werkwoordsvormen. *Tijdschrift voor Taalbeheersing*, *3*, 55-66.

Assink, E. M. H. (1990). Learning to spell. In P. Reitsma & L. Verhoeven (Eds.), *Acquisition of Reading in Dutc*h (pp. 65-76). The Netherlands, Dordrecht: Foris Publications.

Assink, E. M. H., Van Doorn-Van Eijsden, M., & Van Hees, H. (1982). Schrijfstrategieën van onervaren spellers bij lexicaal bepaalde homofonen. *Tijdschrift voor Taalbeheersing, 4*, 314-327.

Backman, J., Bruck, M., Hebert, M., & Seidenberg, M. S. (1984). Acquisition and use of spelling-sound correspondences in reading. *Journal of Experimental Child Psychology, 38,* 114-133.

Bakan, P., & Alperson, B. (1967). Pronounceability, attensity, and interference in the color-word test. *American Journal of Psychology, 80*, 416-420.

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension Processes in Reading* (pp. 9-32). Hillsdale, NJ: Erlbaum.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human perception and Performance 10*, 340-357.

Baluch, B., & Besner, D. (1991).Visual word recognition: evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 644-652.

Baron, J., Treiman, R., Wilf, J. F., & Kellman, P. (1980). Spelling and reading by rules. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 159-194). London: Academic Press.

Barron, R. W. (1980). Visual and phonological strategies in reading and spelling. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 195-213). London: Academic Press.

Barron, R. W. (1986). Word recognition in early reading: a review of the direct and indirect access hypotheses. *Cognition, 24*, 93-119.

Bleichrodt, N., Drenth, P. J. D., Zaal, J. N., & Resing, W. C. M. (1984). *Revisie Amsterdamse Kinder Intelligentie Test, RAKIT*. (Revised Amsterdam Children Intelligence Test). The Netherlands, Lisse: Swets & Zeitlinger.

Booij, G. E. (1981). *Generatieve fonologie van het Nederlands*. The Netherlands, Utrecht: Het Spectrum

Brady, S. A., & Shankweiler, D. P. (1991). *Phonological processes in literacy*. Hillsdale, NJ: Lawrence Erlbaum.

Bruck, M. (1988). The word recognition and spelling of dyslexic children. *Reading Research Quarterly, 23*, 51-69.

Bruck, M., & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics: the case of initial consonant clusters. *Journal of Experimental Child Psychology, 50,* 156-178.

Bryant, P. E., & Bradley, L. (1980). Why children sometimes write words which they do not read. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 355-370). London: Academic Press.

Caesar, F. B. (1975). *Eén-Minuut-Test voor de technische leesvaardigheid.*. The Netherlands, Tilburg: Zwijsen.

Caesar, F. B. (1979). *Veilig leren lezen: Structuurmethode voor het aanvankelijk leesonderwijs*. The Netherlands, Tilburg: Zwijsen.

Caesar, F. B. (1980). *Handleiding bij Veilig leren lezen: Structuurmethode voor het aanvankelijk leesonderwijs*. The Netherlands, Tilburg: Zwijsen.

Carello, C., Turvey, M. T., & Lukatela, G. (1992). Can theories of word recognition remain stubbornly nonphonological? In: R. Frost and L. Katz (Eds.), *Orthography, Phonology, Morphology, and Meaning* (pp. 211-226). The Netherlands, Amsterdam: Elsevier Science Publishers.

Cattell, J. M. (1886). The time taken up by cerebral operations. *Mind: A Quarterly Review of Psychology and Philosophy, XI,* 524-538.

Chase, C. H., & Tallal, P. (1990). A developmental, interactive activation model of the word superiority effect. *Journal of Experimental Child Psychology, 49*, 448-487.

Chastain, G. (1981). Phonological and orthographic factors in the word-superiority effect. *Memory & Cognition, 9*, 389-397.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335-359.

Clarke, R., & Morton, J. (1983). Cross modality facilitation in tachistoscopic word recognition. *Quarterly Journal of Experimental Psychology, 35A*, 79-96.

Cohen, G. (1980). Reading and searching for spelling errors. In U. Frith (Ed.), *Cognitive Processes in Spelling* (pp. 135-155). London: Academic Press.

Cohen J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332-361.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). N.Y.: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589-608.

Coltheart, V., Avons, S. E., Masterson, J., & Laxon, V. J. (1991). The role of assembled phonology in reading comprehension. *Memory & Cognition, 19*, 387-400.

Coltheart, V., Laxon, V. J., Keating, G. C., & Pool, M. M. (1986). Direct access and phonological encoding processes in children's reading: effects of word characteristics. *British Journal of Educational Psychology, 56*, 255-270.

Coltheart, V., Laxon, V., Rickard, M., & Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 387-397.

Corcoran, D. W. (1966). An acoustic feature in letter cancellation. *Nature, 210*, 658.

Corcoran, D. W. (1967). Acoustic factors in proof reading. *Nature, 214*, 851-852.

Cunningham, A. E., & Stanovich, K. E. (1990). Early spelling acquisition; writing beats the computer. *Journal of Educational Psychology, 82*, 159-162.

Daneman, M., & Stainton, M. (1991). Phonological recoding in silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 618-632.

De Groot, A. M. B. (1987). The priming of word associations: A levels-of-processing approach. *The Quarterly Journal of Experimental Psychology, 39A*, 721-756.

De Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 824-845.

Dodd, B. (1980). The spelling abilities of profoundly pre-lingually deaf children. In U. Frith (Ed.), *Cognitive Processes in Spelling* (pp. 423-440). London: Academic Press.

Dumont, J. J. (1984). *Lees- en spellingsproblemen*. The Netherlands, Rotterdam: Lemniscaat.

Drewnowski, A., & Healy, A. F. (1982). Phonetic factors in letter detection: a reevaluation. *Memory & Cognition, 10*, 145-154.

Ehri, L. C. (1980). The development of orthographic images. In U. Frith (Ed.), *Cognitive Processes in Spelling* (pp. 311-338). London: Academic Press.

Ehri, L. C., & Roberts, K. T. (1979). Do beginners learn printed words better in contexts or in isolation? *Child Development, 50*, 675-685.

Ehri, L. C., & Wilce, L. S. (1979). Does word training increase or decrease interference in a Stroop task? *Journal of Experimetal Child Psychology, 27*, 352-364.

Ehri, L. C., & Wilce, L. S. (1982). The salience of silent letters in children's memory for word spellings. *Memory & Cognition, 10*, 155-166.

Ehri, L. C., & Wilce, L. S. (1987). Does learning to spell help beginners learn to read words? *Reading Research Quarterly, 22*, 47-65.

Farnham-Diggory, S., & Simon, H. A. (1975). Retention of visually presented information in children's spelling. *Memory & Cognition, 3*, 599-608.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: evidence from masked non-word priming. *The Quarterly Journal of Experimental Psychology, 45a*, 353-372.

Fisher, F. W., Shankweiler, D., & Liberman, I. Y. (1985). Spelling proficiency and sensitivity to word structure. *Journal of memory and language, 24*, 423-441.

Frith, U. (1979). Reading by eye and writing by ear. In P. A. Kolers, M. Wrolstad, & H. Bouma (Eds.), *Processing of Visible language*, 1 (pp. 379-390). N.Y.: Plenum Press.

Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive Processes in Spelling* (pp. 495-515). London: Academic Press.

Frith, U., & Frith, C. (1980). Relationships between reading and spelling. In: J. F. Kavanagh & R. L. Venezky (Eds), *Orthography, Reading, and Dyslexia* (pp. 287-295). Baltimore: University Park Press.

Frost, R. (in press). Phonological computation and missing vowels: mapping lexical involvement in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Frost, R., & Kampf, M. (1993). Phonetic recoding of phonologically ambiguous printed words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 23-33.

Frost, R., & Katz, L. (1989). Orthographic depth and the interaction of visual and auditory processing in word recognition. *Memory & Cognition, 17*, 302-310.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science, 4*, 294-298.

Gibson, E. J., & Levin, H. (1975). *The Psychology of Reading* (pp. 276-290). Cambridge: MIT Press.

Golinkoff, R. M., & Rosinski, R. R. (1976). Decoding, semantic processing, and reading comprehension skill. *Child Development, 47,* 252-258.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* U.K., Hove: Lawrence Erlbaum

Gough, P. B., Juel, C., & Griffith, P. L. (1992). Reading, spelling, and the orthographic cipher. In: P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 35-48). Hillsdale, NJ: Lawrence Erlbaum.

Grossberg, S., & Stone, G. (1986). Neural dynamics of word recognition and recall: attentional priming, learning, and resonance. *Psychological Review, 93,* 46-74.

Hanson, V. L. (1991). Phonological processing without sound. In: S. A. Brady & D. P. Shankweiler (Eds.), *Phonological Processes in Literacy* (pp. 153-161). Hillsdale, NJ: Lawrence Erlbaum.

Hanson, V. L., & Fowler, C. A. (1987). Phonological coding in word reading: evidence from hearing and deaf readers. *Memory & Cognition, 15,* 199-207.

Hanson, V. L., Liberman, I. Y., & Shankweiler, D. (1984). Linguistic coding by deaf children in relation to beginning reading success. *Journal of Experimental Child Psychology, 37,* 378-393.

Hatfield, F. M., & Patterson, K. E. (1983). Phonological spelling. *Quarterly Journal of Experimental Psychology, 35A*, 451-468.

Hawkins, H. L., Reicher, G. M., Rogers, M., & Peterson, L. (1976). Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 380-385.

Hoffman, P. R., & Norris, J. A. (1989). On the nature of phonological development: evidence from normal children's spelling errors. *Journal of Speech and Hearing Research, 32*, 787-794.

Holligan, C., & Johnston, R. S. (1991). Spelling errors and phonemic segmentation ability: the nature of the relationship. *Journal of Research in Reading, 14*, 21-32.

Holmes, V. M., & Ng, E. (1993). Word-specific knowledge, word-recognition strategies, and spelling ability. *Journal of Memory and Language, 32*, 230-257.

Huey, E. B. (1908). *The Psychology and Pedagogy of Reading*. Cambridge, Mass: The M.I.T. Press (reprinted in 1968)

Humphreys, G. W., & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *The Behavioral and Brain Sciences, 8,* 689-705.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120*, 358-394.

Jorm, A. F. (1981). Children with reading and spelling retardation: functioning of whole-word and correspondence-rule mechanisms. *Journal of Child Psychology and Psychiatry, 22*, 171-178.

Kohnstamm, G. A., Schaerlaekens, A. M., De Vries, A. K., Akkerhuis, G. W., & Froonincksx, M. (1981). *Nieuwe streeflijst woordenschat voor 6-jarigen.*. Lisse: Swets & Zeitlinger.

Kreiner, D. S. (1992). Reaction time measures of spelling: testing a two-strategy model of skilled spelling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 765-776.

Kreiner, D. S., & Gough, P. B. (1990). Two ideas about spelling: rules and word-specific memory. *Journal of Memory and Language, 29*, 103-118.

Lam, A. S. L., Perfetti, C. A., & Bell, L. (1991). Automatic phonetic transfer in bidialectical reading. *Applied Psycholinguistics, 12*, 299-311.

Langereis, M. P., & Elshout, J. J. (1985). *Het orthografisch woordbeeld bij ervaren spellers: waarom spellen studenten zo slecht*. (VF-Memo 47; psy 4.11.85.77). The Netherlands, Amsterdam: University of Amsterdam, Department of Psychology.

Lashley, K. S. (1951). The problem of serial order in behavior. In L.A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior* (pp. 112-136). N.Y.: Wiley

Lesch, M. F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 285-294.

Leybaert, J., & Alegria, J. (1993). Is word processing involuntary in deaf children? *British Journal of Developmental Psychology, 11*, 1-29.

Liberman, I. Y., Shankweiler, D., Fisher, M. F., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18*, 201-212.

Lukatela, G., Carello, C., & Turvey, M. T. (1990). Phonemic priming with words and pseudowords. *European Journal of Cognitive Psychology, 2*, 375-394.

Lukatela, G., & Turvey, M. T. (1993). Similar attentional, frequency, and associative effects for pseudohomophones and words. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 166-178.

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture-word interference. *Canadian Journal of Psychology, 36*, 349-367.

MacKay, D. G. (1968). Phonetic factors in the perception and recall of spelling errors. *Neuropsychologica, 6*, 321-325.

MacLeod, C. M. (1991). Half a century of reseach on the Stroop effect: an integrative review. *Psychological Bulletin, 109*, 163-203.

Marsh, G., Friedman, M., Welch, V., & Desberg, P. (1980). The development of strategies in spelling. In U. Frith (Ed.), *Cognitive Processes in Spelling* (pp. 339-353). London: Academic Press.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: implications for models of pronunciation assembly and the locus of the word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 14-24.

Mezrich, J. J. (1973). The word superiority effect in brief visual displays: elimination by vocalization. *Perception & Psychophysics, 13*, 45-48.

Mommers, M. J. C., & Van Dongen, D. (1986). *Zorgverbreding bij het leren lezen*. The Netherlands, Tilburg: Zwijsen.

Mommers, M. J. C., Van Leeuwe, J. F. J., Oud, J. H. L., & Janssens, J. M. A. M. (1986). Decoding skills, reading comprehension and spelling; a longitudinal investigation. *Tijdschrift voor Onderwijsresearch, 11*, 97-113.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In: D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 148-197). Hillsdale, NJ: Erlbaum.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*, 323-331.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of verbal Learning and Verbal Behavior, 16,* 519-533.

Morton, J. (1985). Criticising dual-route theory: missing the point. *The Behavioral and Brain Sciences, 8,* 718.

Murphy, J. F., Hern, C. L., Williams, R. L., & McLaughlin, T. F. (1990). The effects of the copy, cover, compare approach in increasing spelling accuracy with learning disabled students. *Contemporary Educational Psychology, 15,* 378-386.

Navon, D., & Shimron, J. (1981). Does word naming involve grapheme-to-phoneme translation? Evidence from Hebrew. *Journal of Verbal Learning and Verbal Behavior, 20*, 97-109.

Nelson, H. E., & Warrington, E. K. (1974). Developmental spelling retardation and its relation to other cognitive abilities. *British Journal of Psychology 65*, 265-274.

Newman, S., Fields, H., & Wright, S. (1993). A developmental study of specific spelling disability. *British Journal of Educational Psychology, 63*, 287-296.

Nicholson, T. (1991). Do children read words better in context or in lists? A classic study revisited. *Journal of Educational Psychology, 83*, 444-450.

Ormrod, J. E. (1985). Proofreading *the cat in the hat*: evidence for different reading styles of good and poor spellers. *Psychological Reports, 57*, 863-867.

Ormrod, J. E., & Jenkins, L. (1989). Study strategies for learning spelling: correlations with achievement and developmental changes. *Perceptual and Motor Skills, 68,* 643-650.

Paap, K. R., McDonald, J. E., Schvaneveldt, R. W., & Noel, R. W. (1987). Frequency and pronounceability in visually presented naming and lexical-decision tasks. In M. Coltheart (Ed.), *Attention & Performance XII* (pp. 221-243). Hillsdale, NJ: Erlbaum.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: the word superiority effect. *Psychological Review, 89*, 573-594.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: still a good horse race. *Psychological Research, 53*, 13-24.

Paap, K. R., Noel, R. W., & Johansen, L. S. (1992). Dual-route models of print to sound: red herrings and real horses. In: R. Frost and L. Katz (Eds.), *Orthography, Phonology, Morphology, and Meaning* (pp. 293-318). Amsterdam: Elsevier Science Publishers B.V.

Patterson, K., & Coltheart, V. (1987). Phonological processes in reading: a tutorial review. In M. Coltheart (Ed.), *The Psychology of Reading, Attention and Performance XII* (pp. 421-447). UK, Hove: Lawrence Erlbaum.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In: P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 145-174). Hillsdale, NJ: Lawrence Erlbaum.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: evidence from backward masking and priming. *Journal of Memory and Language, 30*, 473-485.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 633-643.

Perin, D. (1982). Spelling strategies in good and poor readers. *Applied Psycholinguistics, 3*, 1-14.

Phaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: a connectionist model for attention in visual selection tasks. *Cognitive Psychology, 22*, 273-341.

Posnansky, C. J., & Rayner, K. (1977). Visual-feature and response components in a picture-word interference task with beginning and skilled readers. *Journal of Experimental Child Psychology, 24*, 440-460

Posnansky, C. J., & Rayner, K. (1978). Visual vs. phonemic contributions to the importance of the initial letter in word identification. *Bulletin of the Psychonomic Society, 11*, 188-190

Raven, J. C. (1958). *Standard Progressive Matrices*. London: H. K. Lewis.

Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading*. Englewood Cliffs, NJ: Prentice-Hall.

Rayner, K., & Posnansky (1978). Stages of processing in word identification. *Journal of Experimental Psychology: General, 1*, 64-80.

Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition, 24*, 31-44.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*, 275-280.

Reitsma, P. (1983a). Word-specific knowledge in beginning reading. *Journal of Research in Reading, 6*, 41-56.

Reitsma, P. (1983b). *Phonemic and graphemic codes in learning to read*. Doctoral dissertation. Free University of Amsterdam, Amsterdam, The Netherlands.

Reitsma, P. (1984). Sound priming in beginning readers. *Child Development, 55*, 406-423.

Reitsma, P. (1985). De verwerving van orthografische kennis. In E. Assink en G. Verhoeven (Eds.), *Visies op spelling* (pp. 73-82). The Netherlands, Groningen: Wolters-Noordhoff.

Reitsma, P. (1988). Tussen letter en woord: het effect van oefening met wisselrijen. *Pedagogische Studiën, 65*, 344-357.

Reitsma, P. (1989). Het effect van wisselrijen in een vroege fase van het leren lezen. *Pedagogische Studiën, 66*, 421-428

Reitsma, P. (1990). Development of orthographic knowledge. In P. Reitsma & L. Verhoeven (Eds.), *Acquisition of Reading in Dutc*h (pp. 43-64). The Netherlands, Dordrecht: Foris Publications.

Reitsma, P., & Van Daal, V. (1989). Effecten van sequentiële redundantie in de eerste fasen van het leren lezen. *Tijdschrift voor Taalbeheersing, 11*, 204-214.

Reitsma, P., & Verhoeven, L. (1990). Acquisition of written Dutch: an introduction. In P. Reitsma & L. Verhoeven (Eds.), *Acquisition of Reading in Dutc*h (pp. 1-13). The Netherlands, Dordrecht: Foris Publications.

Reitsma, P., & Vinke, J. (1986). Individuele verschillen in direct en indirect lezen. In P. Reitsma, A. G. Bus, & W. H. J. Van Bon (Eds.), *Leren lezen en spellen: ontwikkeling en problemen* (pp. 80-85). The Netherlands, Lisse: Swets & Zeitlinger.

Ridderinkhof, K. R. (1993). Interference from irrelevant information. Doctoral dissertation. University of Amsterdam, The Netherlands.

Roberts, K. T., & Ehri, L. C. (1983). Effects of two types of letter rehearsal on word memory in skilled and less skilled beginning readers. *Contemporary Educational Psychology, 8*, 375-390.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. American Psychologist, 45, 1043-1056.

Roediger, H. L., & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory & Cognition, 15*, 379-388.

Roediger, H. L., Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Lawrence Erlbaum.

Rolf, P. C., & Van Rijnsoever, R. J. (1984). *Positionele letterfrekwenties van het Nederlands*. The Netherlands, Lisse: Swets & Zeitlinger.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development, 46*, 247-253.

Rossmeissl, P. G., & Theios, J. (1982). Identification and pronunciation effects in a verbal reaction time task for words, pseudowords, and letters. *Memory & Cognition, 10*, 443-450.

Rozin, P., & Gleitman, L. R. (1977). The structure and acquisition of reading II: The reading process and the acquisition of the alphabetical principle. In A. S. Reber & D. L. Scarborough (Eds.), *Towards a Psychology of Reading* (pp. 55-141). Hillsdale, NJ: Lawrence Erlbaum.

Rubinstein, H., Lewis, S. S., & Rubinstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10*, 645-657.

Schadler, M., & Thissen, D. M. (1981). The development of automatic word recognition and reading skill. *Memory & Cognition, 9*, 132-141.

Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman & D. R. Davies (Eds.) *Varieties of attention* (pp. 1-27). Orlando, Fl.: Academic Press.

Schrijnemakers, C. (to appear). The storage of newly learned information in semantic memory. Doctoral dissertation. University of Nijmegen, The Netherlands

Sears, N. C., & Johnson, D. M. (1986). The effects of visual imagery on spelling performance and retention among elementary students. *Journal of Educational Research, 79*, 230-233.

Seegers, G., Feenstra, H., & Hooft, M. L. (1984). Automatische woordherkenning en verschillen in leesvaardigheid. In A. J. W. M. Thomassen, L. G. M. Noordman, & P. A. T. M. Eling (Eds.) *Het Leesproces* (pp. 191-204). The Netherlands, Lisse: Swets & Zeitlinger.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

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

Seymour, P. H. K., & Porpodas, C. D. (1980). Lexical and non-lexical processing of spelling in dyslexia. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 443-473). London: Academic Press.

Simon, D. P., & Simon, H. A. (1973). Alternative uses of phonemic information in spelling. *Review of Educational Research, 43*, 115-137.

Simpson, G. B., & Foster, M. R. (1986). Lexical ambiguity and children's word recognition. *Developmental Psychology, 22*, 147-154.

Sloboda, J. A. (1980). Visual imagery and individual differences in spelling. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 231-248). London: Academic Press.

Smith, F. (1973). Alphabetic writing - A language compromise? In F. Smith (Ed.), *Psycholinguistics and Reading* (pp. 116-130). N.Y.: Holt, Rinehart and Winston.

Stone, G. O. (in press). Combining connectionist and symbolic properties in a single process. In R. Corregan, G. Iverson, & S. Lima (Eds.), *The Reality of Linguistic Rules*.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Supramaniam, S. (1983). Proofreading errors in good and poor readers. *Journal of Experimental Child Psychology, 36*, 68-80.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition, 20*, 303-313.

Tenney, Y. J. (1980). Visual factors in spelling. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 215-229). London: Academic Press.

Theloosen, G., & Van Bon, W. H. J. (1993). Herhaling van woorden en lettergroepen: oefeneffecten bij zwakke lezers. *Pedagogische Studiën, 70*, 180-194.

Treiman, R. (1985). Phonemic awareness and spelling: children's judgements do not always agree with adults'. *Journal of Experimental Child Psychology, 39*, 182-201.

Uhry, J. K. , & Shepherd, M. J. (1993). Segmentation/spelling instruction as part of a first-grade reading program: effects on several measures of reading. *Reading Research Quarterly, 28,* 218-233.

Van Bon, W. H. J., & Bouwens, F. G. M. (1990). Brugklassers met discordante lees- en spelprestaties. *Pedagogische Studiën, 67*, 317-325.

Van Daal, V. H. P., Bakker, N. C. M., & Reitsma, P. (1986) Woordfrequentie, repetitie en ernstige leesproblemen. In P. Reitsma, A. G. Bus, & W. H. J. Van Bon (Eds.), *Leren lezen en spellen: ontwikkeling en problemen* (pp. 40-45). The Netherlands, Lisse: Swets & Zeitlinger.

Van Dale (1984). *Groot woordenboek Engels-Nederlands.* The Netherlandas, Utrecht: Van Dale Lexicografie.

Van Doorn-Van Eijsden, M. (1984). Leer je spellen door veel te lezen? *Tijdschrift voor taalbeheersing, 6*, 252-263.

Van Heuven, V. J. (1980). Aspects of Dutch orthography and reading. In J. F. Kavanagh & R. L. Venezky, *Orthography, Reading, and Dyslexia* (pp. 57-73). Baltimore: University Park Press.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15*, 181-198.

Van Orden, G. C. (1991). Phonologic mediation is fundamental to reading. In: D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 77-103). Hillsdale, N.J.: Erlbaum.

Van Orden, G. C., & Goldinger, S. D. (in press a). The interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance*.

Van Orden, G. C., & Goldinger, S. D. (in press b). The interdependence of form and function in cognitive systems explains phonologic mediation in reading. In P. Tallal (Ed.), *The exceptional brain* (Vol. II): Neural and cognitive mechanisms underlying speech, language and reading.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 371-386.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97*, 488-522.

Van Orden, G. C., Stone, G. O., Garlington, K. L., Markson, L. R., Pinnt, G. S., Simonfy, C. M., & Brichetto, T. (1992). "Assembled" phonology and reading: a case study in how theoretical perspective shapes empirical investigation. In: R. Frost and L. Katz (Eds.), *Orthography, Phonology, Morphology, and Meaning* (pp. 249-292). The Netherlands, Amsterdam: Elsevier Science Publishers B.V.

Van Rijnsoever, R. J. (1988). Spelling en leren lezen: schijnbewegingen van orthografische effecten in de ontwikkeling van de leesvaardigheid (Spelling and learning to read: feints of orthographic effects in the development of the reading ability). Doctoral dissertation. University of Nijmegen, The Netherlands.

Vodegel Matzen, L. B. L., Van der Molen, M. W., & Dudink, A. C. M. (1994). Error analysis of Raven test performance. *Personality and Individual differences, 16(3)*, 433-445.

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

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

Weber, R. M. (1970). A linguistic analysis of first-grade reading errors. *Reading Research Quarterly, 5*, 427-451.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, 1, 59-85.

Wike, E. L., & Church, J. D. (1976). Comments on Clark's "The language-as-fixed-effects fallacy." *Journal of Verbal Learning and Verbal Behavior, 15*, 249-255

Wydell, T. N., Patterson, K. E., & Humphreys, G. W. (1993). Phonologically mediated access to meaning for Kanji: Is a *Rows* still a *Rose* in Japanese Kanji? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 491-514.

11

APPENDICES

Appendix A. Stimuli used in Experiment 2.1.

Training word	Pseudohomophone	Non-training word
pauw (peacock)	pouw	schat (treasure)
schep (scoop)	sgep	schaar (scissors)
tijger (tiger)	teiger	schoen (shoe)
schoot (lap)	sgoot	schuim (foam)
schuur (shed)	sguur	schrift (notebook)
augurk (gherkin)	ougurk	auto (car)
schort (apron)	sgort	vouw (crease)
paleis (palace)	palijs	mouw (sleeve)
eiland (island)	ijlant	gebouw (building)
schuld (debt)	sgult	kauwgom (chewing gum)
oerwoud (jungle)	oerwaut	vijver (pond)
goudvis (goldfish)	gautvis	bedtijd (bed time)
nijlpaard (hippopotamus)	neilpaart	glijbaan (slide)
stouterd (naughty boy/girl)	stautert	wijsneus (know-all)
wedstrijd (match)	wetstreit	breinaald (knitting needle)

Appendix B. Stimuli used in Experiment 2.2.

Training word	Pseudohomophone
schuim (foam)	sguim
scheur (crack)	sgeur
augurk (gherkin)	ougurk
auto (car)	outo
schuld (guilt)	sgult
stouterd (naughty boy/girl)	stautert
oerwoud (jungle)	oerwaut
aardbei (strawberry)	aartbij
nijlpaard (hippopotamus)	neilpaart
breinaald (knitting needle)	brijnaalt
schildpad (tortoise)	sgiltpat
wedstrijd (match)	wetstreit

Appendix C. Stimuli used in Experiment 2.3.

Training word	Pseudohomophone
jouw (your)	jauw
wijzen (to point)	weizun
meisje (girl)	mijsju
dichtbij (close by)	digtbei
handdoek (towel)	hantdoek
vouwblad (folder)	vauwblat
altijd (always)	alteit
glijbaan (slide)	gleibaan

tijger (tiger)	teigur
zacht (soft)	zagt
rondje (round/ lap circuit)	rontju
lijmen (to glue)	leimun
stouterd (naughty boy/girl)	stautert
nacht (night)	nagt
einde (end)	ijnde
bladzijde (page)	blatzeide
flauw (silly/unsporting)	flouw
krijgen (to get)	kreigun
optocht (parade)	optogt
paleis (palace)	palijs

Appendix D. Stimuli used in the Experiment 3.1.

strand (beach)	kauwen (to chew)
bedtijd (bedtime)	blauw (blue)
zuchten (to sigh)	schildpad (tortoise)
piano (piano)	ijskoud (ice-cold)
bouwdoos (box of building blocks)	hoofdpijn (headache)
stoplicht (traffic light)	schuld (blame)

Appendix E. Stimuli used in Experiments 4.1 and 4.2.

Easy condition	Difficult condition
reis (journey)	strand (beach)
zoen (kiss)	kauwen (to chew)
vlag (flag)	ijskoud (ice-cold)
lijm (glue)	bedtijd (bedtime)
fruit (fruit)	zuchten (to sigh)
zeil (sail)	schuld (blame)
ziek (ill)	blauw (blue)
vies (dirty)	bouwdoos (box of building blocks)
rijst (rice)	hoofdpijn (headache)
fee (fairy)	stoplicht (traffic light)

Appendix F. Spelling test used in Experiments 4.1 and 4.2. Target words are underlined.

Er zit een gat (hole) in mijn broek.	Brrr.... wat een kou (cold).
Die stapel lijkt een hele boel (a lot).	s'Avonds schijnt de maan (moon).
Er zit een lek (leak) in de kraan.	Dit bord is schoon (clean).
Die jongen heet Guus (Christian name).	In de boom zit een duif (pigeon).
Ik buk (stoop) om de pen van de grond te rapen.	Het been (leg) van Jan doet zeer.
Deze zin (sentence) is heel lang.	Het meisje eet een ei (egg).
Op de vensterbank staat een pot (pot)	Ik heb vijf (five) vingers aan mijn hand.
Hij heeft een bril op zijn neus (nose).	Het pakje valt (falls) van de stoel.
Een poes lust graag melk (milk).	De agent zegt: stop (stop).
Deze banaan is krom (bent).	Dit is een hele zware asbak (ashtray).

Appendix G. Stimuli used in Experiment 5.1.

Words	Legal nonwords	Illegal nonwords
boog (bow)	beeg	bgoo
boek (book)	biek	bkoe
doos (box)	dees	dsoo
deur (door)	doer	dseu
hout (wood)	heut	htou
haas (hare)	huus	hsaa
koek (cake)	keuk	kkoe
kees (Christian name)	kuus	ksee
loop (walk)	laap	lpoo
lief (sweet)	leuf	lfie
maan (moon)	moen	mnaa
miep (Christian name)	maap	mpie
noot (nut)	noet	ntoo
neus (nose)	nees	nseu
poes (cat)	peus	psoe
paal (post/pole)	puul	pkaa
raam (window)	reem	rmaa
reus (giant)	ries	rseu
voet (foot)	voot	vtoe
vuur (fire)	veur	vsuu
wieg (cradle)	woeg	wgie
weeg (weigh)	wuug	wgee
zout (salt)	zuut	ztou
zeef (sieve)	zief	zfee

Appendix H. Stimuli used in Experiment 5.2.

Training Word	Legal PsH	Illegal PsH	Legal NW	Illegal NW	Known Word
brief (letter)	broot	briev	blark	bdral	boek (book)
brood (bread)	draat	dichd	droos	dkril	doos (box)
dicht (closed)	fluid	fietz	fliem	ftsie	fik (fire)
draad (thread)	gelt	grijz	glap	gsla	gaat (goes)
fiets (bike)	hont	huiz	harf	hdem	hoort (hears)
fluit (flute)	klij	kalv	karp	krti	kook (cook)
geld (money)	lant	leech	lons	lmsi	loopt (walks)
grijs (grey)	mont	menz	murg	mstu	maan (moon)
hond (dog)	naalt	nachd	nielt	ntlei	niest (sneezes)
huis (house)	pland	paarz	pralk	prkli	paal (post)
kalf (calf)	rant	rupz	rems	rmsa	rookt (smokes)
klei (clay)	stoud	slurv	stril	sklri	soep (soup)
land (land)	vrei	vijv	vom	vme	veegt (sweeps)
leeg (empty)	zant	zalv	zulf	zfli	zeeft (sieves)
mens (person)					
mond (mouth)					
naald (needle)					
nacht (night)					
paars (purple)					
plant (plant)					
rand (edge/rim)					
rups (caterpillar)					
slurf (trunk)					
stout(naughty)					
vijf (five)					
vrij (free)					

zalf (salve)
zand (sand)

PsH is pseudohomophone; NW is nonword

Appendix I. Stimuli used in Experiment 5.3.

Single cluster	Double cluster	Mixed cluster
ant	aab	auf
arg	aat	aul
ast	aaf	aud
arp	aam	aup
asp	aad	auk
elg	eeb	eum
ers	eep	euf
eps	eek	eul
erp	ees	eug
est	eeg	euk
ost	oos	out
orp	oob	ouf
olf	oop	ouk
org	oof	oup
orf	ool	oul
uls	uuf	uik
urp	uum	uip
ust	uud	uis
urg	uun	uif
urm	uut	uid

Appendix J. Stimuli used in Experiments 6.1 and 6.2.

Word	Pseudohomophone	Spelling control
bij (near/bee)	bei	bijn
kijkt (looks)	keikt	kijnt
zijn (to be)	zein	zijm
weg (road)	wech	wag
zout (salt)	zaut	zolt
klein (small)	klijn	krein
voet (foot)	foet	voot
bijl (ax)	beil	bijk
pijp (pipe)	peip	pijg
hout (wood)	haut	huut
bijt (bites)	beit	bijf
geeft (gives)	geevt	geekt

Appendix K. Stimuli used in Experiments 6.3 and 6.4.

Target stimulus	Pseudohomophone	Spelling control
wijn (wine)	wein	wijg
zout (salt)	zaut	zoum
tand (tooth)	tant	tald
acht (eight)	agt	aft
bij (near/bee)	bei	bijf
krijt (chalk)	kreit	krijl
kind (child)	kint	kild
geit (goat)	gijt	geim
kous (stocking)	kaus	koum

grijs (gray)	greis	grijk

Appendix L. Stimuli used in Experiment 7.1.

Training word	Incorrect alternative used in spelling test
vouwen (to fold)	vauwen
stout (naught)	staut
blauw (blue)	blouw
kauwen (to chew)	kouwen
spijker (nail)	speiker
tijger (tiger)	teiger
klei (clay)	klij
paleis (palace)	palijs
trein (train)	trijn
kwaad (angry)	kwaat
vriend (friend)	vrient
staart (tail)	staard
sloot (ditch)	slood
zacht (soft)	zagt
zuchten (to sigh)	zugten
slecht (bad)	slegt
vegen (to sweep)	vechen
slager (butcher)	slacher

Appendix M. Stimuli used in Experiment 7.2.

In the first and third columns the spelling of the pseudowords as it was instructed in the spelling training. In the second and fourth columns the alternative spellings as used in the forced-choice spelling test.

With an ambiguous spelling cluster		Without an ambiguous spelling cluster	
bassiend	bassient	abeleun	abeleen
bodufijp	bodufeip	bodderas	boddelas
kannieme	kannime	gateling	gatelin
kapplond	kapplont	hovelaar	havalaar
kloumine	klaumine	jardaal	jardaan
kraplout	kraplaut	karpaan	karpan
lagootild	lagootilt	kenelgat	kanelgat
leciene	lecine	kompeur	kempeur
pameint	pamijnt	lenkerof	lenkenof
peifelons	pijfelons	mantoor	mantaar
plauzum	plouzum	nortoom	norteem
ribeluk	riebeluk	petanuis	petanees
schapijt	schapeit	rekulans	rekulant
steimerd	stijmerd	slaakhuif	slaakheuf
uitpijder	uitpeider	sokaleur	sakaleur

Appendix N. Stimuli used in Experiment 7.3.

blij (happy)	mauwt (miaow)	toch (still/yet)	huis (house)	voer (sailed)
krijg (get)	gauw (quickly)	groot (large	duik (dive)	zoen (kiss)
kwijt (lost)	stout (naughty)	vlug (fast)	buis (tube)	vaar (sail)

fijn (nice)	koud (cold)	veegt (sweeps)	sluipt (sneaks)	zeep (soap)
pijn (pain)	goud (gold)	schoen (shoe)	trui (swater)	zes (six)
klein (small)	pauw (peacock)	dicht (closed)	neus (nose)	bang (afraid)
wijs (wise)	touw (rope)	vroeg (early)	reus (giant)	ring (ring)
bijt (bite)	jou (you)	lied (song)	poes (cat)	duw (push)
prei (leek)	hout (wood)	hard (hard)	moe (tired)	zacht (soft)
plein (square)	nou (now)	duif (pigeon)	koek (cooky)	peer (pear)

Appendix O. Stimuli used in Experiment 7.4.

The first word of each stimulus pair is the word's proper spelling and the second the pseudohomophone version of the word as it was presented in the forced-choice test.

steigeren	stijgeren	twijfelen	tweifelen
ammonia	amonia	parallel	paralel
vrouwenstem	vrouwestem	mannenstem	mannestem
dynamiet	dinamiet	synoniem	sinoniem
pelzen	pelsen	halzen	halsen
feeëriek	feëriek	tweeërlei	tweërlei
hysterie	histerie	mysterie	misterie
carrosserie	carosserie	interruptie	interuptie
huwelijk	huwlijk	afschuwelijk	afschuwlijk
polsen	polzen	belevenis	belefenis
bochel	boggel	rochelen	roggelen
fascineren	fassineren	discipline	dicipline
serviezen	serviesen	pyjama	pijama
oeuvre	ouvre	paleizen	paleisen
bisschop	bischop	manoeuvre	manouvre
faillissement	faillisement	misschien	mischien
loochenen	loogenen	applaudisseren	applaudiseren
lineair	liniair	murw	murf
scenario	cenario	goochelaar	gogelaar
wollige	wolige	alinea	alinia
abonnee	abonee	scène	scene
koloniën	kolonieën	mollige	molige
lijzijde	leizijde	abattoir	abbatoir
hartgrondig	hardgrondig	bacteriën	bacterieën
enigszins	enigzins	rijstebrij	rijstebrei
beambte	beamte	asterisk	asteriks
miljard	milliard	envelop	envellop
affaire	afaire	geenszins	geenzins
origineel	orgineel	symptoom	syntoom
nervositeit	nerveusiteit	miljoen	millioen
pincet	pinset	affiche	affische
heiig	heïg	tafereel	tafreel
compromitteren	compromiteren	klauteren	klouteren
luiwammesen	luiwammessen	buiig	buiïg
filiaal	fileaal	perceel	perseel
parafen	paraven	permitteren	permiteren
koningschap	koningsschap	dreumesen	dreumessen
leeuweriken	leeuwerikken	liniaal	lineaal
wroeten	vroeten	paragrafen	paragraven
rapport	raport	louter	lauter
weifelen	wijfelen	monniken	monnikken
takkenbos	takkebos	verleidelijk	verlijdelijk
tiranniseren	tiraniseren	apparaat	aparaat
giechelen	giegelen	weigeren	wijgeren
statieven	statiefen	sterrenbeeld	sterrebeeld
uitweiden	uitwijden	galopperen	galloperen
behaaglijk	behaagelijk	choreografie	goreografie
onmiddellijk	onmiddelijk	octaven	octafen

melodieën	melodiën	vibreren	fibreren
appelleren	appeleren	bedrieglijk	bedriegelijk
erfenis	ervenis	herinneren	herineren
aanvankelijk	aanvanklijk	strategieën	strategiën
plakkaat	plakaat	duelleren	dueleren
gêne	gène	begrafenis	begravenis
religieuze	religieuse	aanhankelijk	aanhanklijk
peloton	peleton	pakket	paket
arrogant	arogant	nerveuze	nerveuse
attribuut	atribuut	grammofoon	grammafoon
portugese	portugeese	arrestant	arestant
sterrenkunde	sterrekunde	attractie	atractie
barak	barrak	overzeese	overzese
krokussen	krokusen	hondehok	hondenhok
glooiing	glooiïng	karikatuur	karrikatuur
millibar	milibar	notarissen	notarisen

Appendix P. Anecdotal evidence for the fundamental role of phonology in reading and spelling.

In his article "A spreading activation theory of memory" (1983) in the Journal of Verbal Learning and Verbal Behavior, 22, 261-295, John Anderson refers to an article by him and his colleague Gordon Bower as (Anderson & Bauer, 1973; p. 292).
(Detected by René Zeelenberg)

The role of phonology is also apparent in word translation. In an experimental task Dutch-speaking subjects had to translate English words, presented visually on a computer screen, into their native Dutch language. Tables 11.1 and 11.2 list examples which suggest that phonology was involved (collected by Annette de Groot and Janet van Hell).

It shows from Table 11.1 that in some cases the subjects activated a homophone word* of the target stimulus; in other cases they coded the phonology of the target word incorrectly#.

Table 11.1

Stimulus	(incorrect) Response	via Phonology of	Proper Dutch translation
AUNT#	mier	/ant/	tante
BREAD#	fokken	/breed/	brood
CHANCE#	wisselgeld	/change/	kans
DESCENT#	netjes	/decent/	daling
DUKE#	eend	/duck/	hertog
FAITH#	noodlot	/fate/	geloof
FATE#	geloof	/faith/	noodlot
HEAT#	hoofd	/head/	hitte
INSIGHT#	binnenin	/inside/	inzicht
KNIGHT*	nacht	/night/	ridder
LAMB#	lamp	/lamp/	lam
MAIL*	mannelijk	/male/	post
NIGHT*	ridder	/knight/	nacht
PEACE*	stuk	/piece/	vrede

PEAR#	paar	/pair/	peer
PIECE*	vrede	/peace/	stuk
SHEEP#	goedkoop	/cheap/	schaap
SWEAT#	zoet	/sweet/	zweet
THREAD#	dreiging	/threat/	draad

Table 11.2 shows examples of English words, some of which were interlexical homographs*, that received incorrect (Dutch) phonologic coding. In a few cases they gave rise to a Dutch synonym via the Dutch phonologic coding of the target#. Other examples show that some of the target words were mistakenly treated as interlexical homographs.

Table 11.2

Stimulus	(incorrect) Response	Meaning of response	Proper Dutch translation
BAKER*	vroedvrouw#	midwife	bakker
BID*	bid	pray	bod
BULL*	diploma#	certificate	stier
DOLL	gek#	crazy	pop
LOAN	loon	salary	lening
MEANING	mening	opinion	betekenis
RIDDLE	riedel	jingle/tune	raadsel
ROOF*	roof	robbery	dak
RUMOUR	rumoer	noise	gerucht
SIN	zin	sentence	zonde
SON	zon	sun	zoon
TEAR	teer	tar	traan
WALL	wal	(em)bank(ment)	muur

<u>Slips of the pen (= key on the computer keyboard)</u>

Typo's that were registered by the author of this dissertation. Some of them seem to be caused by interference due to the application of Dutch grapheme-phoneme correspondence rules*, and others result from a sloppy English pronunciation (i.e., phonologic coding)#.

Actualised	Intended
e-meel*	e-mail
oll*	all
do (is also the name of the author's mother)#	though
frough#	through
bove#	both
vor*	for
...very in price	...vary in price
the cause of which...	the course of which..
foor*	for
fird#	third
de*	the
a*	I

12

SAMENVATTING (SUMMARY IN DUTCH)

Dit proefschrift is de neerslag van experimenteel onderzoek naar het lees- en spellingproces van zowel beginnende als ervaren lezers en spellers. Bij de aanvang van het onderzoeksproject werd een cognitivistisch model, het 'twee-route-model', als theoretisch uitgangspunt genomen. De belangrijkste assumptie van het twee-route-model is dat er twee van elkaar onafhankelijke processen (routes) zijn die een geletterde ten dienste staan bij het lezen van een woord. Er wordt aangenomen dat beginnende lezers een woord eerst fonologisch moeten decoderen door successievelijk alle grafemen om te zetten in hun fonemen, alvorens zij tot betekenisactivatie kunnen komen. Deze manier van lezen wordt ook wel fonologisch gemedieerd lezen genoemd (of indirect lezen). Met toenemende leeservaring zou dit proces van fonologische mediatie echter niet meer nodig zijn, omdat een lezer dan van een groot aantal woorden een orthografische representatie in zijn mentale lexicon heeft opgebouwd. Het lezen van een woord gebeurt dan door het maken van een match tussen het gedrukte woord en de geheugenrepresentatie, zonder tussenkomst van een fonologisch proces. Dit proces wordt ook wel aangeduid met directe lexicale toegang (directe route). Het essentiële verschil tussen beide processen is dus de rol van de fonologie. In het eerste geval speelt de fonologie een fundamentele rol bij het activeren van de betekenis van een woord, terwijl in het tweede geval de fonologie niet fungeert als mediator tussen woord en betekenisactivatie. Een tweede belangrijke assumptie van dit model is dat de indirecte- of verklankingsroute relatief traag is vergeleken met de directe route.

De onderzoeksvraag waarmee dit onderzoeksproject startte was: hoe snel gaan beginnende lezers over van een indirecte (fonologische) leeswijze naar een directe. In Hoofdstuk 2 worden drie experimenten beschreven waarin geprobeerd wordt een antwoord op deze vraag te krijgen. Hiertoe moesten beginnende lezers (Groep 3 van het Basisonderwijs) nieuwe woorden leren lezen. Een aantal van deze woorden lazen ze relatief vaak (hoog-frequente conditie) en een aantal lazen zij minder vaak (laag-frequente conditie). Er werd verwacht dat de kinderen na een zeker aantal presentaties van een woord, met name in de hoog-frequente conditie, in staat waren om deze woorden zonder verklanking te lezen, omdat ze van deze woorden een orthografische representatie hadden opgebouwd die zij direct zouden kunnen consulteren. Om na te gaan na hoeveel aanbiedingen van een woord er een orthografische representatie was ontstaan werden in een testfase zowel de nieuw-geleerde woorden als daarvan afgeleide 'pseudohomofonen' aangeboden. Een pseudohomofoon is een woord dat qua klank identiek is aan een bestaand woord, maar dat in zijn spelling daarvan afwijkt, bijvoorbeeld, het woord "stautert" is een pseudohomofoon van het woord 'stouterd'. Omdat van pseudohomofonen verondersteld wordt dat zij geen representatie in het mentale lexicon hebben, zullen deze dus gelezen moeten worden via de langzame indirecte route. De verwachting was dus dat bij een gering aantal aanbiedingen van een woord geen opleestijdverschil zou ontstaan tussen woorden en pseudohomofonen, maar dat na een bepaald groter aantal aanbiedingen het woord sneller

opgelezen zou worden dan het bijbehorende pseudohomofoon. Met andere woorden, er werd een interactie verwacht tussen de factoren frequentie van aanbieding en type woord (woord vs. pseudohomofoon). De interactie zou dus het aantal aanbiedingen van een woord noodzakelijk voor de opbouw van een orthografische representatie aangeven. Er werden drie experimenten uitgevoerd waarin de variabele frequentie werd gemanipuleerd, maar niet één keer werd er een significante interactie gevonden, zelfs niet na 30 aanbiedingen van een woord. Het was duidelijk dat herhaald aanbieden van een woord de opleestijd verkortte, maar de opleestijd van het corresponderende pseudohomofoon werd ook korter. Het leek er dus op dat er transfer had plaatsgevonden van het woord naar het pseudohomofoon. Dit is niet verrassend, omdat de twee typen stimuli, weliswaar verschillend van elkaar, toch een zekere mate van overlap vertonen wat betreft hun spelling. Kort gezegd, uit de resultaten van Hoofdstuk 2 bleek dat er iets van een representatie was opgebouwd, maar de aard ervan was nog onduidelijk.

Een tweede vraag die in Hoofdstuk 2 centraal stond was in hoeverre beginnende lezers en spellers de door lezen opgebouwde orthografische representatie inzetten bij het spellen. Omdat het antwoord op de eerste vraag - na hoeveel aanbiedingen van een woord heeft een lezer een orthografische representatie opgebouwd - uitbleef was het niet mogelijk om deze tweede vraag te beantwoorden. Het was echter wel duidelijk dat het veelvuldig lezen van een woord niet tot gevolg had dat de kinderen deze woorden ook beter spelden. Tevens bleek dat het spellingproces van deze kinderen duidelijk sporen van fonologische mediatie vertoonde.

Uit de resultaten van Hoofdstuk 2 bleek dus dat lezen geen effectieve manier is om de spelling van woorden te leren. De vraag die hieruit volgde was of deze kinderen de spelling van de aangeboden woorden wel zouden kunnen leren met behulp van een andere instructiemethode dan lezen. Daartoe werd een spellingtraining experiment uitgevoerd, waarvan verslag is gedaan in Hoofdstuk 3. Naast lezen als instructiemethode werden drie niet-lees spellinginstructiemethoden, 'overschrijven', 'probleem noemen', en 'mondeling spellen' onderzocht op hun effectiviteit. Opnieuw bleek de geringe effectiviteit van lezen voor het spellen, en van de drie niet-lees instructiemethoden bleek mondeling spellen de meest effectieve.

In Hoofdstuk 4 stond de aard van het spellingproces centraal. Opnieuw werd een aantal spellinginstructiemethoden met elkaar vergeleken. Nu werd niet alleen de effectiviteit van de methoden nagegaan, maar werd tevens onderzocht of er instructiemethoden zijn die minder aanleiding geven om het spellingproces fonologisch te laten verlopen. Analoog aan het leesproces werd verondersteld dat er voor het spellen twee wegen zijn om de spelling van een woord te genereren. Een directe route, waarbij het orthografische patroon uit het geheugen wordt gelezen, en een indirecte route waarbij fonemen omgezet worden in grafemen. Uit de resultaten van de twee experimenten bleek heel duidelijk dat alle spellinginstructiemethoden uitgevoerd bij kinderen uit Groep 3, een fonologische spellingstrategie induceerden, en geen aanleiding gaven om een directe spellingprocedure te gebruiken.

Voorlopig was er dus geen indicatie dat het lees- en spellingproces van beginnende lezers uit Groep 3 zich ontwikkelt van een proces waarin aanvankelijk een hoofdrol voor de fonologie is weggelegd naar een proces waarin de fonologie geen of een ondergeschikte rol speelt. Aan het eind van Hoofdstuk 2 werd verondersteld dat het 'pseudohomofoon'

paradigma dat gebruikt werd mogelijk niet geschikt was om de overgang van indirect naar direct lezen te onderzoeken. Daarom werd in Hoofdstuk 5 een ander paradigma, de eerste-letter-taak, gebruikt.

In de oorspronkelijke versie van deze taak worden proefpersonen geconfronteerd met woorden, orthografisch legale niet-woorden, en orthografisch illegale niet-woorden. De taak aan de proefpersoon is om zo snel mogelijk de eerste letter van elke aangeboden stimulus te noemen. De benoeming van de eerste letter van woorden en orthografisch legale niet-woorden blijkt sneller te zijn dan van orthografisch illegale niet-woorden. De aanvankelijke interpretatie van dit effect was geformuleerd in termen van contextuele facilitatie. Wanneer ervaren lezers geconfronteerd worden met woorden dan verwerken zij de letters in een woord op parallele wijze. Tevens kunnen zij de orthografische context gebruiken als een onafhankelijke informatiebron om de identiteit van de letters, dus ook van de eerste letter, te bepalen. Dit maakt het ervaren lezers mogelijk om de eerste letter van legale letterreeksen (woorden en orthografisch legale niet-woorden) sneller te bepalen dan van orthografisch illegale niet-woorden. Het eerste-letter-effect is ook een aanwijzing dat ervaren lezers directe toegang tot het lexicon hebben. Als ze immers de woorden volgens de indirecte route zouden lezen, dan zou de orthografische context geen effect hebben op de benoemtijd van de eerste letter, omdat dit process van links naar rechts verloopt.

Als het eerste-letter-effect een indicatie is van direct lezen, en als tevens beginnende lezers uit Groep 3 nog indirect lezen, dan mag worden verwacht dat zij geen eerste-letter-effect vertonen. Uit de resultaten van Hoofdstuk 5 bleek echter dat beginnende lezers precies hetzelfde effect vertonen als ervaren lezers. Dat wil zeggen, beide groepen benoemden de eerste letter van woorden sneller dan van legale niet-woorden en deze werden weer sneller benoemd dan die van illegale niet-woorden. Hoewel de conclusie dat deze groep beginnende lezers reeds de overgang van indirect naar direct lezen hebben gemaakt voor de hand lag, werd deze toch niet getrokken, omdat er een alternatieve interpretatie van het eerste-letter-effect mogelijk is. Uit de overige experimenten bleek namelijk dat het eerste-letter-effect beter verklaard kan worden in termen van respons competitie dan in termen van contextuele facilitatie. Wanneer een lezer geconfronteerd wordt met een woord, dan wordt het leesproces automatisch gestart. Tevens moet de proefpersoon een proces starten om de eerste letter van het woord te bepalen en te benoemen. Beide processen leiden ertoe dat een respons wordt voorbereid, maar de respons als gevolg van de verwerking van de hele stimulus moet verdrongen worden ten gunste van die van de eerste letter. Omdat automatische verwerking van woorden sneller gaat dan van legale en illegale niet-woorden kan de proefpersoon zich bij woorden sneller op de gevraagde taak richten dan bij niet-woorden, waarmee de verschillen in benoemtijden tussen woorden en niet-woorden verklaard kunnen worden. De experimenten uit Hoofdstuk 5 lieten zien dat het niet nodig is om twee leesroutes te veronderstellen, en dat de gevonden effecten het beste verklaard kunnen worden in termen van een 'één-route-model', zoals geformuleerd door Van Orden, Pennington, and Stone (1990). In dit model is een fundamentele rol weggelegd voor de fonologie.

In Hoofdstuk 6 werd voor het eerst geprobeerd positieve evidentie te vinden bij de rol van de fonologie voor lezen. Eerder onderzoek had laten zien dat ervaren lezers tijdens een tekstcorrectietaak vaker een pseudohomofoon (bijvoorbeeld: peip) missen dan een pseudowoord (in de tekst spelling-controle-woord genoemd) dat geen klankovereenkomst

vertoont met een bestaand woord (bijvoorbeeld: pijg). Dezelfde effecten traden op bij een lexicale-decisie-taak en een semantische-categorisatie-taak. In alle gevallen konden de effecten verklaard worden in termen van fonologische mediatie. Immers, wanneer proefpersonen vaker een pseudohomofoon dan een spelling-controle-woord missen wanneer hen expliciet gevraagd wordt om fout-gespelde woorden op te sporen, dan kan dat alleen maar verklaard worden door aan te nemen dat het pseudohomofoon ten onrechte aanleiding heeft gegeven tot betekenisactivatie. Het enige verschil tussen pseudohomofonen en spelling-controle-woorden is dat er een klankovereenkomst bestaat tussen het pseudohomofoon en een bestaand woord, maar niet tussen een spelling-controle-woord en een bestaand woord.

In Hoofdstuk 6 werden dezelfde drie taken gebruikt bij beginnende lezers. Uit de resultaten van de experimenten bleek dat ook beginnende lezers vaker een pseudohomofoon missen dan een spelling-controle-woord wanneer hen gevraagd wordt fout-gespelde woorden op te zoeken. Een belangrijk verschil tussen beide groepen is dat beginners veel meer pseudohomofonen missen dan ervaren lezers, terwijl het gemiddeld aantal gevonden spelling-controle-woorden nagenoeg gelijk is in beide lezersgroepen. Dit verschil werd verklaard in termen van een minder geautomatiseerd spelling-verificatie-mechanisme bij de beginnende lezers dan bij de ervaren lezers. De belangrijkste conclusie in Hoofdstuk 6 was echter dat een één-route-model, waarin de fonologie een fundamentele rol speelt, niet alleen een zuiniger, maar tevens een even adequate verklaring verschaft voor het leesgedrag van zowel beginnende als ervaren lezers.

Het model dat door Van Orden e.a. (1990) is voorgesteld lijkt dus een veelbelovend alternatief. In Hoofdstuk 7 werd daarom geprobeerd om het twee-route-model te contrasteren met het één-route-model van Van Orden e.a. (1990). Uit beide modellen werden hypotheses afgeleid die betrekking hebben op de relatie tussen lezen en spellen. Uit de assumpties van het twee-route-model kan afgeleid worden dat wanneer lezers een orthografische representatie hebben van een woord, dit woord gelezen kan worden door een directe match van het geschreven woord met die representatie, en dat woorden waarvan geen, of een incorrecte orthografische representatie voorhanden is, gelezen moeten worden via de langzame indirecte route. Uit het één-route-model van Van Orden e.a. kan afgeleid worden dat kennis over de spelling van woorden geen voldoende voorwaarde is voor het vlot kunnen lezen van een woord. De hypothese die uit het één-route-model volgde luidde dan ook dat het oplezen van een woord waarover de proefpersonen spellingkennis bezitten niet sneller gaat dan van woorden waarvan zij geen spellingkennis hebben. Inderdaad lazen zowel beginnende als ervaren lezers woorden waarvan ze de spelling kenden even snel op als woorden waarvan de spelling onbekend, incompleet, dan wel incorrect was. Dit was dus een ondersteuning van het model van Van Orden e.a. (1990). Een tweede hypothese die getoetst werd betrof de aard van het spellingproces. Volgens het twee-route-model zou de verdeling van de proporties fonologisch correcte ('staut' in plaats van 'stout') en fonologische incorrecte fouten ('skout' in plaats van 'stout') in woorden waarover de proefpersoon orthografische kennis heeft verschillend zijn van die van woorden waar deze geen orthografische kennis over heeft. Het één-route-model voorspelt echter geen verschillen in de distributies van type fouten. In alle gevallen zouden de proporties fonologisch correcte fouten de proporties fonologische incorrecte fouten moeten overtreffen. De resultaten van

zowel de beginnende als de ervaren lezers lieten precies dit effect zien. Ook in het spellingproces speelt de fonologie dus een fundamentele rol.

In Hoofdstuk 1 werd het één-route-model van Van Orden en zijn collega's nader uitgewerkt. Het is niet verankerd in een symbolische, cognitivistische theorie zoals het twee-route-model, maar in dynamische-systeem-theorie. In tegenstelling tot de symbolische theorie veronderstelt de dynamische-systeem-theorie geen expliciete representaties waarop regelgestuurde operaties worden uitgevoerd, maar vormen activatiepatronen de belangrijkste pijlers van het model. Er worden verbindingen verondersteld tussen orthografische, fonologische, en semantische subsymbolen, waarbij coherentie tussen de subsymbolen (resonantie) zal ontstaan. De reden dat de fonologie een belangrijke determinant is bij het lezen is dat de volgorde waarin coherentie tussen subsymbolen ontstaat bepaald wordt door de mate van zelf-consistentie tussen de subsymbolen. De relatie tussen orthografische en fonologische subsymbolen is consistenter dan tussen orthografische en semantische subsymbolen, omdat de variatie tussen woorden en hun betekenis (betekenis wordt vaak bepaald door de context) groter is dan tussen woorden en hun uitspraak (deze is meestal uniek, op een paar uitzonderingen na, 'regent', 'kantelen', 'bedelen'). Dit verklaart dus de essentiële rol van de fonologie tijdens lezen.

In Hoofdstuk 8, General Discussion, zijn de resultaten van de experimentele hoofdstukken nog eens op een rijtje gezet, en werd geprobeerd aan te tonen dat het niet nodig is om een model te veronderstellen waarin twee routes naar het lexicon gepostuleerd worden, maar dat een zuiniger één-route-model de voorhanden zijnde feiten kan verklaren.