Journal of Educational Psychology
2006, Vol. 98, No. 4, 879–890

0022-0663/06/$12.00 DOI: 10.1037/0022-0663.98.4.879

Learning the Spelling of Strange Words in Dutch Benefits From Regularized Reading

Anna M. T. Bosman, Janet G. van Hell, and Ludo Verhoeven
Radboud University Nijmegen

In 2 experiments, the authors tested the effect of 2 types of reading on the spelling memory of strange or sound–spelling inconsistent words in Dutch students with and without learning disabilities: standard reading and regularized reading. *Standard reading* refers to reading the word the way it has to be read. *Regularized reading* refers to reading a sound–spelling inconsistent word as if it is sound–spelling consistent. In Experiment 1, both groups showed a short-term effect. Shortly after training, all students who participated in the regularized-reading condition showed better spelling performance than students who took part in a standard-reading condition. One week after training, spelling knowledge of students without learning disabilities appeared to be more stable than that of students with learning disabilities. In Experiment 2, only students with learning disabilities participated. The results reveal that more training substantially enhanced spelling performance in the long term, and repetition of the regularized word was even more effective. A recurrent network account served as the guiding principle to explain the results.

Keywords: learning to spell, strange words, consistency, learning disabilities, feedback models

The beginning of Modern Dutch, a member of the West Germanic language group, is strongly related to the emergence of a standard, written language that came into being in the 16th century and started to show uniformity with Siegenbeek's publication in 1804 on how to write certain spelling variants. It was not until 1883, however, that the Dutch government accepted the official spelling rules described by te Winkel and de Vries (1866). Since then, two more official spelling reforms have taken place, one in 1954 and the most recent one in 1995 (Bosman, de Graaff, & Gijsel, 2006). Unlike the French, who appointed a number of people in L'Académie Française to fight the invasion of foreign words into the French language, the Dutch freely adopted words from languages such as Latin, Greek, Celtic, and Yiddish and still do from English, French, German, Italian, Spanish, Arabic, Hebrew, and others. Presently, the Dutch spoken and written language broadly comprises two types of words: indigenous words and borrowings. Indigenous words are words that already existed when the Germanic language group was still a unity. Borrowings have entered the Dutch language from other languages. Loan words are the most common borrowings in the Dutch language and are often subdivided in naturalized words, bastard words, and strange words.

Naturalized words are morphologically and phonologically identical to indigenous words; the borrowing *kerk* (*church*) is an example of a naturalized word. It strongly resembles the indigenous word *werk* (*work*) but is borrowed from the Greek word *kuriakon.* Similarly, the word *beschuit* (*Dutch biscuit*), which is borrowed from the French word *biscuit*, has the appearance of an indigenous Dutch word as a result of the use of the prefix *be-*. Bastard words are loan words with spelling and phonology adapted to indigenous words; examples are *fitheid*, from *fitness*; *empirisch*, from *empirical*; and *citroen*, from the French word *citron* (*lemon*). Strange words usually retain to a large extent the phonology, morphology, and spelling of the source language; examples are *thriller* and *computer* from English, *grammatica* (*grammar*) from Latin, and *bureau* (*desk*) and *douche* (*shower*) from French. For an unrivalled overview of the effects of other languages on Dutch, we refer the reader to van der Sijs (1996).

With respect to the present study, this last group of words is the object of our investigation. From a computation on the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993), it appears that about 15% of Dutch words are strange words (Bosman et al., 2006; Bosman & Mekking, 2005; Nunn, 1998). Dutch strange words are composed of atypical or inconsistent Dutch phoneme–grapheme relations. Phoneme–grapheme consistency is different from grapheme–phoneme consistency, as shown by Stone, Vanhoy, and Van Orden (1997). In almost all alphabetical languages, sound-to-spelling inconsistency is higher than spelling-to-sound inconsistency (Bosman & Van Orden, 1997; Stone et al., 1997; Ziegler, Jacobs, & Stone, 1996). In other words, there are more possible ways to spell a word than there are possible ways to pronounce a

Anna M. T. Bosman, Janet G. van Hell, and Ludo Verhoeven, Department of Special Education, Faculty of Social Sciences, Radboud University Nijmegen, Nijmegen, the Netherlands.

We thank all teachers and students of De Albatros, Prins Constantijn, and 't Gijmink in Goor; 't Eimink and De Telgenkamp in Hengelo; De Ark, De Herman Broerenschool, De Princenhof, and De Werkhaven in Enschede; Wilhelmina Bladergroenschool in Nijmegen; and Meginhardschool in Arnhem, all in the Netherlands. Without their helpful cooperation, this research would not have been possible. We also thank Ilona Schiffelers, Wendy Exterkate, and Anne-Marie Rosink for training the children and for their help with analyzing the data.

Correspondence concerning this article should be addressed to Anna M. T. Bosman, Department of Special Education, Faculty of Social Sciences, Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, the Netherlands. E-mail: a.bosman@pwo.ru.nl

word, which explains why spelling is often more difficult than reading. Below we illustrate the importance of this phenomenon.

Numerous studies in a variety of languages have shown that the reading of spelling-to-sound inconsistent words is slower and less accurate than the reading of spelling-to-sound consistent words. See, for example, Jared (1997) and Seidenberg, Waters, Barnes, and Tanenhaus (1984) for English; Content (1991) for French; Borzone de Manrique and Signorini (2000) for Spanish; Ziegler, Perry, and Coltheart (2000) for German; and Lukatela, Turvey, and Todorovic (1991) for Serbo-Croatian. Research into the question of whether sound-to-spelling inconsistent words are also harder to spell than sound-to-spelling consistent words has attracted much less attention.

The first researchers to study this issue were Fischer, Shankweiler, and Liberman (1985). They had American university students spell three types of words. The first set followed prototypical sound–letter couplings: for example, *harp.* The second set of words had spellings that deviated from the prototypical relations; this group comprised a set of words that required the application of orthographic rules and a set that required morphophonemic knowledge—for example, one can determine the spelling of *muscle* by relating the word to *muscular.* The third and final set of words they used were strange or idiosyncratic words—for example, *bourgeois.* These words require rote memorization, which involves good visual memory, or at least that was the assumption. Fischer et al. found that good spellers outperformed poor spellers on all three categories but particularly on the morphophonemic words, which required the application of morphophonemic knowledge. The authors concluded that the major difference between the good and poor spellers resides in the level of linguistic sensitivity rather than good visual memory.

Holmes and Ng (1993) conducted a similar experiment with Australian university students. They also found that good spellers were better on all three types of words, but in their case, the good and poor spellers were best discriminated on the set of strange words and not on the morphophonemic words. This finding caused the authors to conclude that the major difference between good and poor spellers is the ease with which they retain particular letter sequences in (visual) memory. Our goal is not to unravel the reasons for these diverging findings but to present a theoretical explanation for the difficulty of learning the spelling of strange words and to conduct two empirical studies in which a relatively simple strategy is implemented to ease the spelling process of these words.

The scientific literature on reading has a rich tradition in which computational models (e.g., Coltheart, 1978; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) and connectionist approaches (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) have been developed. Unfortunately, the attention for spelling in these models is limited or absent. Separate models for spelling have been proposed by Brown and Loosemore (1994), Olson and Caramazza (1994), and Snowling (1994). An integration of both aspects of literacy in one account was lacking until Bosman and Van Orden (1997) extended Van Orden, Pennington, and Stone's (1990) recurrent network account for reading to spelling. A decade after the model's introduction, Farrar and Van Orden (2001) simulated five benchmark phenomena of intact and dyslexic word naming. We explain the regularization error, the one relevant for our current study, below.

Our decision to take Van Orden et al.'s (1990) recurrent network account as the guiding principle for the current study is not because we believe that it is the only model that provides an adequate explanation for the effects presented in these experiments (see Van Orden & Kloos, 2005) or that it is the only possible instantiation of an attractive set of design principles (i.e., covariant learning and self-consistency; Stone & Van Orden, 1994; Van Orden & Goldinger, 1994). Rather, we use the account because it incorporates both reading and spelling, because it emphasizes the fundamental relation between reading and spelling performance by means of its feedback connections (i.e., it is fully recurrent) among three families of nodes, and because it is firmly grounded in complex systems theory.

Before turning to the empirical part of this study, we present a concise description of Van Orden et al.'s (1990) recurrent network account. The present recurrent network consists of three families of fully interdependent nodes—orthographic nodes, phonologic nodes, and semantic nodes (see Figure 1)—in which input and output form an irreducible interdependent whole. These node families are sufficient to describe reading and spelling performance in a large variety of reading tasks. An essential aspect of the model is that all node families are connected recurrently or bidirectionally. This means that there is a connection from each of the orthographic nodes to each of the phonologic and each of the semantic nodes, that there are backward connections from each of the phonologic and semantic nodes to the orthographic nodes, and that there are similar connections between phonologic and semantic nodes. On presentation of a printed word, the orthographic nodes get activated, and they, in turn, activate the phonologic and semantic nodes (feedforward activation). The recurrent connections cause the phonologic and semantic nodes to activate the orthographic nodes again (feedback activation). Whenever the feedback activation pattern matches the feedforward activation pattern, a temporarily stable, coherent, dynamic whole emerges. Similarly, when the network is presented with a spoken word, phonologic nodes get activated, and they, in turn, activate the orthographic and semantic nodes. Again, the recurrent connections cause the orthographic and semantic nodes to activate the phonologic nodes, and whenever the feedback pattern matches the feedforward pattern, a temporarily stable, coherent, dynamic whole emerges.

An important property of the present network is the difference in overall strength among node families, as illustrated by the relative boldness of the arrows in Figure 1. *Connection strength* indicates

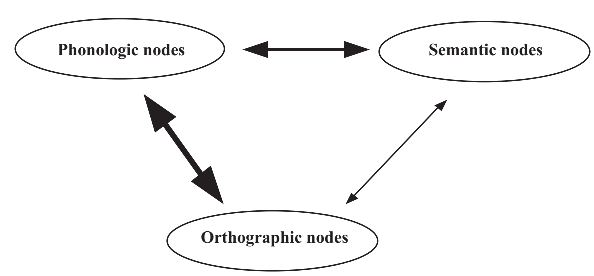

Figure 1. Macrodynamics of a recurrent network account. The thickness of the arrows illustrates the relative strength of the connection between nodes.

the speed with which dynamics cohere. In the present case, the connections between orthographic and phonologic nodes are strongest. In alphabetic languages, letters and phonemes correlate rather strongly. The letter *p* is almost always pronounced as /p/ (exceptions are the *p*s in *phoenix* and *psycho*), and the phoneme /p/ is almost always written with a *p.* The relations between phonemes and semantics are less strong. Knowing that a word starts with the phoneme /p/ does not provide us with much information on its meaning (e.g., *page*, *pure*, and *practical* do not share much meaning). Although phonologic and semantic nodes share only weak connections, these connections are stronger than those between semantic and orthographic nodes. This is primarily because people learn to speak before they learn to read, and they speak more often than they read. Note that before children learn to read, the strongest connections are between phonologic and semantic nodes, but quickly after reading instruction starts, the ones between orthographic and phonologic nodes supersede the other connections as a result of the strong correlations between letters and phonemes.

Thus, this recurrent network predicts that the dynamics involving the relation between orthography and phonology cohere before the dynamics between phonology and semantics, which, in turn, cohere before the dynamics between semantics and orthography. If this is true, it should not surprise us that phonology is an early and omnipresent constraint in reading as well as in spelling. For examples on reading, we refer the reader to Van Orden et al. (1990) and Frost (1998), and for various examples on spelling, we direct the reader to Bosman and Van Orden (1997).

Feedforward networks, pattern associators, and simple recurrent networks (see, e.g., McLeod, Plunkett, & Rolls, 1998), just like our preferred recurrent account, are designed to detect the statistical relations between the input and the output. Any trained network favors consistent relations over inconsistent ones. The difference between feedforward and recurrent or feedback networks is that recurrent networks take into account the asymmetry between spelling-to-sound and sound-to-spelling relations. This characteristic of recurrent models has important implications for reading as well as for spelling. For example, the model predicts that reading should be affected by the relative consistency of phoneme–letter relations. In other words, not only does it matter for reading that a word's spelling may have more than one pronunciation, it also matters that a word's pronunciation may have more than one spelling. Stone et al. (1997) tested this hypothesis using a grain size of spelling to sound that was larger than letters and phonemes—namely, onsets and rhymes (or bodies). The onset of one-syllable words is the initial consonant cluster (*str* in *street*), and the rhyme or body is the vowel and the final consonants (*eet* in *street*). Although letter–phoneme correlations are tracked at the grain size of letters and phonemes, the overall pattern of weights reflects a correlational structure at any larger grain size. It turns out that, for English, correlations at the grain size of onsets and rhymes are highly predictive (Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995).

Stone et al. (1997) used four sets of words in a 2 × 2 design. Bidirectionally consistent words were consistent in both spelling to sound and sound to spelling. An example is the word *lust.* Its spelling body, *-ust*, is only pronounced one way in the various words that share this spelling body, and its pronunciation body, /_ust/, is only spelled one way in the words that share this pronunciation body. A second set of words was bidirectionally inconsistent. For example, the spelling body *-eak* in *bleak* has multiple pronunciations, as in *break* and *leak*, and the pronunciation body /_eak/ has multiple spellings, as in *freak* and *creek*. The third and fourth sets of words were consistent in one direction but inconsistent in the other. The example *heap* has a spelling body, *-eap*, that is always pronounced the same, but the pronunciation body, /_eap/, can be spelled multiple ways, as in *creep* and *leap.* The contrasting example *hull* has a pronunciation body, /_ull/, that can only be spelled one way but a spelling body, *-ull*, that can be pronounced multiple ways, as in *dull* and *pull.* In a lexical decision task, words that were consistent in both directions yielded shorter correct "yes" response times than words that were inconsistent in either direction. These empirical findings agree with our suggestion that spelling and reading are fundamentally related.

If this is so, what is the implication for spelling, particularly Dutch strange words, which are sound-to-spelling inconsistent (and, depending on the spelling–sound grain size, often also spelling-to-sound inconsistent)? Recall that any trained network favors consistent relations over inconsistent ones and can overcome slow or inaccurate processing through frequent presentations. With regard to spelling, frequent spelling trials may compensate for the inconsistency of Dutch strange words. Although repetition is helpful to some extent, earlier work has shown that frequent reading of words is actually the least effective way to learn the spelling of words (e.g., Bosman & van Leerdam, 1993; van Leerdam, Bosman, & Van Orden, 1998).

A robust phenomenon that was established with participants both with and without (acquired) dyslexia and is directly related to the current issue is the so-called regularization error. The *regularization error* refers to the phenomenon of readers who pronounce a (printed) word in accordance with the pronunciations of words with similar and dominant pronunciations. For example, when readers name the word *pint* as rhyming with *hint*, *mint*, and *tint*, they commit a regularization error (e.g., Kawamoto & Kitzis, 1991; Kawamoto & Zemblidge, 1992; Patterson & Hodges, 1992; Warrington, 1975; Watt, Jokel, & Behrmann, 1997). Kawamoto and Kitzis (1991) showed not only that participants in their study committed 25% regularization errors on inconsistent words but that the mean latency of these regularized responses was about 100 ms faster than the mean latency of correct responses (see also Kello & Plaut, 2000). The proper pronunciation of sound–spelling inconsistent words appears to require additional effort, because readers have to suppress the statistically more plausible—that is, regularized (i.e., incorrect)—pronunciation of inconsistent words. Classroom observations also show the occurrence of regularization errors in Dutch beginning readers and spellers in their effort to master the written language. When trying to decode the Dutch strange word *jus* (*gravy*), beginning readers apply Dutch prototypical grapheme–phoneme conversion rules, which leads to the incorrect pronunciation /jYs/ rather than the proper /sjy/. An experimental example in Dutch is a finding by Hasselman (2000). He showed that the number of regularizations in Dutch children with dyslexia who were asked to read as quickly as possible increased markedly compared with a self-paced reading task.

Although the regularization error is a reading error, it may actually support the spelling process. After all, regularizing solves the inconsistencies in the system. In other words, to remedy inconsistencies, the speller has the option of pronouncing odd words in a regularized fashion. For example, before one attempts to spell the word *pint*, it may be helpful to regularize its pronunciation by having it rhyme with *hint*, *mint*, and *tint,* which would result in /pInt/ rather than the standard pronunciation /pAInt/.[1] Regularized pronunciations accord with the prototypical phoneme–grapheme relations, and the spelling of these words should therefore be easier to write down.

Ormrod and Jenkins (1989) were the first to show that applying regularized reading indeed enhances spelling performance. They asked students at various grade levels (Grades 3, 4, 7, and 8) and American undergraduates to practice the spelling of 10 words the students could not spell. While doing this, the students were told to "think aloud." From the tape-recorded utterances and the results on a subsequent spelling test, Ormrod and Jenkins concluded that the most effective method was what they called *overpronunciation*, which we refer to as *regularized reading.* A small minority in the group of undergraduates only (i.e., 14%) appeared to apply this strategy. Students in the other grade levels did not spontaneously use regularized reading; their predominant strategy was either letter rehearsal or standard reading (i.e., using the proper pronunciation). These strategies were also used most often in the group of undergraduates (26% and 39%, respectively). It is interesting that the amount of study time did not correlate with scores on the spelling tests. Stated differently, the amount of time devoted to the task did not compensate for apparently ineffective strategies.

Holmes and Malone (2004) used an approach similar to that of Ormrod and Jenkins (1989) with adult advanced English speakers (i.e., university students) who were either good spellers or weak spellers. The participants were asked to learn the spelling of 10 words that they had spelled incorrectly on a prior spelling test. While studying the words, they had to think out loud. It appeared that both groups used letter rehearsal most often. The second most used strategy was overpronunciation, followed by comparison of the remembered and the correct spelling, morphological analysis, and visualization. On a subsequent spelling test, it appeared that good spellers benefited from all study strategies, whereas weaker spellers were less successful with overpronunciation, comparison, and morphological analysis. Weaker spellers were not always successful in the application of overpronunciation because they sometimes formed incorrect overpronunciations or had forgotten the overpronunciations.

On the basis of theoretical notions derived from our recurrent network and supported by the evidence provided by Ormrod and Jenkins (1989) and Holmes and Malone (2004), we predict that learning the spelling of strange words read in a regularized fashion will enhance spelling performance more than learning the spelling of strange words using a standard-reading strategy. We carried out two experiments to investigate this hypothesis. To satisfy an additional goal—that is, finding an effective strategy for educational practice—we decided (a) to run the experiments with inexperienced spellers from primary education, including students with and without learning disabilities to test for potential differential effects, and (b) to administer an immediate posttest as well as a retention test to test for performance stability.

Experiment 1

Method

Participants. A total population of 187 primary-school students, 84 students without learning disabilities and 93 with learning disabilities, was presented with the Schaal Vorderingen Spellingvaardigheid 2′ (Scale Achievement Spelling Skill 2), a Dutch standardized spelling test by van den Bosch, Gillijns, Krom, and Moelands (1990). The score on this spelling test (minimum = 0, maximum = 36) was used to select a group of students with learning disabilities and a group of students without learning disabilities with similar spelling skills. The selection resulted in 41 students without learning disabilities, with a mean score of 28.9 (SD = 4.0; range = 13; minimum = 20, maximum = 33), and 44 students with learning disabilities, with a mean score of 28.8 (SD = 4.2; range = 15; minimum = 19, maximum = 34). The mean score of the sample of students without learning disabilities who were drawn from their population (28.9) was significantly lower than the mean score of their peers who did not participate in the experiment (31.6), $F(1, 82) = 8.65$, $p < .01$, whereas the mean score of the sample of students with learning disabilities (28.8) was significantly higher than the mean score of their peers who did not participate in the experiment (22.9), $F(1, 101) = 14.43$, $p < .001$. The students without learning disabilities were all recruited from Grade 3 of three different primary schools in the eastern part of the Netherlands. The students officially diagnosed with learning disabilities were all recruited from five schools for special education, also located in the eastern part of the Netherlands. The diagnostic label *learning disabilities* in the Netherlands refers to an educational delay that results from a specific learning disability, such as dyslexia or dyscalculia, or from a general learning disability not due to below-average intelligence level. At the time of the experiment, only students with normal or above-normal intelligence levels were admitted to a school for children with learning disabilities (in the Netherlands, Leer- and OpvoedingsMoeilijkheden [LOM] school). Thus, none of the students in the experiment had a below-average intelligence. No details were available on the types of learning disabilities of each child who participated.

The students from each sample were randomly assigned to one of two conditions—that is, the regularized-reading condition or the standard-reading condition—with the restriction that the spelling scores of all four groups were statistically equal ($F < 1$). Students with learning disabilities had similar educational levels but were, on average, 25 months older than students without learning disabilities, $F(1, 82) = 369.36$, $p < .0001$. The mean age of students without learning disabilities was 110 months (SD = 4.8; range = 22; minimum = 102, maximum = 124), whereas the mean age of students with learning disabilities was 135 months (SD = 7.0; range = 30; minimum = 117, maximum = 147). Table 1 lists the relevant information regarding the participant groups. Note that in both groups, the number of boys was larger than the number of girls, but the difference was considerably larger in the group of students with learning disabilities (80% boys and 20% girls), whereas in the group without learning disabilities, the percentages were 61% boys and 39% girls.

Materials. To select experimental words that students who took part in the experiment did not have any or had little knowledge of, we presented a set 50 Dutch words with inconsistent phoneme–grapheme relations to a group of 20 students from Grade 3 without learning disabilities and with the same educational level as the students who participated in the experiment. To obtain a valid assessment, we ensured that the group of students who participated in the test for word selection represented a typical Grade 3 group. Without prior training or spelling instruction, they were asked to spell all 50 words. Only words that were spelled incorrectly by at least 90% of the students were selected for the training (M = 96% incorrect). The final selection consisted of 16 predominantly nonnative Dutch or strange words, namely, *asperge* (*asparagus*; 95% incorrect), *bungalow* (*bungalow*; 95% incorrect), *champignon* (*mushroom*; 100% incorrect), *douane* (*customs*; 100% incorrect), *giraffe* (*giraffe*; 100% incorrect), *jeans* (*jeans*; 95%

[1] Incidentally, the English word *pint* has the same meaning and spelling in Dutch, but the Dutch word's phonology deviates from the English; it actually received the English regularized pronunciation /pInt/.

Table 1
Spelling Scores, Age, and Gender Ratio of the Experimental Groups in Experiment 1

	Without LD				With LD			
	Standard		Regularized		Standard		Regularized	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Spelling (max. = 36)	28.8	4.2	29.0	3.9	29.1	3.7	28.6	4.8
Age (months)	109.0	4.1	111.0	5.4	137.0	6.9	134.0	7.2
Girls–boys	9–12		7–13		2–20		7–15	
n	21		20		22		22	

Note. LD = learning disability.

incorrect), *jungle* (*jungle*; 90% incorrect), *jus* (*gravy*; 95% incorrect), *kangoeroe* (*kangaroo*; 100% incorrect), *kievit* (*lapwing*; 95% incorrect), *milieu* (*environment*; 95% incorrect), *niveau* (*level*; 100% incorrect), *onmiddellijk* (*immediately*; 90% incorrect), *passagier* (*passenger*; 100% incorrect), *populair* (*popular*; 95% incorrect), and *station* (*station*; 90% incorrect).[2] The set of 16 experimental words was printed three times on three A–4-size pieces of paper in three different orders, resulting in a list of 48 words presented in one single column.

Recall that students who participated in the spelling-selection task did not take part in the experimental training. To ensure that any (little) preexisting knowledge of the experimental stimuli did not affect posttest outcomes, we randomly assigned students to either of the two conditions. Moreover, the students without learning disabilities who participated in the experiment were likely to have even less knowledge of the experimental stimuli than the students who were pretested on the materials, because their mean spelling level was significantly below that of their population.

Procedure. All students were trained and tested in individual sessions by one of two experimenters in a separate and quiet room in the school designated for the study. Teachers were not involved in the experiment or in the testing in any way. Two training conditions were developed according to a written protocol: standard reading and regularized reading. In both conditions, five practice words (all strange words) preceded the experiment proper to explain the task and explicate anything that was unclear. The practice words were *toilet* (*toilet*), *vakantie* (*holidays*), *horloge* (*watch*), *journaal* (*news*), and *computer* (*computer*).

In the standard-reading condition, students were asked to read each word aloud according to the proper pronunciation. For example, the proper reading of the word *toilet* in Dutch is its French pronunciation /twA/ /lEt/. If a word was read incorrectly, the experimenter provided the proper reading. While performing the first reading of the word, students often had to correct themselves or received help from the experimenter, because this was usually the first time that they were exposed to the spelling of these words (they often regularized the pronunciation on first reading). The second and third reading of the words did not present serious difficulties.

In the regularized-reading condition, students were asked to read the words according to prototypical Dutch letter–sound relations. Stated differently, they were asked to apply phonological or regularized reading. For example, the regularized reading of the word *toilet* in Dutch is more in common with the English pronunciation, that is, /tçi/ /let/. Primary-grade students spontaneously apply prototypical Dutch letter–sound couplings, because the majority of Dutch words have regular pronunciations. Some students realized after reading the word which word they had actually read and corrected themselves. In this condition, however, the experimenter stressed that the students were required to read in a regularized fashion (it was called the "funny way of reading"). In all cases, the standard pronunciation of the word—for example, /twA /lEt/ for *toilet*—was provided by the experimenter in the first reading only. In the incidental case in which the student did not apply regularized reading, the experimenter corrected the response.

In both conditions, students were asked to read all 48 words (each word three times) in one session. In incidental cases, a word's meaning required explanation by the experimenter; this pertained to the word *milieu* (*environment*) almost exclusively. Completion of the task in both conditions, including the instruction, the practice words, and the reading of the experimental trials, required, on average, 10 min.

Two spelling tests followed the training. The first one, the posttest, was presented in the afternoon of the day the training took place. The second one, a retention test, was presented 1 week after the training. Both tests required students to spell each of the experimental words that were read to them using the standard pronunciation. Thus, students who were trained to apply regularized reading had to spell the word on oral presentation of the standard pronunciation.

Results

Two observers (both experimenters) determined the accuracy of the spelling of each word. In most cases this was not a problem, because students' handwriting was generally perfectly legible; in the odd case in which it was not, they discussed the identity of the spelling until agreement was reached, which caused the level of agreement to be 100%. For each student, the percentages of correctly spelled words on the posttest and the retention test were computed. The mean percentages correct are displayed in Table 2.

A 2 (group: without learning disabilities vs. with learning disabilities) × 2 (training condition: standard reading vs. regularized reading) × 2 (test: posttest vs. retention test) analysis of variance (ANOVA) on the percentages correct, with group and condition as between-subjects variables and test as a within-subject variable, revealed a significant three-way interaction, $F(1, 81) = 4.19$, $p = .04$, partial $\eta^2 = .05$. Because one of the additional issues was the potential performance difference between students with and students without learning disabilities, we decided to investigate the source of this interaction by performing separate analyses on the data of the two groups.

[2] The words *kievit* and *onmiddellijk* are actually not strange words in the linguistic sense, but they both present spelling problems as a result of their odd pronunciations. Moreover, the word *kievit* was phoneme–grapheme inconsistent when Experiment 1 was performed. After the last spelling reform, the spelling of the word *kievit* was changed into *kieviet*, and the word now has consistent phoneme–grapheme relations. Experiment 2 was performed after the spelling reform, but we forgot to check whether the spelling of this word had been changed; therefore, students in Experiment 2 learned the former spelling of the word *kievit.*

Table 2
Percentage Correct Spelling on Posttest and Retention Test in the Two Training Conditions of Experiment 1

	Without LD				With LD			
	Standard		Regularized		Standard		Regularized	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Posttest	17.6	12.8	45.3	27.8	19.3	13.8	38.1	21.2
Retention test	14.6	11.8	39.1	25.7	20.7	13.4	27.8	21.8

Note. LD = learning disability.

Students without learning disabilities. A 2 (training condition: standard reading vs. regularized reading) × 2 (test: posttest vs. retention test) ANOVA was performed on the percentages correct of the students without learning disabilities, with condition as a between-subjects variable and test as a within-subject variable. The interaction effect between training condition and test did not reach significance, $F(1, 39) = 1.67$, $p = .20$, partial $\eta^2 = .04$, whereas both main effects did. The main effect of training condition indicated that students without learning disabilities showed better spelling performance in the regularized-reading condition than in the standard-reading condition on both posttest and retention test, $F(1, 39) = 17.02$, $p = .0002$, partial $\eta^2 = .30$. The main effect of test showed better spelling performance on the posttest than on the retention test, $F(1, 39) = 13.36$, $p = .001$, partial $\eta^2 = .26$. Thus, in both conditions, spelling knowledge of the trained words had dropped a week after the training ended. To test whether performance on the retention test by students who participated in the regularized-reading training was still superior to the performance of those who took part in the standard-reading training, we conducted an independent samples *t* test. This appeared to be the case, $t(26.3) = -3.89$, $p = .001$. This *t* test was an unequal variance test (Levene's test for equality of variances was used). In sum, the significant difference between the training conditions on the posttest indicates that, in the short term, regularized reading was superior to standard reading. Performance appeared to be unstable, because in both training conditions posttest scores were better than retention test scores. However, the fact that performance on the retention test by students who took part in the regularized-reading training was still better than the performance of those who participated in the standard-reading training indicates a more robust effect of regularized reading in students without learning disabilities.

Students with learning disabilities. A 2 (training condition: standard reading vs. regularized reading) × 2 (test: posttest vs. retention test) ANOVA was performed on the percentages correct of the students with learning disabilities, with condition as a between-subjects variable and test as a within-subject variable. The main effect of task was significant, $F(1, 42) = 6.21$, $p = .02$, partial $\eta^2 = .13$, as was the main effect of test, $F(1, 42) = 7.71$, $p = .008$, partial $\eta^2 = .16$. Because the interaction effect between training condition and test also reached significance, it qualifies the main effects, $F(1, 42) = 13.49$, $p = .001$, partial $\eta^2 = .24$. We made subsequent pairwise comparisons to investigate the source of this interaction. It appeared that performance on the posttest by students who took part in the regularized-reading training was superior to the performance of those who participated in standard-reading training ($p = .0001$), whereas this difference was not significant on the retention test ($p = .20$). Moreover, two additional pairwise comparisons indicated that the performance of students in the regularized-reading training dropped significantly from posttest to retention test ($p = .0001$), whereas the performance of students in the standard-reading training remained stable ($p = .53$). In sum, the significant difference between the training conditions on the posttest indicates that, in the short term, regularized reading was superior to standard reading. However, the performance of students who took part in the regularized-reading training appeared to be rather unstable. Performance on the retention test did not just drop significantly, it dropped to the level of the students in the standard-reading condition, whose performance was similar in both the posttest and the retention test.

Discussion

The results of this short training confirm the short-term superiority of regularized reading for students both with and without learning disabilities. Even though spelling performance was not perfect immediately after training, scores on the posttest of both groups were better for students who took part in the regularized-reading training than for students who participated in the standard-reading training. Given that a similar group of students was virtually unable to spell any of the words and that only three presentations established this effect, the conclusion is warranted that regularized reading is an effective and efficient spelling training for words with inconsistent phoneme–grapheme relations.

In the long term, however, performance was not stable for students who took part in the regularized-reading training. In both groups, scores on the retention test were lower than on the posttest. In the case of the students with learning disabilities, the performance of those who took part in the regularized-reading training actually dropped to the level of students in the standard-reading condition, but the performance of students without learning disabilities who took part in the regularized-reading training was still superior. Although neither of the two groups showed a stable long-term effect, students without learning disabilities retained more of their spelling knowledge than students with learning disabilities.

The spelling knowledge of students without learning disabilities who took part in the standard-reading training dropped from posttest to retention test, whereas it remained stable for students with learning disabilities. A possible explanation for this effect relates to memory capacity. Memory capacity increases with age (Gathercole, Pickering, Ambridge, & Wearing, 2004). It is likely that the memory capacity of the students with learning disabilities, who were, on average, more than 2 years older than the students without learning disabilities, was sufficiently large to enable them to remember the limited amount of spelling knowledge (about 20%) they had acquired.

In sum, the short-term effect of the training was comparable for both types of students, but the long-term effects were superior in the group of students without learning disabilities. Two possible reasons for the difference in performance come to mind. First, it may well be that students with learning disabilities required more training to enable stable development of orthographic knowledge. Second, perhaps students with learning disabilities had more difficulty remembering the regularized pronunciation of the phoneme–grapheme inconsistent word, a necessary prerequisite for a successful training result, a problem that also occurred in the university student sample of Holmes and Malone (2004).

Experiment 2

To evaluate the explanations we have given for the decline in performance of students with learning disabilities, we ran Experiment 2. Two extended regularized-reading training conditions were developed. One training was identical to the regularized version of Experiment 1 but lasted 3 weeks rather than 1. The effectiveness of this original regularized-reading training was compared with a training with one additional component, namely, practicing the regularized pronunciation of the phoneme–grapheme inconsistent words, implemented to enhance participants' memory for the regularized pronunciation. We called this training *enhanced-regularized reading*. Furthermore, we investigated the effect of spelling level, training effectiveness, and transfer. After all, training may be most (and sometimes only) effective for students whose spelling level is already high compared with that of their less-skilled classmates.

Method

Participants. From a group of 76 students with learning disabilities who were attending special education, we selected a group of 33 students who met three requirements.[3] First, students with learning disabilities participating in Experiment 2 had to come from the same category as those from Experiment 1. This was ensured by the fact that students in this experiment were recruited from the same type of school (i.e., LOM school) as were those from Experiment 1. Second, the present group had to be of statistically similar age as the students with learning disabilities from Experiment 1. A one-way ANOVA proved that to be the case ($F < 1$). Third, the spelling level of the group in Experiment 2 had to be statistically the same as that of the students with learning disabilities who took part in Experiment 1. Spelling level was again assessed by means of Scale Achievement Spelling Skill 2, by van den Bosch et al. (1990). The mean scores of the students with learning disabilities in Experiments 1 and 2 were indeed similar ($F < 1$). Thus, with regard to the basic requirements, the two groups of students in Experiments 1 and 2 were similar.

Subsequently, each of the 33 students was randomly assigned to either the regularized-reading condition ($n = 17$) or the enhanced regularized-reading condition ($n = 16$). Again, the mean scores on the spelling test and the mean ages of the two experimental groups in Experiment 2 were statistically identical to each other (spelling level, $F < 1$; mean age, $F < 1$). To be able to further differentiate between good and poor spellers in the present group of participants, we designated the 50% of the students who performed best on the spelling test as students with a high spelling level, and we designated the 50% of the students who performed worst as students with a low spelling level. In this experiment, we acquired additional information on reading performance. We administered a standardized word-reading test, the Eén-minuut-test (One-Minute Test), by Brus and Voeten (1973), and a standardized pseudoword-reading test by van den Bos, lutje Spelberg, Scheepstra, and de Vries (1994). The two experimental groups showed statistically equal performance on both tests ($F < 1$ in both cases). As in Experiment 1, more boys than girls participated in Experiment 2. All information regarding the experimental groups is presented in Table 3.

Materials. The same set of words that was used in Experiment 1 was used in this experiment.

Procedure. All students were trained and tested in individual sessions by one experimenter in a separate and quiet room in the school designated for the study. Again, teachers were not involved in the experiment or in the testing in any way. Moreover, to prevent intervention by teachers during the extended training procedure, we did not inform them about the materials or the content of the training until after completion of training and testing. The same training protocol used in Experiment 1 was applied in this experiment. The regularized-reading condition was identical to the one used in Experiment 1. The procedure of the enhanced regularized-reading condition was identical to the regularized version, except for the addition of the memory component. After each regularized-reading training session, the experimenter named each of the words according to its standard pronunciation and asked the students to say the regularized, or "funny," pronunciation. For example, "What is the best way to pronounce /twA/ /lEt/ (i.e., *toilet*) such that it is easy to remember how to spell it?" The student then had to answer "/tçi/ /let/."

The important difference between the two experimental groups in this experiment was the fact that the students in the enhanced regularized-reading group after the reading training were asked to say the regularized pronunciation of each of the 16 words on oral presentation of the standard pronunciation, which required less than a minute. We want to emphasize that the students in this condition never saw the words again. Thus, additional visual-information processing, which could have enhanced the memory for spelling, did not occur; this was purely an auditory task that required little additional time.

Both groups received training once a week during a period of 3 weeks. After each training session, they received a spelling test on all 16 trained words. The words to be spelled were pronounced according to the standard pronunciation. The three tests that followed the training are referred to as Posttest 1, Posttest 2, and Posttest 3. In the 4th week, participants did not receive training but were asked to spell all 16 words again; this test is referred to as Retention Test 1. Additionally, all students were presented with a remembrance test—that is, they were asked whether they could remember the funny way of pronouncing the words they had studied in the preceding weeks (i.e., the regularized pronunciations). Finally, a month after the first retention test and without any training, the students had to spell the words one more time; this test is referred to as Retention Test 2. A transfer test concluded the final test session. All students were presented with a list of 20 words; half of these contained phoneme-to-grapheme consistent relations, and half contained phoneme-to-grapheme inconsistent relations (i.e., strange or nonnative Dutch words). The students were asked to read each word and indicate on paper which of these words were and which words were not suitable for application of the regularized method of pronunciation to better remember the spelling.

[3] Initially, 40 students were selected, but 7 students failed one or more training sessions. The data of these students had to be discarded from the analyses, which left us with the data of 33 students.

Table 3
Test Scores, Age, and Gender Ratio of the Experimental Groups in Experiment 2

Variable	Regularized reading *M*	Regularized reading *SD*	Enhanced regularized reading *M*	Enhanced regularized reading *SD*
Spelling (max. = 36)	26.4	7.0	26.4	7.4
Word reading (wpm)	51.3	13.3	52.4	15.4
Pseudoword reading (wpm)	17.8	7.7	17.8	7.4
Age (months)	136.0	8.0	135.0	9.0
Girls–boys	8–9		4–12	
n	17		16	

Note. LD = learning disability.

Results

Two observers (the experimenter and Anna M. T. Bosman) determined the accuracy of the spelling of each word. Because students' handwriting was generally perfectly legible, this did not present a problem. In the odd case in which a difference between observers emerged, they discussed the identity of the spelling until agreement was reached, which caused the level of agreement to be 100%. For each student, we computed the percentages of correctly spelled words on all three posttests and the two retention tests.

Experiment 2 was designed to investigate five questions. The first question was whether students with learning disabilities benefited from extended training in regularized reading. The second question concerned the issue of whether practicing the funny pronunciation (i.e., enhanced regularized reading) added to the effectiveness of regularized reading. The third issue was whether students in the enhanced regularized-reading condition were better at remembering the funny pronunciation than students in the regularized-reading condition. The fourth issue concerned the question of to what extent students with learning disabilities were capable of transferring knowledge from the training to a new set of materials. The fifth and final question was whether the training was equally effective for those with a relatively high and a relatively low spelling level.

Before putting these questions to test, we performed an additional analysis to ensure that the learning-disabled students in Experiments 1 and 2 indeed had similar spelling levels. In the current situation, the most appropriate test is to compare the mean percentage correct of the posttest of students with learning disabilities who participated in the regularized-reading condition of Experiment 1 (M = 38.1, SD = 21.2) with the results of the first posttest of students with learning disabilities who participated in the same regularized-reading condition of Experiment 2 (M = 29.4, SD = 21.1). The students in Experiment 2 appeared to have more limited knowledge than those in Experiment 1, but the results of a t test indicated that this difference was not statistically significant, $t(37) = 1.26$, $p = .22$, $d = 0.41$.

Extended-training results. To establish whether longer training led to better short-term and long-term spelling results, we made a comparison between the mean percentage correct scores on the posttest and the retention test of the students with learning disabilities who took part in the regularized-reading training in Experiment 1 and the scores on Posttest 3 and Retention Test 1 of the students who participated in the same regularized-reading training of Experiment 2. A 2 (experiment: 1 vs. 2) × 2 (test: posttest vs. retention test) ANOVA on the participants' mean percentages correct, with experiment as a between-subjects variable and test as a within-subject variable, showed a significant main effect of experiment, $F(1, 36) = 16.98$, $p = .0001$, partial $\eta^2 = .32$. This effect indicated that more words were spelled correctly in Experiment 2 (M = 65%, SD = 27%) than in Experiment 1 (M = 33%, SD = 21%). The main effect of test was also significant, indicating that more words were spelled correctly on the posttest (M = 49%, SD = 28%) than on the retention test (M = 43%, SD = 29%), $F(1, 36) = 9.23$, $p = .004$, partial $\eta^2 = .20$.

Because of the significant interaction effect, both main effects had to be qualified, $F(1, 36) = 7.95$, $p = .008$, partial $\eta^2 = .18$. Subsequent t tests indicated that students in Experiment 2, who had had more training, spelled more words correctly on the posttest (M = 65%, SD = 29%) than students in Experiment 1 (M = 38%, SD = 21%), $t(36) = -3.31$, $p = .002$, $d = 1.10$. A similar effect emerged on the retention test 1 week later: Experiment 2 (M = 65%, SD = 25%) and Experiment 1 (M = 28%, SD = 22%), $t(36) = -4.77$, $p < .0001$, $d = 1.59$. Moreover, paired-sample t tests on the mean percentages correct on the posttest and retention test of students in Experiment 2 showed that their performance remained stable, $t(15) < 1$, whereas the scores of the students in Experiment 1 dropped significantly (see the *Results* section of Experiment 1). Thus, students in Experiment 2, who had 3 weeks of training, showed better short-term and long-term performance than students in Experiment 1, who just had 1 week of training.

Enhanced-training results. For each student, we computed the percentages of correctly spelled words on all three posttests and both retention tests. The mean percentages correct are displayed in Table 4. A 2 (spelling level: high vs. low) × 2 (condition: regularized reading vs. enhanced regularized reading) × 5 (test: Posttest 1 vs. Posttest 2 vs. Posttest 3 vs. Retention Test 1 vs. Retention Test 2) ANOVA on mean percentages correct, with spelling level and condition as between-subjects variables and test as a within-subject variable, revealed no significant interaction effects, indicating that all main effects could be interpreted unambiguously. The Greenhouse–Geisser procedure had to be used to adjust the degrees of freedom for the F tests, because Mauchley's sphericity test was significant (i.e., the homogeneity of variance assumption was violated). All three main effects reached significance. Students with a relatively high spelling skill (M = 65%, SD = 19%) showed better performance on the training materials than their less-skilled peers (M = 34%, SD = 19%), $F(1, 29) = 23.21$, $p < .0001$, partial $\eta^2 = .45$. The performance of students in the enhanced regularized-reading condition (M = 56%, SD = 25%) surpassed that of students in the regularized-reading condition (M = 44%, SD = 22%), $F(1, 29) = 4.06$, $p = .05$, partial $\eta^2 = .12$. Thus, students who practiced the odd pronunciation benefited even more from the regularized-reading training than students who did not. The main effect of the repeated measure test was significant, $F(3, 88) = 31.69$, $p < .001$, partial $\eta^2 = .52$. Post hoc tests (Bonferoni corrected) indicated that mean percentages correct on Posttest 3 and Retention Test 1 were higher than those on Posttest 2 and Retention Test 2, which, in turn, were higher than those on Posttest 1. None of the other comparisons reached significance. Because the interaction between spelling level and test was not

Table 4
Percentage Correct Spelling on Posttests and Retention Tests in the Two Training Conditions of Experiment 2

	Regularized reading				Enhanced regularized reading			
	High		Low		High		Low	
Spelling level	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Posttest 1	42	21	15	7	52	28	27	24
Posttest 2	58	17	26	16	73	18	41	26
Posttest 3	67	22	38	17	79	15	51	33
Retention Test 1	66	22	34	12	80	14	49	26
Retention Test 2	60	21	26	17	72	18	38	15
Overall	59	19	28	11	71	17	41	24

significant, we can conclude that both the high-performing and the low-performing spellers benefited from the training.

Remembrance-test results. To ascertain whether students in the enhanced regularized-reading condition better remembered the regularized pronunciation than students in the regularized-reading condition, we compared mean percentage correct scores on the remembrance test (administered during Retention Test 1). A 2 (spelling level: high vs. low) × 2 (condition: regularized reading vs. enhanced regularized reading) ANOVA, in which both factors were treated as between-subjects variables, on the mean remembrance scores revealed significant main effects and no interaction effect. Students who took part in the enhanced regularized-reading condition correctly remembered 87% (*SD* = 14%) of the regularized pronunciations, whereas students in the regularized-reading condition correctly remembered 78% (*SD* = 15%), $F(1, 29) = 4.26$, $p = .05$, partial $\eta^2 = .13$. Students with high spelling levels remembered 89% (*SD* = 10%) of the regularized pronunciations, and students who were less skilled remembered 75% (*SD* = 16%), $F(1, 29) = 10.20$, $p = .003$, partial $\eta^2 = .26$.

Transfer-test results. To test whether there was a differential effect for the two experimental groups on the transfer test (administered during Retention Test 2), we compared mean percentages correct. A 2 (spelling level: high vs. low) × 2 (condition: regularized reading vs. enhanced regularized reading) × 2 (phoneme–grapheme consistency: consistent vs. inconsistent) ANOVA on the percentage correct was performed, with spelling level and condition as between-subjects variables and phoneme–grapheme consistency as a within-subject variable. None of the interaction effects was significant. The main effect of spelling level showed that highly skilled spellers (*M* = 94%, *SD* = 8%) better knew which words were and which were not suitable for a regularized-reading strategy than their less-skilled peers (*M* = 88%, *SD* = 11%), $F(1, 29) = 4.43$, $p = .04$, partial $\eta^2 = .13$. The main effect of training condition suggested that students in the enhanced regularized-reading condition (*M* = 94%, *SD* = 7%) knew better which words were and which were not suitable for regularized pronunciation than students in the regularized-reading condition (*M* = 88%, *SD* = 11%). This effect, however, failed to reach significance, $F(1, 29) = 3.94$, $p = .06$, partial $\eta^2 = .02$. The main effect of phoneme–grapheme consistency was significant, $F(1, 29) = 5.31$, $p = .03$, partial $\eta^2 = .16$, indicating that students were less able to judge that inconsistent words (*M* = 88%, *SD* = 13%) were suitable for the regularized-pronunciation strategy than that consistent words were unsuitable (*M* = 95%, *SD* = 10%).

Discussion

The results of Experiment 2 show that the spelling performance of students with learning disabilities benefited greatly from an extended training. On the basis of the comparison between Experiments 1 and 2, short-term effects revealed that 3 weeks of training (practicing regularized reading of words nine times rather than three) resulted in spelling knowledge that was twice as good. More important, a week after training had stopped, spelling knowledge had not deteriorated. A month after the training, however, spelling knowledge had dropped to the level of the Posttest 2, but the results were still better than those of Posttest 1. Thus, there was a long-term effect after a week and a diminished long-term effect after a month. Rehearsing the funny or regularized pronunciation aided the learning process even more. Not only was the performance of this group better than that of the students who did not practice the regularized pronunciation, these students were also better at remembering the regularized pronunciation. The results on the transfer task suggested that the metacognitive awareness of students who took part in the enhanced regularized-reading condition had increased more than that of students who participated in the regularized-reading condition. However, the 6% difference was too small to reach significance. Relatively highly skilled spellers had better scores on all posttests and retention tests, better remembered the funny pronunciations, and performed better on the transfer task than their less-skilled peers. Note, however, that the spelling performance of the low-skilled spellers increased the same amount as that of highly skilled spellers.

An interesting and partly unexpected result is that the students appeared to be better at judging that consistent words were not suitable than that inconsistent words were suitable for the regularized-reading strategy. After all, if students encounter an inconsistent word for the first time and apply the prototypical grapheme–phoneme relations to decode it, the result should be the pronunciation of a word that does not exist. This would then be an indication that the word is probably inconsistent and therefore suitable for the regularized-reading strategy. However, the students found it more difficult to decide that inconsistent words were suitable for the strategy than that consistent words were unsuitable. Careful inspection of the item data revealed a possible explanation for our finding. The total number of errors on the transfer task was 58. In 18 cases, students indicated wrongly that consistent words were suitable for the regularized-reading strategy. In the remaining 40 cases, students indicated wrongly that inconsistent words were unsuitable for the regularized-reading strategy. Unlike the even distribution of errors made in the set of consistent words, the majority of errors (65%) in the set of inconsistent words were committed on two words. The word *shampoo* (*shampoo*) was erroneously evaluated by 13 students to be unsuitable for the regularized-reading strategy, and the word *chips* (*potato chips*) was erroneously evaluated by 10 students. Both words were highly familiar to these students, and a substantial number of students were probably able to read these words without difficulty and as a result did not become aware of the words' inconsistent status.

General Discussion

Regularized reading proved to be a valuable spelling training for learning the spelling of Dutch sound-to-spelling inconsistent or strange words in students with and without learning disabilities. For both the relatively highly skilled and the less skilled students with learning disabilities, extended training consolidated the learning process, and auditory practice of the regularized pronunciation added to this. Previous work by Bosman and de Groot (1992; Bosman & van Leerdam, 1993; van Leerdam et al., 1998) also revealed that standard reading is not as effective for learning to spell words with ambiguous phoneme–grapheme relations compared with almost any other spelling training. Words with ambiguous phoneme–grapheme relations are a subclass of words with inconsistent phoneme–grapheme relations; they contain one or more phonemes that can be spelled multiple ways. For example, the /iˇ/ in the English word *cheap* is an ambiguous phoneme, because there is also the *ee* as in *keep*, *ey* as in *key*, *ie* as in *chief*, *e* as in *here*, and *y* as in *entry*.

These examples show that the pronunciation (i.e., reading) of these words is relatively unambiguous, whereas the spelling of the words is not. As we have stated, in most alphabetic languages, sound-to-spelling inconsistency is higher than spelling-to-sound inconsistency. In other words, there are more possible ways to spell a word than there are possible ways to pronounce a word (Stone et al., 1997; Waters, Bruck, & Seidenberg, 1985). In terms of our recurrent network, this constitutes the microbasis for the fact that spelling is almost always more difficult than reading (Bosman & Van Orden, 1997). To understand the macrobasis for the asymmetry between spelling and reading and why regularized reading is a more effective means for learning the spelling of inconsistent words than standard reading, we focus again on the dynamics of the recurrent network.

As we have noted, both spelling and reading are constrained by the relatively strong dynamic between phonologic and orthographic nodes. However, inconsistencies in these relations must be resolved by different sources of constraint. When a model or a reader reads a low-frequency, letter-to-phoneme inconsistent word, such as *pint*, for example, the more consistent relation between spelling and phonology would rhyme with *mint.* In this case, the relatively strong dynamic between semantic and phonologic nodes (compared with the weaker one between semantic and orthographic nodes) may supply sufficient constraints for the appropriate phonology. Additionally, contextual sources of semantic activation may also contribute via the relatively strong connections between semantic and phonologic nodes.

With respect to spelling, however, the model must resolve the inverted pattern of ambiguity in the dynamic between phonologic and orthographic nodes. When a model or a speller spells a low-frequency, phoneme-to-letter inconsistent word, such as *heap*, the more consistent spelling for /–hiṗ/ would be as in *deep*. In this case, correct spelling relies on relatively weak relations between semantic and orthographic nodes to supply sufficient activation of the appropriate letters (as illustrated in Figure 1 in the relative boldness of the arrows). Even contextual support is filtered through the weaker connections between semantic and orthographic nodes. This relatively weak support for spelling (the semantic–orthographic relations), compared with the stronger one for reading (the semantic–phonologic relations), is the macrobasis for the asymmetry between spelling and reading. Thus, spelling is more difficult than reading because phoneme–letter relations are more inconsistent than letter–phoneme relations and because the phoneme–letter inconsistencies must be resolved by the relatively weak semantic–letter dynamic, whereas in reading, letter–phoneme inconsistencies are resolved by the stronger semantic–phoneme dynamic. The fact that people engage less in spelling than in reading enhances this fundamental asymmetry between spelling and reading.

Having established that spelling is inherently more difficult than reading and knowing that people engage less in spelling than in reading indicates that enhancing spelling performance requires building stronger relations between words' meanings and their spellings or creating a language with less inconsistent phoneme-to-letter relations. One strategy to strengthen the relation between a word's pronunciation and its spelling has been provided in the introduction—that is, frequent presentations of a word's spelling and repetition of its (regular) pronunciation. A second and apparently more efficient means is provided by our experimental findings: regularized reading. This strategy enables the speller (or the model) to safely rely on his or her strongest dynamic, the one between phonologic and orthographic nodes. After all, regularizing a word's pronunciation allows the speller to rely on the statistically most probable relations. This conclusion not only concerns the spelling of Dutch strange words, it is equally applicable to other alphabetic languages, including English (see Holmes & Malone, 2004). In all written languages in which the spelling of words deviates from the prototypical phoneme–grapheme relations, it is possible to apply the strategy of regularized reading (for an example from the German language, see Thaler & Landerl, 2005)

With this in mind, we briefly return to the explanations of Fischer et al. (1985) and Holmes and Ng (1993) for the distinction between spellers who have relatively good skills and those who do not. Fischer et al. (1985) maintained that good spellers are more linguistically sensitive, and Holmes and Ng (1993) claimed that good spellers have better (visual) knowledge of unusual letter sequences. According to our recurrent network account, there is no need to decide between these two explanations. Although the network distinguishes among orthography, phonology, and semantics, it is impossible to separate these aspects. Stated differently, it does not make sense to attribute any effect to just one of these three, because they all are interconnected as a result of their bidirectional relations. Better visual knowledge cannot be attributed to single causation of the orthographic nodes, because the orthographic nodes are recurrently connected to the phonologic and semantic nodes. Thus, to be linguistically sensitive or to have better (visual) knowledge of unusual letter sequences translates, in terms of our network, into a speller (or a model, for that matter) who has learned the subtle correlations or statistical regularity of phoneme, grapheme, and semantic covariations that occur in written languages.

Before concluding this article with some important educational implications, we discuss some limitations and potential improvements for future research. First, the one-session training of Experiment 1 revealed a short-term learning effect favoring regularized reading in students both with and without learning disabilities. The absence of a robust long-term effect indicates that a spelling training in which words are practiced only three times does not induce permanent spelling knowledge. Second, although standard reading was not as effective as regularized reading, the effects for students with learning disabilities revealed that these students acquired relatively stable spelling knowledge of an average of three words spelled correctly on posttest and retention test. Future research might be directed at a more detailed comparison of regularized and standard reading. Third, we recommend that future researchers administer a pretest to a group with similar educational levels (to pilot the materials) as well as to the experimental groups. Note that we do not believe that our results are compromised, because the goal of the study was a comparison between two study methods, not the effect of a spelling training per se. However, a pretest would have provided us with another measure to assess learning gain.

This experimental study was performed in a natural setting and has potentially important educational implications. One is that the regularized-reading strategy was effective not only for students without learning disabilities but also for students with learning disabilities. Second, unlike most intervention programs, in which the better achiever gains relatively more than the poor one, our training shows that it is possible to develop a training condition that affects these groups similarly. Third, students with learning disabilities, including the ones who did not engage in extra practice of the regularized pronunciation, knew quite well which words were and which were not suitable for the regularized-reading strategy. Generalization of newly acquired knowledge or strategies usually only occurs when students receive elaborate, explicit instruction and practice of an effective strategy (e.g., O'Sullivan & Pressley, 1984). Our finding indicates that even students with learning disabilities are capable of generalizing newly acquired knowledge, although further research is required to establish whether young, inexperienced spellers would be able to apply the regularized-reading strategy effectively with a new set of words.

Finally, initially all readers apply the (proto)typical grapheme–phoneme rules to any word they encounter, because most words (in Dutch, about 85%) obey the standard grapheme–phoneme correspondence rules, and this almost always leads to a satisfactory outcome. When the application of typical correspondences, in the case of strange words, causes the reading of a nonword, a small group of children are able to correct themselves, either on the basis of the context in which the word is presented or because they figure it out anyway. The majority of children, however, need the help of a teacher who tells them which word they are actually reading. It is interesting that the actual occurrence of a regularization error provides the teacher with the opportunity (a) to explain the strategy of regularized reading for spelling and (b) to emphasize the asymmetric relation between reading and spelling. The regularization error is therefore a perfect occasion for the development of metalinguistic skills, which enable the students to become the teacher of their own learning process.

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database* [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Borzone de Manrique, A., & Signorini, A. (2000). *Lectura y prosodia: Una via para el estudio del procesamiento cognitivo* [Reading and prosody: A way to study cognitive processing]. *Interdisciplinaria, 17,* 95–117.

Bosman, A. M. T., de Graaff, S., & Gijsel, M. A. R. (2006). Double Dutch: The Dutch spelling system and learning to spell in Dutch. In R. M. Joshi & P. G. Aron (Eds.), *Handbook of orthography and literacy* (pp. 135–150). Mahwah, NJ: Erlbaum.

Bosman, A. M. T., & de Groot, A. M. B. (1992). Differential effectiveness of reading and non-reading tasks in learning to spell. In F. Satow & B. Gatherer (Eds.), *Literacy without frontiers* (pp. 279–289). Widnes, Cheshire, England: United Kingdom Reading Association.

Bosman, A. M. T., & Mekking, T. (2005). *Statistical structure of spelling-sound and sound-spelling relations in Dutch words.* Manuscript in preparation.

Bosman, A. M. T., & van Leerdam, M. (1993). *Aanvankelijk spellen: De dominantie van de verklankende spelwijze en de geringe effectiviteit van lezen als spelling-instructie methode* [Beginning spelling: Prevalence of the phonological strategy and the scant effectiveness of reading as a spelling instruction method]. *Pedagogische Studiën, 70,* 28–45.

Bosman, A. M. T., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 173–194). Hillsdale, NJ: Erlbaum.

Brown, G. D. A., & Loosemore, R. P. W. (1994). Computational approaches to normal and impaired spelling. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 319–335). New York: Wiley.

Brus, B. T., & Voeten, M. J. M. (1973). *Een-Minuut-Test* [One-Minute Test]. Nijmegen, the Netherlands: Berkhout.

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Content, A. (1991). The effect of spelling-to-sound regularity on naming in French. *Psychological Research, 53,* 3–12.

Farrar, W. T., & Van Orden, G. C. (2001). Errors as multistable response options. *Nonlinear Dynamics, Psychology, and Life Sciences, 5,* 223–265.

Fischer, F. W., Shankweiler, D., & Liberman, I. Y. (1985). Spelling proficiency and sensitivity to word structure. *Journal of Memory and Language, 24,* 423–441.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40,* 177–190.

Hasselman, F. (2000). *Subtypes of developmental dyslexia: A study of the effects of strategic control on the overt reading behaviour of dyslexic children.* Unpublished master's thesis, University of Nijmegen, Nijmegen, the Netherlands.

Holmes, V. M., & Malone, N. (2004). Adult spelling strategies. *Reading and Writing: An Interdisciplinary Journal, 17,* 537–566.

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

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36,* 505–529.

Kawamoto, A. H., & Kitzis, S. N. (1991). Time course of regular and irregular pronunciations. *Connection Science, 3,* 207–217.

Kawamoto, A. H., & Zemblidge, J. (1992). Pronunciation of homographs. *Journal of Memory and Language, 31,* 349–374.

Kello, C. T., & Plaut, D. C. (2000). Strategic control in word reading: Evidence from speeded responding in the tempo-naming task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 719–750.

Lukatela, G., Turvey, M. T., & Todorovic, D. (1991). Is alphabet biasing in bialphabetical word perception automatic and prelexical? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 653–663.

McLeod, P., Plunkett, K., & Rolls, E. T. (1998). *Introduction to connectionist modelling of cognitive processes.* Oxford, England: Oxford University Press.

Nunn, A. M. (1998). *Dutch orthography: A systematic investigation of the spelling of Dutch words.* Den Haag, the Netherlands: Holland Academic Graphics.

Olson, A., & Caramazza, A. (1994). Representation and connectionist models: The NETspell experience. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 337–363). New York: Wiley.

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

O'Sullivan, J. T., & Pressley, M. (1984). Completeness of instruction and strategy transfer. *Journal of Experimental Child Psychology, 38,* 275–288.

Patterson, K. E., & Hodges, J. R. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologica, 30,* 1025–1040.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular designs. *Psychological Review, 103,* 56–115.

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

Seidenberg, M. S., Waters, M., 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.

Siegenbeek, M. (1804). *Verhandeling over de spelling der Nederduitsche taal, ter bevordering van eenparigheid in dezelve* [Essay on the spelling of the Dutch language, encouragement to uniformity]. Amsterdam: J. Allart.

Snowling, M. J. (1994). Towards a model of spelling acquisition: The development of some component processes. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 111–128). New York: Wiley.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36,* 337–359.

Stone, G. O., & Van Orden, G. C. (1994). Building a resonance framework for word recognition using design and systems principles. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1248–1268.

te Winkel, L. A., & de Vries, M. A. (1866). *De grondbeginselen der Nederlandsche spelling: Ontwerp der spelling voor het aanstaande Nederlandsch woordenboek* [Basic principles of Dutch spelling: Design of the spelling of the forthcoming Dutch dictionary]. Leiden, the Netherlands: D. Noothoven van Goor.

Thaler, V., & Landerl, K. (2005, June). *The influence of spelling pronunciations on the orthographic spelling competence.* Poster presented at the meeting of the Society for the Scientific Study of Reading, Toronto, Ontario, Canada.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124,* 107–136.

van den Bos, K. P., lutje Spelberg, H. C., Scheepstra, A. J. M., & de Vries, J. R. (1994). *De Klepel: Een test voor de leesvaardigheid van pseudowoorden* [De Klepel: A test for pseudoword-reading skill]. Nijmegen, the Netherlands: Berkhout.

van den Bosch, L., Gillijns, P., Krom, R., & Moelands, F. (1990). *Schaal vorderingen Spellingvaardigheid* 2 [Scale Achievement Spelling Skill 2]. Arnhem, the Netherlands: Cito.

van der Sijs, N. (1996). *Leenwoordenboek: De invloed van andere talen op het Nederlands* [Loan-word dictionary: The effect of other languages on Dutch]. Den Haag, the Netherlands: SDU.

van Leerdam, M., Bosman, A. M. T., & Van Orden, G. C. (1998). The ecology of spelling instruction: Effective training in first grade. In P. Reitsma & L. Verhoeven (Eds.), *Problems and interventions in literacy development* (pp. 307–320). Dordrecht, the Netherlands: Kluwer Academic.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1269–1291.

Van Orden, G. C., & Kloos, H. (2005). The question of phonology and reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 61–78). Oxford, England: Blackwell.

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.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 27*(A), 635–657.

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.

Watt, S., Jokel, R., & Behrmann, M. (1997). Surface dyslexia in nonfluent progressive aphasia. *Brain and Language, 56,* 211–233.

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments & Computers, 28,* 504–515.

Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology, 12,* 413–430.

Received April 22, 2005
Revision received April 13, 2006
Accepted April 19, 2006 ■