Overruled!: Implicit cues rather than an orthographic rule determine Dutch children's vowel spelling

de Bree, E.; Geelhoed, J.; van den Boer, M.

DOI
10.1016/j.learninstruc.2018.03.006

Publication date
2018

Document Version
Final published version

Published in
Learning and Instruction

Learning and Instruction 56 (2018) 30–41

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Overruled!: Implicit cues rather than an orthographic rule determine Dutch children's vowel spelling

Elise de Bree[a,*], Jan Geelhoed[b], Madelon van den Boer[a]

[a] University of Amsterdam, Research Institute of Child Development and Education, PO Box 15780, 1001 NG, Amsterdam, The Netherlands
[b] Vrije Universiteit Amsterdam, PI Research, De Bascule, Rijksstraatweg 145, 1115 AP, Duivendrecht, The Netherlands

ARTICLE INFO

Keywords:
Spelling
Vowels
Spelling rule
Instruction
Implicit cues

ABSTRACT

This study addressed the question why vowel spelling acquisition is relatively difficult for young Dutch spellers. A spelling rule guides vowel spelling, but implicit cues could also play a role. We evaluated the role of phonology, morphology, and orthography. Grade 1 (N = 113) and 2 (N = 59) children were presented with dictations of real and pseudowords differing in the degree of consistency and familiarity. Correct scores of consistent vowel spelling in Grade 1 and 2 students were near ceiling, whereas those for inconsistent vowels were low, even in Grade 2 children, who have had explicit instruction of the spelling rule. Correct scores were affected by phonological and morphological consistency, and orthographic familiarity. Effects of these implicit cues were even more pronounced in Grade 2. Findings indicate that vowel length spelling is difficult to acquire because the explicit spelling rule is overruled by various sources of implicit information.

1. Introduction

Part of becoming literate is being able to spell. Spelling ability influences both the writing process and the perception of written texts by others (Graham, Harris, & Hebert, 2011). It has furthermore been proposed to influence reading ability, the other essential aspect of literacy (Ehri, 2000; Graham & Hebert, 2011). Spelling ability is thus of great importance for academic success. Although attention to acquisition of spelling is increasing, surprisingly few studies have focused on spelling compared to reading.

Spelling is complex and error prone. How it can best be taught has been an important discussion in the literature. One debate has been whether spelling should be *taught*, that is, whether spelling should be learnt through directly and systematically teaching children the spelling rules, or whether it should be *caught*, that is, whether it should be acquired incidentally, and indirectly. A recent meta-analysis of Graham and Santangelo (2014) found support for better spelling acquisition when spelling was taught rather than caught, pointing towards the importance of instruction in spelling acquisition.

A related (and partly overlapping) debate has been how spelling is learnt. Broadly speaking, a division can be made between models and interpretations of a phase-or-stage-like development (Ehri, 1992; Nunes, Bryant, & Bindman, 1997), which assume that increasingly more knowledge is used in spelling and that this knowledge leads to abrupt changes in spelling. On the other hand are models that assume that different linguistic and orthographic cues play a role from early development onwards (Deacon, Conrad, & Pacton, 2008; Treiman & Kessler, 2014). The currently dominant interpretation is that multiple implicit cues contribute to spelling outcomes as well as instruction (Treiman & Kessler, 2014). However, not much is known about the influence of these different sources of information on spelling outcomes. In this study, we compare vowel spelling in Dutch Grade 1 and 2 children, referring to the phases before and after instruction of the vowel spelling rule. We assess the influence of different sources of implicit cues.

Vowel spelling is dependent on different elements. For instance, vowel duration can influence spelling: errors can occur if the phonetic distinction of long and short vowels is not clear-cut (e.g. Lehtonen & Bryant, 2005; Nag, Treiman, & Snowling, 2010). Related, Landerl (2003) found that in contrast to good spellers, German poor spellers showed poorer vowel categorization. Furthermore, vowel spelling is easier if it is more predictable, when there is a straightforward phonology-orthography conversion, than when it is less predictable, when there is no 1:1 mapping between phonology and orthography, as in English (Wimmer & Landerl, 1997). Young English spellers have been found to initially rely on phonological information and only later shift to considering orthographic information in vowel spelling (Treiman & Kessler, 2006; Varnhagen, Boechloer, & Steffler, 1999). Varnhagen et al. (1999), for instance, found that children initially spell the vowel /ɑ/ with 'o', even when this is incorrect (correct: *sock*, incorrect: *swap*).

* Corresponding author.
E-mail addresses: e.h.debree@uva.nl (E. de Bree), jwgeelhoed@gmail.com (J. Geelhoed), m.vandenboer@uva.nl (M. van den Boer).

https://doi.org/10.1016/j.learninstruc.2018.03.006
Received 22 May 2017; Received in revised form 19 March 2018; Accepted 23 March 2018

They gradually also incorporate orthographic knowledge, by spelling the /ɑ/ words with the 'a' when needed. In addition, acquisition of orthographic conventions is necessary for acquiring the vowel spellings, because vowel spellings are also influenced by the surrounding word-or-syllable context (e.g. Landerl, 2003) and, vice versa, can influence the spelling of the surrounding consonants (e.g., Deacon, Leblanc, & Sabourin, 2011). Finally, vowel spelling can be influenced by morphological consistency: children spell morphemes better when the underived (base) word does not demand a change in the inflection. Deacon et al. (2011) found that children spelled both words with long (*skater*) and short vowels (*knitter*) with one consonant (correct *skater* and incorrect **kniter*). The children thus relied on this consistency even when the orthographic convention of doubling the consonant when preceded by a short vowel pointed them to the correct spelling (*knitter*).

In the present study, the focus was on Dutch vowel spelling of long vowels. Long vowels can be spelled two ways, either with two graphemes (vowel /a:/ in *maan* (moon)) or with one (vowel /a:/ in *water* (water) or *manen* (moons)). The former spelling pattern relies on consistent phoneme to grapheme conversion. The spelling of a long vowel with one grapheme is dependent on an orthographic rule of vowel degemination. This rule demands the conversion of a phonological representation (from /a:/ to 'a') on the basis of a spelling rule. It is generally taught as the letter-thief rule, as one of the letters for the vowel is 'stolen' when the vowel occurs at the end of an open syllable. This is the case both for mono-morphemic words, such as *water*, and words that are inflected and have a monomorphemic counterpart (*manen*).

The spelling convention applies only to vowels A/E/O/U, not for those with diphthongs, such as *buik-buiken* (bellies) and not for words with closed syllables (e.g. *taart-taarten*, cakes). In Grade 1, children are taught transparent spellings (*tas* 'bag', *maan*). In the second half of Grade 2, they are taught the vowel degemination rule (*manen*, *water*). The rule is not fully acquired at least until Grade 4 (Landerl & Reitsma, 2005). Errors made are often writing two graphemes for the long vowel, i.e, **waater* (Landerl & Reitsma, 2005). Studies have looked into improving spelling instruction to increase vowel spelling, but to our knowledge, there are no studies investigating the joint role of phonological, morphological and orthographic cues potentially contributing to vowel spelling. Prior to explicit instruction of this rule, these cues might already affect spelling and might continue to do so even when the rule has been taught explicitly. This is what we aimed to investigate.

Two interventions to promote learning of this Dutch spelling pattern have been reported. They focused on learning through explicit or implicit instruction. Hilte and Reitsma (2011) taught second grade children the rule of vowel length spelling before they were taught this rule in the school curriculum They compared a control group (education-as-usual) and four intervention groups, divided into rule (present or absent) and number of items during intervention. The intervention groups outperformed the control group, indicating that both implicit and explicit instruction were beneficial. The intervention consisting of an expanded set of training items with presentation of the spelling rule was most supportive in learning, although accuracy was still not near ceiling. At first glance, these findings suggest that it is the explicit instruction that aided learning most. However, implicit exposure and analogy to a larger set of targets could also have contributed to this higher outcome, as number and frequency of exemplars presented during intervention affected learning. Because of this ambiguity and because this study collapsed findings on vowel spelling in open and closed syllables further investigation into the role of implicit and explicit cues in vowel spelling is needed.

Kemper, Verhoeven, and Bosman (2012) conducted a short-term intervention study in first grade on vowel length spelling with an explicit, implicit and control condition (education-as-usual). Both implicit and explicit instruction were more effective than the control condition in children with average-to-above-average spelling abilities. In the explicit condition, there was transfer to untrained words, but not to pseudowords. In the implicit condition, there was no transfer. Kemper et al. (2012) take these findings to mean that explicit instruction rendered knowledge that was of higher quality than implicit instruction. Furthermore, they interpret their findings to align with a stage-based approach of spelling; rules are learned in stages and generalization to abstract knowledge about spelling and correct use of these rules takes time and effort.

The findings of both Hilte and Reitsma (2011) and Kemper et al. (2012) indicate that, although phonologically inconsistent vowel length spelling can be learnt, spelling scores remain relatively low. In the present study, we aim to assess which cues might cause the difficulty in spelling this pattern. As the aim of the two intervention studies was to assess whether spelling could be improved, the number of trained items was quite low (5 phonologically inconsistent targets in Hilte and Reitsma, and 8 in Kemper et al., 2012) as was the variety of the targets. For instance, there were no targets that are phonologically consistent. Such targets would provide a baseline of vowel spelling. It is important to confirm that children are able to spell vowels correctly when phonology and orthography are consistent, as this precludes other difficulties, such as difficulties in perception of vowel duration. As we outline the potential contribution of different sources of information contributing to vowel length spelling below, the case for an analysis into these different cues is made.

Vowel distinction could influence spelling of the long vowel. In Dutch, the distinction between short and long vowels is based on duration as well as spectral composition. Children are able to classify /A/ as in man and /a:/ as in *maan* (Gerrits, 2001). Furthermore, five-year-olds have generally acquired the phonemes /A/ and /a:/ (Beers, 1995). Spelling of consistent long vowels has not been found to be problematic in Grade 2-to-4-children (Landerl & Reitsma, 2005). Although difficulties are not anticipated in vowel duration, spelling of both consistent short and long vowels needs to be assessed before assessing inconsistent long vowel spelling.

Morphological consistency can also contribute to vowel spelling. Phonologically inconsistent long vowel spelling occurs in monomorphemic (uninflected) targets, such as *water* but also in inflected words, such as *manen* ('moons') or *koken* ('to cook'). Based on findings that show a preference for morpheme consistency (Deacon et al., 2011), it can be anticipated that young spellers might make more errors in targets that have an uninflected counterpart. They might aim to be consistent in their (phonology and) morphology (*maan*-**m<u>aa</u>nen*) rather than use the correct orthographical rule. This pattern has indeed been reported by Landerl and Reitsma (2005). They tested Grade 2 to 4 children's ability to spell and identify plural words and pseudowords. Both spelling and identification of the correct spelling was difficult. Morphological and phonological consistency influenced the outcomes more than the orthographic rule. This finding is reflected in a study by Verhoeven, Schreuder, and Baayen (2006), who found that Grade 3 and 4 children as well as adults were slower and less accurate in lexical decision when the target was a word that undergoes vowel change due to pluralisation (i.e., *manen*) than when there were no changes or other changes. Although this provides important information on the role of morphological consistency in long vowel spelling, a comparison of vowel spelling between monomorphemic (*water*) and inflected targets (*manen*) was not made. This is relevant for assessing the influence of a morphological pattern.

Furthermore, knowledge of the meaning of the word might aid spelling. Studies have reported weak to moderate relationships between vocabulary knowledge and spelling for beginning spellers (Kim, 2010; Sénéchal & LeFevre, 2002). This could mean that the vowel in a word such as *water* ('water'), a word familiar in meaning and phonology to children might be easier to spell than *krater* ('crater'), which is less frequent and therefore less likely to be familiar to children. It would also imply that pseudowords, which lack meaning, are more difficult to spell than words that are familiar.

Finally, orthographic exposure could affect vowel spelling. Previous exposure to the orthographic form of the word as a whole might

influence spelling of the vowel. For instance, both the word *water* ('water') and *bami* ('bami') are likely to be familiar to children in Grade 1. However, they are more likely to have encountered *water* orthographically than *bami*. This exposure might lead to better spelling of *water* than of *bami* and of pseudowords. Additionally, orthographic patterns *within* words might influence spelling. For instance, Pacton, Perruchet, Fayol, and Cleeremans (2001) found that French children from Grade 1 onwards were aware that 1) certain consonants can be doubled in writing whereas others cannot, 2) that vowels cannot be doubled and 3) that consonant doubling was dependent on the position in the word. With respect to inconsistent long vowel spelling, this matter has not been addressed yet. The first question is whether the frequency of occurrence of the vowel in general aids spelling. Frequent exposure might aid in spelling in general: vowels that are most frequent might be spelled better than those that children see less frequently. Furthermore, the occurrence of inconsistent long vowels could affect the spelling. Vowels that occur more often in this open syllable position might be spelled better than those that do not. Finally, the relative frequency with which long vowels are spelled with one grapheme could influence spelling. If inconsistent long vowel spelling (e.g. *manen*, *water*) is relatively more frequent than consistent long vowel spelling (*maan*), then this could render better spelling of inconsistent syllables than if the pattern is the other way around.

In sum, acquisition of the rule of phonologically inconsistent vowel spelling demands knowledge of an explicit orthographic rule which violates phonological and morphological consistency. Word familiarity and orthographic exposure could facilitate spelling. These different types of cues need to be taken into account to answer the question of whether the rule has been acquired.

1.1. Present study

We compared long vowel spelling in Grade 1 and 2 children, referring to the phases before and after explicit instruction on the rule. Specifically, we assess the influence of vowel discrimination through dictation of consistent long and short vowels. We also examine effects of phonological and morphological consistency and orthographic exposure. The expectation was that children in Grade 2 outperformed those in Grade 1, as they have been taught the spelling rule and have had more exposure to orthography, more experience with spelling and know more words than the Grade 1 children. Accordingly, the difference between vowel spelling in phonologically consistent and inconsistent targets should be smaller in this group. However, children in Grade 1 might already rely on implicit cues in spelling the different vowels, which would point to gradual implicit acquisition of the vowel pattern.

With respect to the different sources of information for vowel spelling, there are three expectations. First, correct vowel spelling is higher in real words than in pseudowords; real words have semantic content and children can have encountered them as a whole. Second, correct vowel spelling is higher in phonologically consistent than in inconsistent words. In other words, vowel discrimination and consistent phoneme-grapheme-conversion are not deemed problematic. Third, vowel spelling is affected by complexity (word meaning, orthographic exposure, and morphology): Vowels in words that are unfamiliar in meaning are more difficult to spell than those words that are familiar. Vowels in words that have not been presented orthographically to the children are more difficult to spell than those in words they have encountered in writing. Vowels in inflected words that have an uninflected counterpart are more difficult to spell than those that have no counterpart. The expectation would thus be that the ranking is: consistent vowel spelling (*tas*, *maan*) > inconsistent vowel spelling (*water*, *manen*). Inconsistent vowel spelling is divided into familiar words encountered in orthography (*water*) > familiar words not encountered in orthography (*bami*) > unfamiliar words not encountered in orthography (*krater*, crater), and familiar plurals that demand violating phonological and morphological consistency (*manen*). For pseudowords, the assumption is that vowels in consistent targets (*naag*) are spelled correctly more often than pseudowords presented as a monomorphemic target (*fatel*), which in turn are spelled better than targets that are presented as inflected (*nagen*).

It is an open question whether vowel frequency influences vowel spelling. The general frequency of occurrence of the vowel (A/E/O/U) might influence vowel spelling, as might the *relative* occurrence of consistent and inconsistent vowel spellings.

2. Method

2.1. Participants

The sample consisted of 113 children (45 girls) in Grade 1, mean age 6 years 10 months (SD = 4.1 months). There were 59 children (27 girls) in Grade 2, mean age 7 years 9 months (SD = 5.4 months). The children, who came from six different schools, were all fluent speakers of Dutch. They were tested in the spring of the school year (March-May). The Grade 1 sample contained 9% of children (11/113) who also spoke another language at home; the Grade 2 sample 10% (6/59).

2.2. Instruments

2.2.1. Dictation

Vowel spelling was measured through a dictation task. Stimuli consisted of real and pseudowords. With respect to the *real words*, there were items with four different vowels: A, E, O, and U, and eight target types. The first four types had consistent vowel spelling. These were divided in words with a long vowel spelled with corresponding two graphemes (e.g., *maan,* moon), words with a short vowel spelled with one grapheme (*tas*, bag), compounded words with a long vowel spelled with corresponding two graphemes (*spaarpot,* piggy bank) and compounded words with a short vowel spelled with one grapheme (*rugzak*, backpack). There were also four types that were phonologically inconsistent as they contained vowels in open syllables. These were divided in inflected familiar words (*manen,* moons), monomorphemic familiar words (*water,* water), monomorphemic familiar words that children have not encountered often in their reading (*bami*, bami), and monomorphemic words unfamiliar to children in spoken and written form (*krater*, crater).

The aim was to include two items for each vowel per word category, with 2 (items) x 4 (vowels) x 8 (word types) = 64 vowel instances. The total number of words, however, was 58. As some compounds contained both short and long vowels, 58 items was enough to obtain at least two spellings of each short and long vowel in a compound. The real word stimuli are presented in Appendix 1. One short vowel U target (*bus*, bus) was accidentally omitted from the dictation and one additional U target was added in the category monomorphemic familiar words that children have not encountered often in their reading (*muziek*, music). It should be noted that the omission involved a consistently spelled vowel, which is unlikely to have affected the results. Furthermore, the results were the same when the additional target (*muziek*) was excluded.

For want of a better child-directed measure, we used frequencies of CELEX (Baayen, Piepenbrock, & Van Rijn, 1993) and the age of acquisition (AoA) ratings of Brysbaert, Stevens, De Deyne, Voorspoels, and Storms (2014). These were available for the majority of words. We also calculated frequencies of Subtlex (Keuleers, Brysbaert, & New, 2010). As these were highly correlated to those of CELEX (r = 0.932, p < 001), only those of CELEX were reported. Mean frequencies of the targets and age of acquisition per vowel category (A/E/O/U) are presented in Table 1. There is no main effect of word frequency and AoO for vowel type (CELEX: $F(3, 56) = 1.446$, $p = .239$; AoA: $F(3,60) = 0.514$, $p = .674$).

Mean frequencies per word type (8 different types) are presented in

Table 1
Frequencies and age of acquisition for the dictation words per vowel category.

Words with vowel	CELEX frequencies per million	AoA mean age
A	121.5 (222.7)	6.2 (2.7)
E	65.2 (110.1)	5.6 (2.4)
O	36.4 (40.6)	6.7 (2.4)
U	38.9 (42.2)	6.2 (2.3)

Table 2
Frequencies and AoA for the experimental words per word category.

Consistency	Word category	CELEX frequencies	AoA
Consistent	Long vowel (*maan*)	147.2 (155.5)	4.6 (2.0)
	Short vowel (*tas*)	59.4 (56.7)	5.0 (0.6)
	Long vowel: Compound (*spaarpot*)	9.7 (7.0)	7.0 (1.4)
	Short vowel: Compound (*rugzak*)	9.3 (5.9)	6.8 (1.3)
Inconsistent	Inflected familiar word (*manen*)	194.7 (300.1)	5.1 (2.6)
	Monomorphemic familiar word (*water*)	90.5 (108.6)	5.0 (1.0)
	Monomorphemic familiar word in meaning, unfamiliar in orthography (*bami*)	28.1 (38.5)	7.6 (2.1)
	Monomorphemic unfamiliar word in meaning and orthography (*krater*)	7.3 (7.5)	8.8 (3.8)

Table 2. Given the small sample size, statistical comparison is not useful. The findings do show a decrease in frequency and an increase in AoA in the inconsistent condition, as planned. Note that the frequency of category 5 (*manen*) words is higher than that of the other categories. Their AoA is as low as that of the consistent vowels. The frequency of the words with consistent vowels in compounds is low, and the AoA of these targets is relatively high.

Stimuli were embedded in short sentences. First, the target was pronounced, followed by the sentence, and the task to spell the target word. An example is "*Maan. Vannacht is het volle maan. Spel 'maan'*". ["Moon. Tonight it is full moon. Spell 'moon'"]. Children's vowel spellings were scored as correct or incorrect. Proportion correct target vowel spelling was calculated; other errors (e.g. *rugzak* as **rugsak*) were ignored. The items were presented in three dictation sessions, in which vowels and word types were balanced. Reliabilities between dictations 1, 2, and 3 ($\alpha = 0.862$) and within dictations were sufficient (dictation 1: $\alpha = 0.831$, dictation 2: $\alpha = 0.796$, dictation 3: $\alpha = 0.770$). However, some items were removed for calculation of these reliabilities as they all had a perfect vowel score and therefore zero variance (dictation 1: *sok* 'sock' and *zes* 'six', dictation 2: *leesles* 'reading lesson', dictation 3: *pen* 'pen' and *tas* 'bag'). These ceiling scores were only found for consistent vowels. Note that these scores were not omitted in calculating children's proportions correct vowel spelling, only for calculating the reliabilities.

With respect to the *pseudowords*, spelling of the same four vowels (A/E/O/U) was assessed. There were three target types: pseudowords with a long vowel spelled with corresponding two graphemes (e.g., *snaap*), inflected pseudowords with long vowels that lose one vowel grapheme to morphological inflection (*snapen*), and monomorphemic pseudowords with long vowels spelled with one grapheme (*fatel*). The category of short vowels was omitted in the pseudoword dictation to reduce demands on the children. There were two items for each vowel per word category, so 2 (items) x 4 (vowels) x 3 (target types) = 24 pseudoword items, see Appendix 2. Similar to words, pseudowords were embedded in short sentences. First, the target was pronounced, followed by the sentence, and the task to spell the target word. An example is "Snaap. Ik vind die snaap lelijk. Spel 'snaap'". ["Snaap. I think that snaap is ugly. Spell 'snaap'"]. Children's vowel spelling was scored as correct or incorrect. Proportions vowel spelling correct were calculated. Reliability within the pseudoword dictation was good ($\alpha = .865$).

2.3. Procedure

Testing took place in the spring of the school year (March-May). The dictations were conducted in four class-based sessions, each taking approximately 40 min. Sessions 1 to 3 included real word dictation and session 4 the pseudowords. Sessions 1 to 3 were conducted at all schools, but session 4 only at 5 of these schools. An individual test session tapping different cognitive skills was also conducted for all children (except 30 students in Grade 1) as part of a study into cognitive correlates of spelling reported elsewhere (see de Bree & van den Boer, in preparation). These tasks included word reading, phoneme awareness, rapid automatized naming, and visual attention span. Data was collected by four trained assistants. Teachers were present during the dictations. The study was approved by the ethics committee of the University of Amsterdam (2015-CDE-4151).

2.4. Analysis

In order to assess whether vowel type and target type influence spelling outcomes and whether Grade affects spelling outcomes, repeated measures ANOVAs were conducted on the real word and pseudoword results. Within-subject factors were vowel (A/E/O/U) and target type (real words: 8, pseudowords: 3). Grade was the between-subjects factor (Grade 1, 2). Huynh-Feldt corrections were applied in cases of asphericity of the data. Bonferroni corrections were applied in pairwise comparisons.

3. Results

3.1. Data screening

Dictation data of real words was available of 113 Grade 1 and 59 Grade 2 students. Mean proportions vowel correct per child were calculated and used for the analyses. There were no outliers on mean proportion vowel correct and no multivariate outliers on correct scores for each category. However, Shapiro Wilk results on the distribution of the mean proportion correct of the real words is significant ($SW = 0.944$, $p < .001$). As repeated measures ANOVAs are robust and the sample size is sufficient, this was not considered problematic.

Dictation data of pseudowords was available of 82 Grade 1 and 54 Grade 2 students. There were no outliers on mean proportion vowel correct. However, there were two indications against normality of the data, as 1) there were 3 multivariate outliers on correct scores for each category and 2) Shapiro Wilk results on the distribution of the mean proportion vowel correct is significant ($SW = 0.839$, $p < .001$). As analyses with and without the outliers rendered the same results, outcomes of the full dataset were reported.

In order to confirm that the pattern of findings on the real words was the same for the entire dataset (the set reported here) and the subset containing the data of children who completed both the word and pseudoword dictation, we ran all analyses on vowel spelling in real words for the entire dataset and the subset. The pattern of findings remained the same and is therefore not addressed further.

3.2. Vowel spelling in words and pseudowords across grades

Spelling of words and pseudowords was compared for the children who completed both types of dictations. Results are presented in

Table 3
Mean proportion correct vowel spelling per word type and Grade.

	Grade 1	Grade 2
Real words	.55 (.11)	.71 (.14)
Pseudowords	.52 (.22)	.51 (.21)

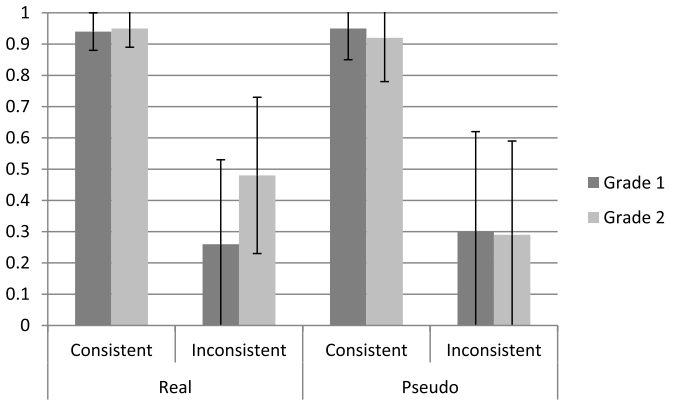

Fig. 1. Proportion vowel spelling correct (SD) per Grade.

Table 3. The first question, whether Grade 2 children obtained higher vowel correct scores than Grade 1, is examined in a repeated measures ANOVA with Word type as within-subjects variable and Grade (Grade 1, 2) as between-subjects variable. Results showed an effect of word type, F (1, 134) = 48.157, $p < .001$, $\eta_p^2 = 0.266$), Grade, $F(1, 134) = 6.252$, $p = .142$, $\eta_p^2 = 0.045$, and an interaction between word type and Grade, $F(1, 134) = 29.278$, $p < .001$, $\eta_p^2 = 0.180$. The interaction is due to the fact that there is a large increase in proportion correct for real words from Grade 1 to Grade 2, but not for pseudowords. In sum, vowels in words are easier to spell than those in pseudowords. This gap increases in Grade 2.

3.3. Vowel spelling for phonologically consistent and inconsistent targets

Spelling results divided by consistency are presented in Fig. 1. For words, a repeated measures ANOVA with proportion vowel correct per Consistency (consistent, inconsistent) and Grade (1,2) shows a main effect of consistency, $F(1, 170) = 697.444$, $p < .001$, $\eta_p^2 = 0.804$, an effect of Grade, $F(1, 170) = 27.060$, $p < .001$, $\eta_p^2 = 0.137$, and an interaction between consistency and Grade, $F(1, 170) = 21.713$, $p < .001$, $\eta_p^2 = 0.113$. The main effects indicate that consistent vowels are spelled correctly more often than inconsistent ones and that Grade 2 children show higher correct scores than those in Grade 1. The interaction is due to the fact that there is an increase in vowel spelling correct for inconsistent targets between Grade 1 and Grade 2, but not for consistent targets, the scores of which are already high in Grade 1.

For pseudowords, there was also a main effect of consistency, $F(1, 134) = 485.066$, $p < .001$, $\eta_p^2 = 0.784$, with high vowel spelling correct of consistent targets compared to much lower vowel spelling correct of inconsistent targets. There was, however, no effect of Grade, $F(1, 134) = 0.432$, $p = .512$, and no interaction between Grade and consistency, $F(1, 134) = 0.031$, $p = .861$. For both words and pseudowords, the vowel errors for inconsistent long vowels were primarily spelling the vowel with two graphemes (words: 98%, pseudowords: 96%, with similar percentages for both grades separately). In sum, words with consistent vowel spellings are easier to spell than those with inconsistent ones. Inconsistent vowel spelling of words increases over time, but this is not the case for inconsistent vowel spelling of pseudowords.

3.4. Vowel spelling for different categories

Table 4 shows the mean proportion vowel correct for the different categories of the real words. Scores for consistent targets are high; those for inconsistent targets are much lower and are characterized by substantial standard deviations. For words, a repeated measures ANOVA shows an effect of category, $F(7, 1183) = 476.666$, $p < .001$, $\eta_p^2 = 0.738$, of Grade, $F(1, 170)$ 25.445, $p < .001$, $\eta_p^2 = 0.131$, and an interaction between category and Grade, $F(7, 1183) = 24.747$, $p < .001$, $\eta_p^2 = 0.128$. The interaction was unpacked through a repeated measures ANOVA for each Grade separately. Similar to the overall ANOVA, there was an effect of category for both Grade 1, $F(7,784) = 478.075$, $p < .001$, $\eta_p^2 = 0.810$, and Grade 2, $F(7, 399) = 133.357$, $p < .001$ $\eta_p^2 = 0.733$. In both grades, there was no difference between different categories of consistent short vowel words (*tas*, *rugzak*; $p = 1.0$ in all cases) and between different categories of consistent long vowel words (*maan*, *spaarpot*; $p > .05$). In Grade 1, consistent short vowels were spelled better than consistent long vowels, although proportion vowel correct was high for both. This difference was not significant in Grade 2. In Grade 1, scores between inconsistent uninflected familiar words (*water*) and inconsistent unfamiliar words (*krater*) did not differ ($p = 1.0$). In Grade 2, all comparisons for inconsistent vowels were significantly different. Ranking in Grade 1 is thus *tas, rugzak* (consistent short vowels) > *maan, spaarpot* (consistent long vowels) > *water, krater* > *bami* > *manen* (long vowel inflected). In Grade 2, this is *tas, rugzak* = *maan, spaarpot* > *water* > *bami* > *krater* > *manen*. In other words, the scores in Grade 2 are higher than in Grade 1 and show more differentiation of target type than in Grade 1.

Table 4
Proportion vowel spelling correct (SD) for word type, category, and Grade.

Consistency	Word category	Real		Pseudo	
		Grade 1	Grade 2	Grade1	Grade 2
Consistent	Long vowel (*maan*)	.92 (.09)	.92 (.13)	.95 (.10)	.93 (.14)
	Short vowel (*tas*)	.98 (.05)	.99 (.02)		
	Long vowel: Compound (*spaarpot*)	.89 (.13)	.90 (.02)		
	Short vowel: Compound (*rugzak*)	.96 (.08)	.99 (.03)		
Inconsistent	Inflected familiar word (*manen*)	.19 (.26)	.25 (.26)	.29 (.31)	.26 (.32)
	Monomorphemic familiar word (*water*)	.30 (.31)	.68 (.31)		
	Monomorphemic familiar word in meaning, unfamiliar in orthography (*bami*)	.25 (.30)	.54 (.29)		
	Monomorphemic unfamiliar word in meaning and orthography (*krater, waber*)	.29 (.29)	.41 (.31)	.32 (.34)	.31 (.31)

With respect to pseudowords, there is an effect of category, $F(2, 268) = 433.986$, $p < .001$, $\eta_p^2 = 0.764$, with consistent > inconsistent uninflected > inconsistent inflected, i.e., *naag* > *waber* > *nagen* (all $p < .001$). There is no effect of Grade, $F(1, 134) = 1.134$, $p = .307$, and no interaction between Grade and Category, $F(2, 268) = 0.025$, $p = .975$.

Finally, in order to test whether real words that are unfamiliar in meaning and orthography (*krater*) and pseudowords differ from each other, a paired samples t-test was conducted. There was no difference between unfamiliar real words and pseudowords $t(101) = -1.701$, $p = .092$. In sum, consistent words are easier to spell than inconsistent ones. In Grade 2, differences between target types become more pronounced as influences of word and orthographic familiarity surface. Consistent pseudowords are also easiest to spell. For pseudowords, the three tested target types differ from Grade 1 onwards (consistent > inconsistent uninflected > inconsistent inflected).

3.5. Vowel spelling for different vowels

Prior to turning to the outcomes on spelling per vowel, we obtained the occurrences of the different vowel graphemes of interest. On the basis of the Age of Acquisition list by Brysbaert et al. (2014), we selected the words with an AoA up to 9 years of age, rendering 9423 words. The 159 words that only occur in Flemish Dutch, not in the

Table 5
Proportion correct vowel spelling (SD) for vowel type and Grade.

Word type	Vowel	Grade 1	Grade 2
Real words	A	.64 (.15)	.73 (.14)
	E	.57 (.15)	.69 (.16)
	O	.60 (.17)	.72 (.15)
	U	.57 (.16)	.69 (.16)
Pseudowords	A	.61 (.29)	.55 (.23)
	E	.45 (.21)	.48 (.23)
	O	.43 (.26)	.41 (.29)
	U	.59 (.20)	.57 (.24)

Netherlands, were excluded; 9264 words remained. These words were scored for the presence of the graphemes. Per vowel there were three categories: 1) the consistent spelling of the short vowel with one grapheme (e.g. /A/ as 'a' as in *man*), 2) the consistent spelling of the long vowel with two graphemes (e.g. /a:/ as 'aa' as in *maan*) and 3) the inconsistent spelling of the long vowel one with one grapheme (e.g. /a:/ as 'a' as in *manen*). Findings on the relative frequency of each vowel in relation to the other three vowels are reported in Appendix 3A. These counts show that in the AoA corpus, the vowel A is more frequent than E and O, which are more frequent than U. Table 5 displays the vowel correct scores per vowel in the children's data.

We examined whether the general frequency of occurrence of the vowel affected vowel spelling. For words, a repeated measures ANOVA with vowel (A, E, O, U) as within-subjects variable and Grade (Grade 1,2) as between-subjects variable shows an effect of vowel category, $F(3, 510) = 13.714$, $p < .001$, $\eta_p^2 = 0.075$, and Grade, $F(1,170) = 24.015$, $p < .001$, $\eta_p^2 = 0.124$, but no interaction between the two, $F(3,510) = 1.283$, $p = .280$. Grade 2 children obtained higher scores than Grade 1. Pairwise comparisons of vowel type show that A is spelled correctly more often than E/O/U ($p < .01$). O is spelled significantly better than E ($p < .01$). U, the least frequent vowel, does not differ from E and O ($p > .05$). Excepting U, the general pattern mimics the frequency in the AoA count.

For pseudowords, the repeated measures ANOVA showed a main effect of vowel, $F(3, 402) = 47.617$, $p < .001$, $\eta_p^2 = 0.262$, but no effect of Grade, $F(1, 134) = 0.348$, $p = .556$, or interaction between vowel and Grade, $F(3, 402) = 2.377$, $p = .078$. Pairwise comparisons led to a ranking of A, U > E, O. The spelling of U is unexpected in light of the frequency counts. Thus, for both words and pseudowords, vowels that are more frequent are spelled correctly more often, with the exception of high accuracy for the relatively infrequent U.

On the basis of the AoA corpus, we also calculated the distribution of consistent short vowel, consistent long vowel and inconsistent long vowel spelling per vowel (Appendix 3B). These vowel counts (Appendix 3B) show that *within* the four vowels, the consistent short vowel always occurs most often (*tas*), followed by the inconsistent long vowel spellings (*manen*) and the consistent long vowel spellings (*maan*). The difference between inconsistent long vowel spellings and consistent long vowel spellings is most pronounced for long vowels /o:/ and /y:/ (the phoneme for 'uu') and almost absent for/a:/, which is spelled approximately equally often as 'aa' and 'a'. If this orthographic information influences children's spelling, then one effect could be that children spell long vowels with one grapheme. It could also be the case that the effect of consistency might not be the same across vowels, as this spelling with one grapheme would occur more often for /o:/ and /y:/ than for /a:/. Results of the children's data are presented in Fig. 2.

A repeated measures ANOVA on real words, with Long vowel (/a: e: o: u:/) and Consistency (consistent (*maan*), inconsistent (*manen*)) as within-subjects variable, and Grade as between-subjects variable shows a main effect of vowel, $F(3, 495) = 4.499$, $p = .004$, $\eta_p^2 = 0.027$, of consistency, $F(1, 170) = 747.818$, $p < .001$, $\eta_p^2 = 0.820$, and of Grade, $F(1, 170) = 23.007$, $p < .001$, $\eta_p^2 = 0.123$. The interactions between vowel, category and Grade, $F(3, 495) = 1.063$, $p = .364$, and

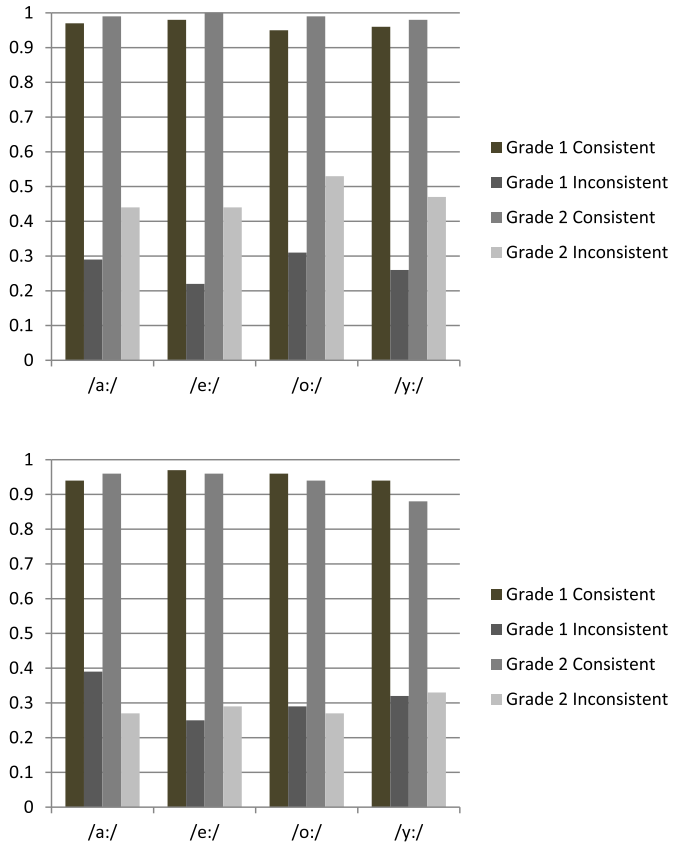

Fig. 2. Proportion Correct Vowel Spelling for Long Vowels per Vowel Type, Consistency, and Grade. Top panel presents findings for real words; bottom panel for pseudowords.

vowel and Grade, $F(3, 495) = 1.934$, $p = .125$ are not significant. There are significant interactions between consistency and Grade, $F(1,170) = 16.487$, $p < .001$, $\eta_p^2 = 0.091$, and between vowel and consistency, $F(3, 495) = 11.033$, $p < .001$, $\eta_p^2 = 0.364$. The consistency and Grade interaction is due to the larger discrepancy between consistent and inconsistent long vowel spelling in Grade 1 compared to Grade 2. The vowel and consistency interaction is due to an overall higher correct score of the consistently spelled long vowels compared to the inconsistent long vowels, with a smaller difference for /o:/. Thus, the frequency within the vowels of the AoA counts is not replicated in terms of consistency, as these counts showed that inconsistent long vowel spelling was more frequent than consistent long vowel spelling. Nevertheless, the findings on spelling of long vowel /o:/ do relate to these frequency findings, as the difference between inconsistent long vowel spelling and consistent long vowel spelling was larger for/o:/ than for the other vowels. In children's spelling, there is a relatively high proportion correct for the/o:/in inconsistent words.

For pseudowords, there is no effect of vowel, $F(3, 402) = 1.135$, $p = .335$, Grade, $F(1, 134) = 0.522$, $p = .471$, and no interaction between vowel and Grade, $F(3, 402) = 2.031$, $p = .109$, and consistency and Grade, $F(1, 134) = 0.014$, $p = .906$. There is a main effect of consistency, $F(1, 134) = 478.344$, $p < .001$, $\eta_p^2 = 0.781$, as well as a significant three-way interaction between vowel, consistency and Grade, $F(3,402) = 4.165$, $p = .016$, $\eta_p^2 = 0.030$. The three-way interaction is unpacked through separate repeated measures ANOVAs for Grade 1 and Grade 2. In Grade 1, a main effect of vowel is found, $F(3, 243) = 4.182$, $p = .009$, $\eta_p^2 = 0.049$, but not for Grade 2, $F(3, 159) = 0.325$, $p = .807$. In Grade 1, long vowel spelling of /a:/ is significantly higher than of /e:/ ($p = .015$). In both Grades, a main effect of consistency (Grade 1: $F(1, 81) = 297.316$, $p < .001$, $\eta_p^2 = 0.786$; Grade 2: $F(1, 53) = 201.126$, $p < .001$, $\eta_p^2 = 0.791$), and an

interaction between vowel and consistency (Grade 1: $F(3, 243) = 6.607$, $p < .001$, $\eta_p^2 = 0.075$; Grade 2: $F(3, 159) = 3.133$, $p = .030$, $\eta_p^2 = 0.056$) are found. In Grade 1 this is due to a relatively low score of /a:/ in the consistent, but a relatively high score in the inconsistent targets, and the mirror image for /e:/, and in Grade 2 this is due to the relatively low score of /o:/ and /y:/ in the consistent targets in contrast to the relatively high score in inconsistent targets. Similar to the real words, the frequency within the vowels of the AoA counts is not replicated in terms of consistency, as consistent long vowel spelling is correct more often than inconsistent long vowel spelling, opposite to the AoA frequencies. However, the Grade 2 findings on spelling of /o:/ and /y:/ do match these frequencies.

In sum, the word and pseudoword results on vowel proportions correct generally align with the overall frequency of occurrence of written vowels, except for U. *Within* the vowels, the findings on vowels and consistency (consistent > inconsistent) are different from the AoA frequencies (consistent < inconsistent). At the same time, there are indications that this distribution does influence children's spelling: vowels with the most pronounced difference between frequency of inconsistent and consistent long vowel spelling are also the ones that children spell correctly more often in the inconsistent condition.

4. Discussion

In this study, we assessed vowel spelling of Dutch children who had and had not yet been taught the spelling rule. We compared vowel spelling in targets that differed in lexicality (real and pseudowords), in phonological consistency (consistent vowel spelling vs inconsistent), in morphology (uninflected vs inflected) and in semantic and orthographic familiarity. Furthermore, we looked into the role that the specific vowel (A, E, O, U) played, especially in terms of frequency of occurrence. The aim was to evaluate whether these factors contribute to children's vowel spelling.

The study yielded five main conclusions. First, spelling of consistent vowels is already at ceiling in Grade 1, even in less frequent compound words, indicating that vowel length discrimination was not problematic, as expected. Second, spelling of inconsistent vowels is difficult, even after the orthographic rule that guides vowel spelling has been explicitly taught to children. Third, spelling is influenced by implicit cues, specifically, morphological consistency, semantic (and subsequent phonological) familiarity and orthographic familiarity. This is most visible in Grade 2 data. Fourth, phonological and morphological consistency also play a role in vowel spelling of pseudowords. Fifth, orthographic vowel frequency influences children's spellings: Written vowels that children encounter more frequently are generally spelled correctly more often and the frequency difference between inconsistent and consistent long vowel spelling impacts on children's inconsistent long vowel spelling.

Acquisition of the spelling rule would entail high correct scores in Grade 2, regardless of target type. This was not the case. Although vowel spelling of inconsistent vowels in Grade 2 was higher than in Grade 1, the proportion correct was still low (max 0.50). The errors that were made were overwhelmingly vowel length errors: spelling of the long vowel with two graphemes when this should be one grapheme (e.g., /a:/ as '*aa*' as in **waater* 'water'). In both grades, there was a gradual decline in correct scores as phonological and orthographic familiarity decreased. An exception to this pattern was inconsistent vowel spelling in morphologically inflected targets: these scores were very low, despite high familiarity as indicated by high lexical frequency and low AoA. The inconsistent vowels in real morphologically inflected words were spelled even more poorly than their pseudoword counterparts. These findings agree with studies that have shown the importance of morphological consistency (root consistency) in children's spellings (Deacon et al., 2011; Landerl & Reitsma, 2005).

Furthermore, Grade 2 students showed *more* sensitivity towards the implicit cues than Grade 1 students. These findings indicate that there is no across-the-board acquisition of the rule but instead an increasing use of implicit information. They mirror results by others. For instance, Van der Ven and de Bree (in revision) found that Grade 6 students, who have been taught the explicit rule on Dutch past tense spelling, rely much more strongly on an implicit cue in past tense spelling than Grade 3 students. Similarly, for English, Kemp and Bryant (2003) found that children's spelling of plurals and non-plural targets were influenced by graphotactic frequencies, co-occurring graphemes, rather than on their knowledge of a morphological rule. For French, Pacton, Fayol, and Perruchet (2005) found that French children relied on a morphological rule as well as graphotactic regularities when spelling derivations; the graphotactic regularities continued to influence this spelling at primary school age and even into adulthood.

Our findings suggest that increased exposure to orthography (i.e., through reading and spelling) leads to perception of more implicit information. It is an open question when and how children learn to overrule phonological and morphological consistency and rely on the explicit spelling rule. More in-depth assessment of the vowel patterns children are exposed to in the school curriculum is needed to establish whether children's eventual acquisition is related to their orthographic exposure or whether the rule is actually acquired explicitly at some point. Tallying the actual orthographic exposure at school is needed to provide information on the true exposure per vowel and target type. This knowledge could then be related to children's spelling abilities of long vowels and the instruction thereof.

We also explored the spelling patterns for different vowels. Children's mean spelling correct scores of words and pseudowords generally agree with the corpus frequency counts, except for U. If children's spellings followed the frequencies entirely, then inconsistently spelled long vowels should be spelled correctly more often than consistent ones, as inconsistent long vowels occur more often. However, this is not the case, as phonological and morphological information affect vowel spelling, as does word frequency. Nevertheless, vowels with the most pronounced difference between frequency of inconsistent and consistent long vowel spelling are also the ones that children spell correctly more often in the inconsistent condition, especially in Grade 2. It thus seems that children track orthographic information at different levels. Contextual frequency might also play a role, such as sequences of graphemes (e.g., -ate with /a:/ (*water*) and -ang with /A/ *bang*, afraid). More detailed analysis of the patterns in the materials that these children are exposed to could provide insight into these contributions to spelling.

Although the current findings indicate that vowel spelling is driven by different implicit sources of information, more research is needed to support this finding as well as to establish the relative contributions of these sources of information. First, this study could be repeated with a larger sample of children across more grade levels to confirm the findings as well as evaluate the development of the different target types for the different vowels over time. Another way is to assess vowel spelling in many more words than those targeted here, and analyze the contributions at the item level. Such assessment is complicated by the demanding nature of the dictations. Child-friendly assessment is required, perhaps through an individual computer game. Another path would be to track children's spelling longitudinally, from Grade 1 to Grade 5, and incorpore the different potential cues across these ages. A third approach would be to assess word recognition, by asking children to select the correct spelling (*manen* - **maanen*) in targets divided by consistency and familiarity. Such a task does not tap spelling as such, but does evaluate orthographic knowledge and, in turn, the knowledge about the spelling rule. A drawback to this approach is that children are exposed to the incorrect target, which is didactically undesirable. Finally, it would be informative to ask children about the rule by asking them, for instance, 'For both *taart* and *manen*, you hear the same vowel. How come they are spelled differently?' It can be assessed whether children who have not acquired the spelling pattern and rule are able to formulate the rule. If the knowledge is implicit, they might not be able

to say why they are spelled differently. Alternatively, children might be able to formulate the rule, but fail to explicitly apply the rule during spelling. Knowing which of the two is more prominent is important to improve their spelling by adapting instruction accordingly.

The findings suggest that acquisition of the orthographic rule for long vowel spelling that demands overruling linguistic knowledge is difficult for Grade 1 and 2 children. These results agree with other findings that indicate that different cues play a role in children's and adults' spelling and spelling acquisition (Deacon et al., 2008; Pacton et al., 2005, 2001; Treiman & Boland, 2017; de Bree, van der Ven, & van der Maas, 2017) and models that account for the contributions of these different sources of information (Treiman & Kessler, 2014). In terms of practical implications, general principles of spelling instruction should be applied in teaching long vowel spelling. For instance, repeated instruction and practice is important. Additionally, because children often make vowel spelling errors, they will often be exposed to the incorrect forms. Therefore, exposure to the correct orthographic form is essential and errors should be prevented by providing direct feedback during instruction. Offering the meaning of the words is also important, as it also aids spelling acquisition. With respect to the rule itself, previous studies have shown that explicit instruction does not necessarily yield better scores than implicit instruction or no instruction (Hilte & Reitsma, 2011; Kemper et al., 2012). This could mean that children memorize the rule, but do not apply it. Instruction and practice could be improved by forcing the children to explicate the rule every time they are asked to spell long and short vowels, as also pointed out by a reviewer. Furthermore, feedback should be provided throughout. Such feedback is especially important for inflected targets (*manen*, *koken*) in relation to their monomorphemic counterpart (*maan*, *kook*).

In sum, our findings show that linguistic information as well as orthographic familiarity contribute to vowel spelling. It is difficult to overrule these implicit cues in favour of an explicit orthographic rule. The findings underscore the importance of assessing spelling patterns in depth and adjusting instruction accordingly, both for theory and educational practice.

Acknowledgements

We are grateful to the children and their teachers for participating. We would like to thank the student assistants involved in data collection. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We would like to thank the reviewers for their suggestions.

Appendix D. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.learninstruc.2018.03.006.

Appendix 1. Overview of Real Word Stimuli

Target	Translation	Consistency	Word type	Vowel	CELEX frequency per million	AoA
maan	moon	Consistent	Long vowel	A	64,9980	5,34
vraag	question	Consistent	Long vowel	A	475,9923	5,00
been	leg	Consistent	Long vowel	E	188,0182	4,09
speel	play	Consistent	Long vowel	E	n.a.	n.a.
droom	dream	Consistent	Long vowel	O	118,9871	5,36
noot	nut	Consistent	Long vowel	O	19,0020	5,54
duur	expensive	Consistent	Long vowel	U	120,0052	6,46
stuur	steer	Consistent	Long vowel	U	44,0048	5,60
tas	bag	Consistent	Short vowel	A	38,0014	5,37
warm	warm	Consistent	Short vowel	A	158,0156	4,71
pen	pen	Consistent	Short vowel	E	19,0020	5,73
zes	six	Consistent	Short vowel	E	129,9870	4,50
sok	sock	Consistent	Short vowel	O	12,9987	4,65
zon	sun	Consistent	Short vowel	O	46,0045	4,35
krul	mark as correct	Consistent	Short vowel	U	12,0005	5,73
bus	motor bus	Consistent	Short vowel	U		
haaksteek	Crochet stitch	Consistent	Long vowel compound	A	n.a.	n.a.
spaarpot	piggy bank	Consistent	Long vowel compound	A	1,0000	5,99
feestdag	national holiday	Consistent	Long vowel compound	E	8,9991	7,90
leesles	reading lesson	Consistent	Long vowel compound	E	5,0001	n.a.
boomhut	treehouse	Consistent	Long vowel compound	O	n.a.	6,47
doolhof	labyrinth/maze	Consistent	Long vowel compound	O	7,0000	6,96
koopman	merchant	Consistent	Long vowel compound	O	16,9981	9,78
buurman	neighbour	Consistent	Long vowel compound	U	19,0020	5,75
vuurwerk	fireworks	Consistent	Long vowel compound	U	5,0003	6,65
buurman	neighbour	Consistent	Short vowel compound	A	19,0020	5,75
feestdag	national holiday	Consistent	Short vowel compound	A	8,9991	7,90
koopman	merchant	Consistent	Short vowel compound	A	16,9981	9,78
rugzak	backpack	Consistent	Short vowel compound	A	8,0002	5,83
leesles	reading lesson	Consistent	Short vowel compound	E	5,0001	n.a.
vuurwerk	fireworks	Consistent	Short vowel compound	E	5,0003	6,65
doolhof	labyrinth/maze	Consistent	Short vowel compound	O	7,0000	6,96
spaarpot	piggy bank	Consistent	Short vowel compound	O	1,0000	5,99
boomhut	treehouse	Consistent	Short vowel compound	U	n.a.	6,47
rugzak	backpack	Consistent	Short vowel compound	U	8,0002	5,83
manen	moons	Inconsistent	Inflected familiar word	A	8,0002	9,17
vragen	question	Inconsistent	Inflected familiar word	A	804,0812	4,90
benen	legs	Inconsistent	Inflected familiar word	E	2,9999	5,56
spelen	play	Inconsistent	Inflected familiar word	E	380,0143	3,90
dromen	dreams	Inconsistent	Inflected familiar word	O	65,9933	4,84
noten	nuts	Inconsistent	Inflected familiar word	O	1,0000	6,67
dure	expensive	Inconsistent	Inflected familiar word	U	n.a.	0,00
sturen	steer	Inconsistent	Inflected familiar word	U	100,9951	5,90
hagel	hail	Inconsistent	Monomorphemic familiar word	A	4,0004	7,02
water	water	Inconsistent	Monomorphemic familiar word	A	363,9988	3,67
lepel	spoon	Inconsistent	Monomorphemic familiar word	E	18,0011	4,19
negen	nine	Inconsistent	Monomorphemic familiar word	E	51,0035	4,86
vogel	bird	Inconsistent	Monomorphemic familiar word	O	96,0064	4,24
zomer	summer	Inconsistent	Monomorphemic familiar word	O	81,0028	5,33
juni	June	Inconsistent	Monomorphemic familiar word	U	41,9952	6,05
muziek	music	Inconsistent	Monomorphemic familiar word	U	115,0006	4,84
ruzie	fight	Inconsistent	Monomorphemic familiar word	U	43,0031	4,90
adem	breath	Inconsistent	Monomorphemic familiar word in meaning, unfamiliar in orthography	A	86,9961	5,68
bami	bami	Inconsistent	Monomorphemic familiar word in meaning, unfamiliar in orthography	A	1,0000	11,84

Appendix 2. Overview of Pseudoword Stimuli

Target	Consistent	Vowel
naag	Consistent	A
snaap	Consistent	A
reen	Consistent	E
veek	Consistent	E
goop	Consistent	O
woom	Consistent	O
kluur	Consistent	U
muut	Consistent	U
fatel	Inconsistent uninflected	A
waber	Inconsistent uninflected	A
gener	Inconsistent uninflected	E
grepel	Inconsistent uninflected	E
dovel	Inconsistent uninflected	O
gover	Inconsistent uninflected	O
dupie	Inconsistent uninflected	U
wubus	Inconsistent uninflected	U
nagen	Inconsistent inflected	A
snapen	Inconsistent inflected	A
renen	Inconsistent inflected	E
veken	Inconsistent inflected	E
gopen	Inconsistent inflected	O
women	Inconsistent inflected	O
kluren	Inconsistent inflected	U
muten	Inconsistent inflected	U

Appendix 3. Overview of Vowel Frequencies in the Age of Acquisition Corpus

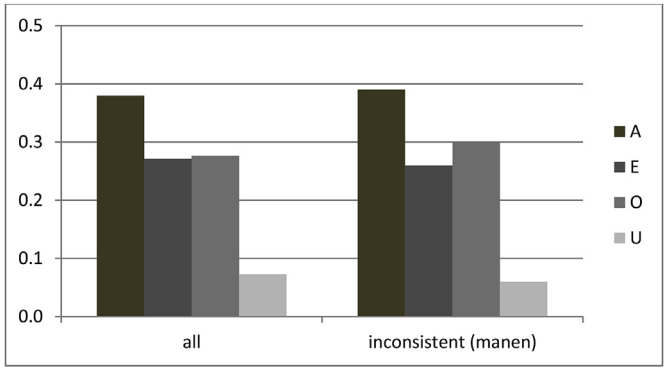

Fig. A. The relative proportions of A, E, O and U vowels in the AoA corpus. The total scores thus amount to 1.0. The left panel refers to all words in the corpus for the four vowel types (total = 1.0). The right panel refers to the inconsistent words in the corpus for the four vowel types (total = 1.0).

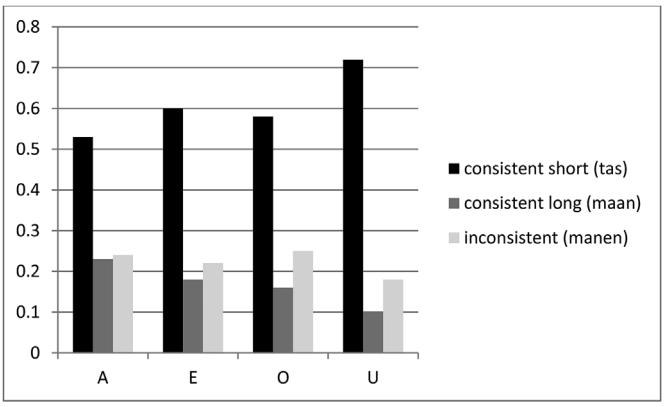

Fig. B. The relative proportions consistent short, consistent long, and inconsistent long vowels per vowel (A, E, O, U) in the AoA corpus. The total scores per vowel score amount to 1.0.

References

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

Beers, M. (1995). *The phonology of normally developing and language-impaired children*Doctoral dissertation. University of Amsterdam.

de Bree, E., van der Ven, S., & van der Maas, H. (2017). The voice of Holland: Allograph production in written Dutch past tense inflection. *Language Learning and Development.* http://dx.doi.org/10.1080/15475441.2016.1217777.

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica, 150*, 80–84. http://dx.doi.org/10.1016/j.actpsy.2014.04.010.

Deacon, S. H., Conrad, N., & Pacton, S. (2008). A statistical learning perspective on children's learning about graphotactic and morphological regularities in spelling. *Canadian Psychology, 49*, 118–124.

Deacon, S. H., Leblanc, D., & Sabourin, C. (2011). When cues collide: Children's sensitivity to letter- and meaning patterns in spelling words in English. *Journal of Child Language, 38*, 809–827. http://dx.doi.org/10.1017/S0305000910000322.

Ehri, L. C. (1992). Review and commentary: Stages of spelling development. In S. Templeton, & D. R. Bear (Eds.). *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 307–332). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders, 20*, 19–36. http://dx.doi.org/10.1097/00011363-200020030-00005.

Gerrits, P. A. M. (2001). *The categorisation of speech sounds by adults and children*Doctoral dissertation. Utrecht: LOT Publications.

Graham, S., Harris, K. R., & Hebert, M. (2011). It is more than just the message: Analysis of presentation effects in scoring writing. *Focus on Exceptional Children, 44*, 1–12. http://dx.doi.org/10.1201/b10962-2.

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*, 710–744. http://dx.doi.org/10.17763/haer.81.4.t2k0m13756113566.

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703–1743. http://dx.doi.org/10.1007/s11145-014-9517-0.

Hilte, M., & Reitsma, P. (2011). Effects of explicit rules in learning to spell open- and closed-syllable words. *Learning and Instruction, 21*, 34–45. http://dx.doi.org/10.1016/j.learninstruc.2009.10.002.

Kemp, N., & Bryant, P. (2003). Do beez buzz? Rule-based and frequency-based knowledge in learning to spell plural -s. *Child Development, 74*, 63–74. http://dx.doi.org/10.1111/1467-8624.00521.

Kemper, M. J., Verhoeven, L., & Bosman, A. M. T. (2012). Implicit and explicit instruction of spelling rules. *Learning and Instruction, 22*, 639–649. http://dx.doi.org/10.1016/j.lindif.2012.06.008.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods, 42*, 643–650. http://dx.doi.org/10.3758/BRM.42.3.643.

Kim, Y. S. (2010). Componential skills in early spelling development in Korean. *Scientific Studies of Reading, 14*, 137–158. http://dx.doi.org/10.1080/10888430903034812.

Landerl, K. (2003). Categorization of vowel length in German poor spellers: An orthographically relevant phonological distinction. *Applied PsychoLinguistics, 24*, 523–538 doi: 10.1017.S0142716403000262.

Landerl, K., & Reitsma, P. (2005). Phonological and morphological consistency in the acquisition of vowel duration spelling in Dutch and German. *Journal of Experimental Child Psychology, 92*, 22–44. http://dx.doi.org/10.1016/j.jecp.2005.04.005.

Lehtonen, A., & Bryant, P. (2005). Doublet challenge: Form comes before function in children's understanding of their orthography. *Developmental Science, 8*, 211–217. http://dx.doi.org/10.1111/j.1467–7687.2005.00409.x.

Nag, S., Treiman, R., & Snowling, M. J. (2010). Learning to spell in an alphasyllabary: The case of Kannada. *Writing Systems Research, 2*, 41–52. http://dx.doi.org/10.1093/wsr/wsq001.

Nunes, T., Bryant, P., & Bindman, M. (1997). Learning to spell regular and irregular verbs. *Reading and Writing: An Interdisciplinary Journal, 9*, 427–449.

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic

and morphological regularities. *Child Development, 76*, 324–339. http://dx.doi.org/10.1111/j.1467-8624.2005.00848_a.x.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130*, 401–426.

Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73*, 445–460.

Treiman, R., & Boland, K. (2017). Graphotactics and spelling: Evidence from consonant doubling. *Journal of Memory and Language, 92*, 254–264. http://dx.doi.org/10.1016/j.jml.2016.07.001https://doi.org/.

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology, 98*, 642–652. http://dx.doi.org/10.1037/0022-0663.98.3.642.

Treiman, R., & Kessler, B. (2014). *How children learn to write words.* Oxford, UK: Oxford University Press.

Van der Ven, S. H. G. & de Bree, E. H.. (in revision). Variability is the spice of spelling: The effect of implicit cues on Dutch past tense spelling is dependent on age, literacy, and task format.

Varnhagen, C. K., Boechloer, P. M., & Steffler, D. J. (1999). Phonological and orthographic influences on children's vowel spelling. *Scientific Studies of Reading, 3*, 363–379. http://dx.doi.org/10.1207/s1532799xssr0304_3.

Verhoeven, L., Schreuder, R., & Baayen, R. H. (2006). Learnability of graphotactic rules in word identification. *Learning and Instruction, 16*. http://dx.doi.org/10.1016/j.learninstruc.2006.10.003 538–448.

Wimmer, H., & Landerl, K. (1997). How learning to spell German differs from learning to spell English. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.). *Learning to spell. Research, theory, and practice across languages*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.