Variation is the Spice of Spelling: The Effect of Implicit Cues on Dutch Past Tense Spelling is Dependent on Age and Literacy, but Not on Task Format

Van Der Ven, S.; de Bree, E.

DOI
10.1080/10888438.2019.1579217

Publication date
2019

Document Version
Final published version

Published in
Scientific Studies of Reading

Scientific Studies of Reading

ISSN: 1088-8438 (Print) 1532-799X (Online) Journal homepage: https://www.tandfonline.com/loi/hssr20

Variation is the Spice of Spelling: The Effect of Implicit Cues on Dutch Past Tense Spelling is Dependent on Age and Literacy, but Not on Task Format

Sanne Van Der Ven & Elise de Bree

To cite this article: Sanne Van Der Ven & Elise de Bree (2019): Variation is the Spice of Spelling: The Effect of Implicit Cues on Dutch Past Tense Spelling is Dependent on Age and Literacy, but Not on Task Format, Scientific Studies of Reading

To link to this article: https://doi.org/10.1080/10888438.2019.1579217

Published online: 19 Feb 2019.

SCIENTIFIC STUDIES OF READING
https://doi.org/10.1080/10888438.2019.1579217

Routledge
Taylor & Francis Group

Variation is the Spice of Spelling: The Effect of Implicit Cues on Dutch Past Tense Spelling is Dependent on Age and Literacy, but Not on Task Format

Sanne Van Der Ven[a] and Elise de Bree[b]

[a]Utrecht University; [b]University of Amsterdam

ABSTRACT

Spelling is influenced by implicit cues, but less is known about variability in this reliance. We assessed whether the influence of three implicit cues on Dutch past tense spelling was moderated by grade, literacy, and format. An Auditory infinitive, Written infinitive, and Picture+cloze format was completed by 68 third-graders and 47 sixth-graders. The implicit cues voicing probability and token frequency affected accuracy, especially in Grade 6 and in children with higher literacy skills. There was no task format effect. This shows that children's inflection accuracy is affected by implicit cues, but the degree of reliance on these cues is variable.

Spelling matters beyond spelling: spelling quality has been found to influence both the writing process and other people's perception of the quality of written texts (Graham, Harris, & Hebert, 2011). At the same time, spelling acquisition is a prolonged process, in which instruction is important (Graham & Santangelo, 2014), particularly for words that cannot be spelled on the basis of phoneme-to-grapheme conversion. However, instruction does not guarantee perfect acquisition, as even adults make errors (e.g. De Schryver, Neijt, Ghesquière, & Ernestus, 2013; Kemp & Bryant, 2003; Mitchell, Kemp, & Bryant, 2011). Different factors might relate to this protracted acquisition. In this study, we looked into the spelling of Dutch regular past tense spelling and three potentially relevant types of factors: implicit cues (voicing probability, token frequency and relative graphotactic bigram proportion), child-based factors (grade, reading ability and phoneme awareness), and task format.

One reason why some spelling patterns are difficult to acquire is because learners do not necessarily exclusively rely on the rules that govern the correct spelling, but also on information about words and patterns they acquire through both spoken and print exposure. Print exposure allows children to acquire orthographic representations of (novel) words (Share, 1999). This exposure impacts on the lexical quality of words in memory: when children see words in context, they can store word forms, connect these with contextual information and enrich these associations (Perfetti & Hart, 2002). Repeated exposure thus enhances lexical quality, which contributes to writing skills (Share & Shalev, 2004). A meta-analysis by Mol & Bus (2011) confirmed that print exposure facilitates children's spelling development.

Second, both printed and spoken language entail exposure to sublexical patterns or implicit cues: statistical regularities in the language that are not explicitly taught (Deacon, Leblanc & Sabourin, 2011; Kemp & Bryant, 2003; Mitchell et al., 2011). These implicit cues have also been found to influence spelling acquisition and performance (Cassar & Treiman, 1997; de Bree,

CONTACT Elise de Bree e.h.debree@uva.nl Research Institute of Child Development and Education, University of Amsterdam, The Netherlands

Authors contributed equally to the study

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hssr.

Geelhoed, & van Den Boer, 2018; Treiman & Boland, 2017). The pattern that is deduced is, however, not always fully correct (Verhaert, 2016). Although it has been shown how implicit cues influence children's Dutch regular past tense spelling (de Bree, van der Ven, & van der Maas, 2017), less is known about the degree to which this reliance on implicit cues is variable between individuals and across task formats. This joint assessment will be done by looking into Dutch regular past tense spelling.

Dutch regular past tense spelling

Dutch regular past tense spelling is marked by two allomorphs: depending on whether the verb stem ending is voiced or unvoiced, either <te> (unvoiced) or <de> (voiced) is added. The term allomorph thus refers to the fact that two morphemes can be used; the distinction between the sounds does not alter the meaning. Past tense spelling is taught in Grade 6 with an explicit rule. This rule states that the writer should check the verb stem, i.e., the infinitive without the final <en>. For example, the stem of <snoepen> "to eat sweets", is <snoep>. If the final letter of this stem is part of the consonants in the mnemonic "'t kofschip" (a historic ship), the past tense is formed by adding <te> (<snoepte>); otherwise by adding <de> (e.g. <rennen> "to run" becomes <rende>). The mnemonic contains all voiceless consonants in Dutch, but this is not pointed out in the instruction of the spelling rule and the use of the mnemonic.

Despite the apparent ease of this rule, past tense spelling is error-prone (De Schryver et al., 2013) and often leads to incorrect allomorph selection, such as <eisde> for <eiste> (demanded) and <schrobte> for <schrobde> (scrubbed) (de Bree et al., 2017; De Schryver et al., 2013; Neijt & Schreuder, 2007). This may be due to the fact that syllable-final voiced plosives (/d,b/) and fricatives (/v,z,ɣ/) are devoiced in Dutch: these phonemes lose their voicing value in speech. For example, the verb <schrobben> (to scrub) has orthographic stem <schrob>.[1] The pronunciation, however, is /p/,/sXrOp/. A learner needs to know that this devoiced final consonant stems from an underlyingly voiced consonant in the infinitive (<schrobben>) to form the correct past tense (<schrobde>). Errors such as <schropte> or <schrobte> can occur if learners rely on the stem pronunciation/p/rather than the spelling (<b>). This devoicing thus complicates phoneme-to-grapheme conversion. For/v/and/z/, syllable-final devoicing occurs not only in speech, but also in spelling. The verb <reizen> (to travel) takes <de>, as <z> is not part of "'t kofschip". However, the stem ends in an /s/(/rEis/) and is written as <reis>. The past tense is spelled as <reisde>, although the pronunciation is /rEizd@/. A stem ending in /s/ can thus come from a verb with a/z/that takes <de> or from a verb with a/s/ that takes <te>. Spellers thus need to know the underlying voicing value.

Item-based factors in Dutch past tense spelling: implicit cues

Because Dutch past tense spelling contains ambiguous cases, a speller may not only rely on the explicit rule to select an allomorph, but also on implicit cues. In the present paper, we investigate the influence of three such cues. The first is the phonological cue of voicing probability (VP; Ernestus & Baayen, 2001, 2003). VP refers to the probability with which the final obstruent of a specific pattern of stem-final phonemes is underlyingly voiced. For example, in the verb <scheppen> (to scoop), the stem rhyme is /Ep/. The /p/ of <schep> is part of "'t kofschip" and its past tense takes <te> The pattern <ep> has a VP of 0.135, meaning that in 13.5% of the words with such a pattern the /p/ is a devoiced /b/. This aligns with the rule: the incorrect, voiced inflection <schebde> has a probability of only 0.135. A stem with a rhyme with a high VP is likely to generate <de>. The rhyme of <zorgen> (to care for) is <org>,[2] with a high VP (.953), pointing towards the correct allomorph <de> (<zorgde>). In some cases, VP and the rule conflict. The verb <eisen> (to demand) contains final

[1]The fact that there are two b's in *schrobben* but only one in *schrob* is due to other Dutch spelling rules that govern the presence of single and double vowels and consonants, depending on word length and place in the syllable.
[2]Note that <g> in Dutch refers to a fricative: /ɣ/.

<s> and should therefore take <te> (<eiste>). Nevertheless, the VP of the rhyme *eis* is .755, pointing towards <de> (incorrect <eisde>). Both children's and adults' past tense spelling performance has been shown to be affected by VP (de Bree et al., 2017; De Schryver et al., 2013).

The second implicit cue that may affect Dutch past tense spelling is relative graphotactic bigram proportion (RGBP; de Bree et al., 2017): the frequency of the sequence of the final grapheme of the stem and the consonant of the allomorph. In the verb <schepte> (scooped), for instance, <pt> is the sequence of interest. On the basis of a CELEX count (Baayen, Piepenbrock, & Van Rijn, 1993) of the relative frequency of <pt> and <pd> in Dutch past tense and participle verb spelling, <pt> occurs 100% and <pd> 0%. This means that RGBP points towards allomorph <te> for <schep>, matching the past tense spelling rule (with in "'t kofschip"). However, RGBP can also run counter to the rule: the grapheme sequence <fd> occurs 85% in Dutch past tense and participle verb spelling and <ft> only 15%. The verb <blaffen> (to bark) takes <te>, but because <fd> is much more frequent than <ft>, relying on the RGBP could steer the speller towards incorrect <blafde>. Such errors have been reported (de Bree et al., 2017).

The third implicit cue potentially affecting inflection accuracy is verb token frequency. More frequent words are likely to be more familiar and therefore easier to inflect. Indeed, in oral productions children inflected more frequent verbs correctly more often (Matthews & Theakston, 2006; Norbury, Bishop, & Briscoe, 2001). Such benefits might also influence written language, as written exposure might enhance orthographic familiarity with the target words, as well as activate the pronunciation. Token frequency has indeed been found to affect spelling accuracy and recognition in adults (Abrams & White, 2011; Mitchell et al., 2011).

Child-based factors in spelling

Spelling performance does not only vary across items, but also between children. Another approach is therefore to establish the contribution of child-based variables (e.g., Lervåg & Hulme, 2010; Veber Nielsen & Juul, 2016). In the present study, we investigate the contribution of age/grade, phoneme awareness (PA) and word reading ability. We are not only interested in the main effects of these child-based factors on spelling performance, but especially in the effect that these factors may have on the degree with which children rely on the implicit cues described in the previous section. As this issue has received limited attention, our expectations are speculative.

First of all, age is a potential factor. Children in older grades could show better spelling due to two reasons. First, older children's knowledge of the rule may lead them to rely less on implicit cues than younger children who do not know the rule yet. This knowledge of the rule should lead to increased correct scores for older children, regardless of the verb that needs to be inflected. The increased lexical quality of the older children could further lead to better past tense spelling. However, these older children also have more detailed phonological representations (Bybee, 1995; Matthews & Theakston, 2006), which may lead them to rely more strongly on VP, and they have more orthographic knowledge (Perfetti & Hart, 2002; Share, 1999), allowing them to rely more strongly on RGBP. This means that older children might rely more strongly on implicit cues that could in some instances also negatively affect spelling performance. This assumption agrees with the Integration of Multiple Patterns view of spelling (Treiman & Kessler, 2014), which states that multiple patterns (as well as instruction) play a role in spelling acquisition.

Individual differences in children's reading skills may also affect spelling outcomes. Reading ability influences spelling (Leppänen, Nieme, Aunola, & Nurmi, 2006). The lexical representations of children who are better readers might be of higher quality, facilitating spelling. Better readers might also be more sensitive to the orthographic patterns in the language. This assumption is supported by findings that better spellers rely on context and orthographic information more than other spellers (Treiman & Kessler, 2006; Varnhagen, Boechloer, & Steffler, 1999). This would mean that all better readers are more strongly affected by all three implicit cues: token frequency, VP and RBGP.

A similar line of reasoning can be made for PA, a skill known to predict spelling accuracy (e.g. Harrison et al., 2016). The ability to perform operations on detailed phonological representations assists grapheme-phoneme correspondences and is likely to contribute to lexical quality of specific words. It might also influence sensitivity to implicit cues: as VP demands specification of phonological rhyme environments, past tense spelling of children with better PA might be affected more strongly by implicit cues. A recent study by Kim, Petscher and Park (2016) looked into the joint contributions of item-based and child-based factors and found that emergent literacy (of which PA is a part) impacts on spelling acquisition. Specifically, emergent literacy had a substantial effect for children with lower spelling ability on spelling easier words. However, children with average emergent literacy skills showed more growth in spelling in more complex patterns. These findings indicate that children's spelling-related skills positively impact on spelling outcomes.

Task-based factors in spelling

A third type of factor involved in spelling might be the way in which output is demanded, such as filling in a missing grapheme or writing a word from memory. Such different task formats might influence children's inflection ability. This issue has received little attention in spelling research, to our knowledge (cf. Landerl & Reitsma, 2005; Puranik & Apel, 2010; Sarris & Panagiotakopoulos, 2010; Treiman & Bourassa, 2000). There are two approaches to the influence of such task demands.

First, situated learning theories state that children's skills depend on the context (e.g., Fischer, Bullock, Rotenberg, & Raya, 1993). Task format may impact overall accuracy, e.g., writing from dictation is more difficult than selecting the correct spelling (Landerl & Reitsma, 2005; Sarris & Panagiotakopoulos, 2010). Moreover, task format may also affect the importance of implicit cues. Compared to relying on implicit cues, the application of explicit spelling rules requires more resources from working memory, a limited capacity system that is strongly related to attention and to language processing (e.g., Baddeley, 2003; Engle, 2002; Erçetin & Alptekin, 2013). Cognitive Load Theory (Paas, van Gog, & Sweller, 2010; Sweller, 2010) shows how these limited working memory resources have to be divided over the intrinsic load of a task (inflecting the verb, in the present context), the germane load (building knowledge schemas), and the extraneous load (all other task demands, such as listening or reading). When the extraneous task load is heavier, less working memory capacity remains for the intrinsic and germane load (Paas et al., 2010; Sweller, 2010). This means that with increasing extraneous task load, children might rely less on the explicit rule and more on implicit cues, as the latter require fewer working memory resources.

Typically in spelling research, spelling is assessed in an isolated setting, in which the task demands are limited to spelling and explicitly stated. For instance, past tense has been tested by asking children to circle the correct consonant or <d> (Neijt & Schreuder, 2007) or by presenting them with the infinitive (de Bree et al., 2017). These tasks might differ in the degree of extraneous load from a perhaps more ecologically valid situation, when someone is writing his/her own text and thus has to use many different skills besides spelling, such as creating a narrative, building syntactically correct sentences, and retrieving vocabulary from long-term memory. Based on Cognitive Load Theory, we predict that in such a more open-ended assignment, more working memory resources are necessary for these other skills. As a consequence, the balance between relying on explicit rules and implicit cues to select the allomorph should shift towards a stronger reliance on implicit cues.

A second approach on task demands is that vocabulary, pronunciation and the orthographic representation all need to be activated and/or retrieved when spelling a word. These are associated: knowledge of the orthographic system and of word-specific orthographic representations can function as an aid to tie spellings of words to their pronunciations in memory (Ehri, 2005, 2014; Perfetti & Hart, 2002; Share, 2008).This connection enhances recall of the pronunciation of a word and might thus facilitate meaning retrieval (Chambré, Ehri, & Ness, 2017). When the orthographic presentation of a target is provided, this can facilitate retrieval of both the pronunciation as well as the meaning of the word. When no spelling is provided, this might be more difficult than when partial or complete

spellings of the infinitives are provided. It might also mean that the effect of implicit cues is more dominant in environments in which no or partial spellings are provided: if the lexical presentation is not provided, the implicit patterns might come into play.

The present study

The present study was designed to investigate differences, both between and within individuals, in the effect of implicit cues on past tense inflection in Dutch children. We compared two age groups: a group that had not been taught the rule (Grade 3; 8–9 year-olds) and a group that had (Grade 6; 11–12 year-olds). We assessed word reading ability and PA, as well as past tense spelling performance in three different task formats. Two were traditional, structured formats in which the target verb was presented as an infinitive, either auditorily (Auditory infinitive) or written (Written infinitive), and the task was explicitly to inflect these verbs to past tense. The third (Picture+cloze condition) was a more open-ended format that was designed to come closer to a more productive situation, while still giving cues such that all children were likely to produce and inflect the same verbs. In this condition, an image was given with an accompanying sentence in which the target verb was partly given: the onset and nucleus were given but the coda was omitted (see Table 1).

Two expectations were evaluated. First we expected that implicit cues, child-based factors and task-based factors all contribute to spelling outcomes: we expected implicit cues, grade, reading and PA to be positively related to accuracy. We also expected the Auditory infinitive condition to be more difficult than the other two, as no (partial) spellings were provided.

Second, we explored interactions between the implicit cues and the child- and task-based factors. If instruction of the past tense spelling rule is successful, Grade 6 should outperform Grade 3 across-the-board, without interactions with implicit cues, but if spoken and print exposure and lexical quality influence past tense spelling, these interactions are likely present. Furthermore, we expected the Picture+cloze condition to require more resources from working memory to generate the verb that has to be inflected. This reduction of available resources for the rest of the task should result in a reduced reliance on the rule and thus a stronger effect of the implicit cues.

Table 1. Task Formats Used in the Study.

Condition	Example	Correct answer
Written infinitive	*Een mug had mij in de zomer gestoken. Ik kreeg heel veel jeuk en …… (krabben) zo erg dat mijn arm ging bloeden.* "A mosquito bit me in the summer. It itched a lot and I …….. (scratch) so hard my arm started to bleed."	Krabde
Auditory infinitive	*Gisteren …… ik aan een muggenbult.*  *krabben* ("scratch")	Krabde
Picture+cloze	 'Yesterday …… I on a mosquito bite *Ik kra … e op mijn buik, want ik had jeuk* "I scra …d my belly, because I had an itch"	Krabde

Method

Participants

Children in Grade 3 and 6 were recruited from three schools, in different parts of The Netherlands. Reading and PA data were missing from one child in Grade 6 due to illness; data from this child were removed from all analyses. The final sample consisted of 68 children (41% girls) in Grade 3 (mean age = 9.1 years; SD = 0.42 year) and 46 (46% girls) children in Grade 6 (mean age = 12.1 years; SD = 0.45 year). All participating children had active parental consent.

Instruments

Word reading

Word reading was assessed with two tasks: a timed word reading task (*Eén Minuut Test;* EMT; Brus & Voeten, 1999) and a timed pseudoword reading task (*Klepel;* van Den Bos, Lutje Spelberg, Scheepstra, & de Vries, 1994). The child has one and two minutes, respectively, to read aloud as many words/pseudowords as possible. Targets increase in length from one to four syllables. The score is the number of correctly read (pseudo)words (maximum of 116 in both tasks). Reliability of both tasks is good (0.90 for EMT and 0.92 for Klepel; Evers et al., 2009–2012)). The tasks correlated strongly, $r(112) = 0.91$, $p < 0.001$. Therefore a composite score was constructed: both were standardized ($M = 0$, SD = 1) and the mean score was taken.

Phoneme awareness

PA was assessed using an abbreviated version containing 17 items from the phoneme deletion task by de Jong and van der Leij (2003). The child heard a pseudoword and had to repeat it. After repetition, phonemes had to be deleted or reversed. The first nine items demanded a single deletion, the next four items a double deletion (the phoneme was present twice), and the final four items demanded phoneme reversal (bolteps → bolpets). The score was the total number of items answered correctly, which was standardized ($M = 0$; SD = 1).

Past tense spelling

The past tense spelling task consisted of 27 different verbs in three different sentence formats: (1) the infinitive in brackets (Written infinitive), (2) auditory input of the infinitive (Auditory infinitive), and (3) a sentence in which part of the verb was given together with an image of the verb (Picture +cloze). Examples are provided in Table 1.

Different types of verbs were included. Thirteen verbs take <te>, and these verb stems ended on a voiceless plosive ($n = 4$) or fricative ($n = 9$). The other 14 verbs take <de>; their stems ended on a voiced plosive ($n = 2$), fricative ($n = 9$), nasal ($n = 1$), glide ($n = 1$), or trill ($n = 1$). Nasals, vowels, and glides always take <de>, as these are always voiced in Dutch. Stimuli are presented in the Appendix.

Implicit cues related to past tense spelling

For each verb, three measures of implicit cues were derived. First, the VP value was derived from Ernestus and Baayen (2003). The correct answer coincides with a low VP for <te> verbs but with a high VP for <de> verbs. Therefore, the VP values for the verbs taking <te> were recoded as 1-VP, such that a high value corresponds with a greater alignment with the correct answer. Second, the RGBP of the stem-final grapheme and both the correct and the incorrect consonant (, <d>) was determined tokenwise, based on all past tenses and participles from the

CELEX database (Baayen et al., 1993). The proportion in which the stem-final grapheme was followed by the correct consonant was then taken. Third, token frequencies of the verbs were derived from CELEX and natural log-transformed.

Procedure

Data were collected by trained assistants in spring. In three class sessions, each lasting approximately 30 minutes, children were asked to inflect the verbs. Each session consisted of 27 items: nine in each task format. After the three sessions, all 27 verbs had been inflected in all three task formats. Two parallel versions were created, to counterbalance the presentation order of the verbs. Word reading was tested in an individual session in the same period. The procedure adheres to the ethical guidelines of Utrecht University and conforms to the standards of the Declaration of Helsinki.

Data screening and analysis

All 9234 past tense realizations were entered manually. Data from 230 realizations (2.5%) were missing, because of illness (162 realizations: 4 children missed one session; 1 child missed two sessions), or illegible handwriting (68 realizations). All remaining 9004 realizations were scored automatically using R (R Core Team, 2017) as correct, allomorph error, or other. Errors classified as "other" involve all other types of errors, often related to violations of other Dutch spelling rules, such as rules involving doubling (e.g., snoeoepte for snoepte "ate candy", lefde for leefde), selection of the wrong verb (pulkte "picked" for pufte "puffed") or failure to inflect the verb at all. The result of the classification of these 9004 realizations is presented in the descriptive results.

Subsequently, the effects of the implicit cues, presentation format and child characteristics on the accuracy of children's past tense realizations were analyzed. For these analyses, all errors of type "other" were removed, leaving the 7713 past tense realizations scored as "correct" or "allomorph error". Then multilevel analyses were carried out, with child, item and presentation format as levels. As the outcome variable was dichotomous (correct/allomorph error), multilevel logistic regression analyses were performed, with the R package lme4 (Bates, Mächler, Bolker, & Walker, 2014). In Model 1, only main effects of the predictors were investigated. In Model 2, interaction terms were added between the three implicit cues on the one hand and the child- and task-based factors on the other. Significant interaction effects mean that the effect of the implicit cues depends on these child characteristics and task formats, and thus indicate variability of the effect of the implicit cues.

Results

Descriptive statistics

Proportion correct scores and allomorph errors per format, grade and VP are presented in Figure 1. On the left-hand side are verbs that take past tenses with <te>; on the right-hand side those that take <de>. Within both these <te> and <de> verbs, the verbs are plotted from low to high VP. Thus, for the <te> verbs, the verbs on the left-hand side have VPs that agree with the actual inflection, such as <bukken> "to bend" and <scheppen> "to scoop" whereas those on the right side of the axis have conflicting VPs, such as <eisen> "to demand" and juichen "to cheer". The pattern is the opposite for the verbs taking <de> past tenses; those on the right-hand side have VPs that agree with the actual inflection, such as <gooien> "to throw" and <horen> "to hear", whereas those on the left side of the axis have conflicting VPs, such as <krabben> "to scratch" and <schrobben> "to scrub" . The verbs in the middle of the chart thus most clearly show the impact of VP as conflicting cue. The graphs show

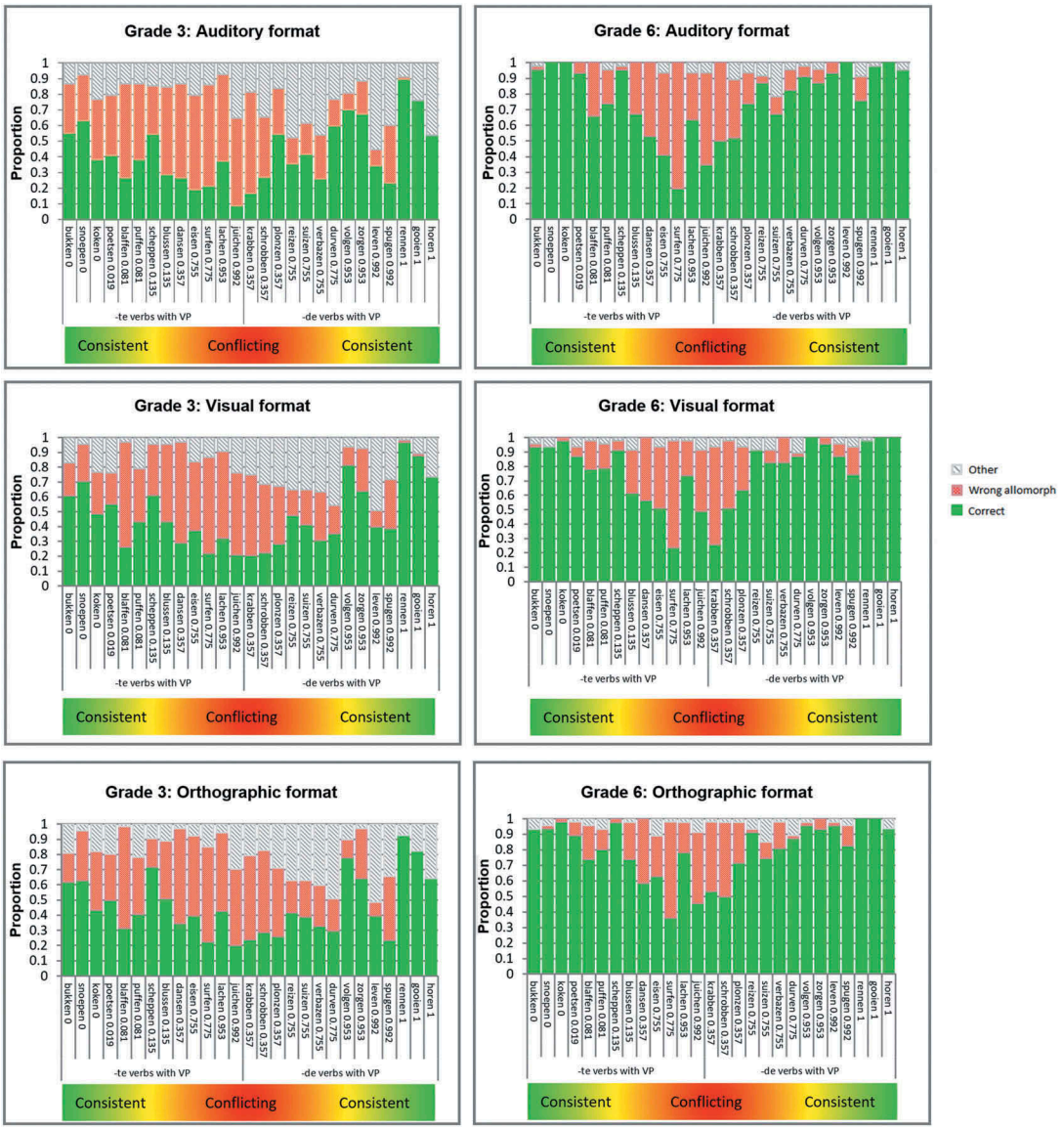

Figure 1. Past tense accuracy and error types, divided by condition and grade.

the proportion of correct allomorphs, incorrect allomorphs (e.g., <de> as <te> and vice versa) and other errors. Note that the statistical analyses below are conducted only on the incorrect allomorph/correct realizations.

A full overview of the proportion correct scores for each verb per presentation format and grade, as well as the VP, RGBP and token frequency, is presented in the Appendix. Children in Grade 6 outperformed the children in Grade 3. Moreover, while in Grade 3 both allomorph errors and other errors were common, children in Grade 6 made predominantly allomorph errors.

The children's mean reading and PA scores are shown in Table 2. Grade 6 outperformed Grade 3 on all measures. Table 3 shows the correlations between the implicit cues and the child-based factors. All child-based factors as well as RGBP and token frequency correlated significantly.

Table 2. Means and Standard Deviations of Reading and Phoneme Awareness by Grade.

	Raw score		Transformed score[c]	
	Grade 3	Grade 6	Grade 3	Grade 6
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)
Reading (composite)	–	–	−0.37 (0.90)	0.55 (0.81)
Words[a]	60.56 (16.27)	77.50 (13.24)	–	–
Pseudowords[a]	50.68 (18.95)	68.26 (18.23)	–	–
Phoneme Awareness[b]	9.85 (3.09)	11.78 (2.32)	−0.26 (1.04)	0.39 (0.78)

Note. [a]Maximum possible score = 116; [b]Maximum possible score = 17;[c] M = 0, SD = 1 in entire sample.

Table 3. Correlation Matrices for the Implicit Cues and the Child-based Factors.

	Implicit cues (item-based factors) (n = 27)			Child-based factors (n = 114)	
	VP	Token frequency		Grade	Reading
Token frequency	0.20	–	Reading	0.46**	–
RGBP	0.10	0.39*	PA	0.32**	0.53**

*p <0.05, **p <0.001

Predicting accuracy with multilevel logistic regression analyses

The results of the multilevel logistic regression analyses are displayed in Table 4. In Model 1, only main effects were included. Two of the three implicit cues significantly affected accuracy: both consistent VP and higher token frequency increased the probability of a correct answer. The effect of RGBP was not significant. Furthermore, two of the three user characteristics were significant: children in Grade 6 and children with better reading abilities were more likely to answer correctly, while PA did not significantly predict accuracy. There was no significant effect of task format.

In Model 2, it was investigated whether the effects of the three implicit cues differed across format and different levels of the child characteristics. Interaction terms between the three implicit cues on the one hand and all child- and task-based factors on the other were added to the model (Table 4, Model 2). The significant interaction effects are displayed in Figure 2. In each plot the predicted probability of a correct answer by the two variables in the interaction effect is shown, while all other variables are kept at their mean value.

The first three significant interaction effects illustrate that the effect of VP was stronger in children of higher age and ability. Accuracy in Grade 6 was always higher than in Grade 3 (Figure 2a) and better readers also always outperformed poorer readers (Figure 2b), but both plots show a widening gap as VP increases. Figure 2c shows an even more extreme effect for PA: for lower, less consistent VPs the probability for a correct answer even decreased with increasing PA. Note that the main effect of PA in Model 2 is negative, but the inclusion of interaction terms in this model means that this main effect should be interpreted as the effect of PA only when VP = 0; at low VP levels, the effect of a higher PA is indeed negative. The final two significant interaction effects show that the effect of token frequency was affected by age and ability in a pattern similar to VP: with increasing token frequency, the advantage of Grade 6 (Figure 2d) and better readers (Figure 2e) increased. Note that on average, children in Grade 6 outperformed children in Grade 3 in both reading ability and PA (see Table 2) meaning that the different effects add up in practice.

Discussion

In this study we investigated whether implicit cues (VP, token frequency, RGBP) influence spelling accuracy and whether their effects differ across learners (age, reading ability and phonological awareness)

Table 4. Multilevel Logistic Regression Outcomes on Past Tense Spelling.

	Model 1				Model 2			
	B	SE	*Z*	*p*	*B*	SE	*z*	*p*
Intercept	−1.58	0.67	−2.37	0.018	−1.41	0.75	−1.90	0.058
Item characteristics (implicit cues)								
Voicing probability	2.94	0.64	4.59	<0.001	2.70	0.72	3.75	<0.001
Token frequency	0.80	0.23	3.46	<0.001	0.73	0.25	2.88	0.004
Relative Graphotactic bigram proportion	0.15	0.59	0.25	0.805	0.23	0.66	0.35	0.724
Child characteristics								
Grade 6[b]	1.28	0.13	9.60	<0.001	0.54	0.24	2.22	0.026
Reading	0.27	0.07	3.67	<0.001	0.28	0.14	2.04	0.041
Phoneme Awareness	0.01	0.07	0.17	0.863	−0.30	0.12	−2.43	0.015
Format characteristics								
Auditory infinitive format[a]	−0.11	0.06	−1.57	0.117	0.12	0.22	0.54	0.587
Written infinitive format[a]	0.08	0.06	1.17	0.241	0.20	0.22	0.93	0.353
Cross-level interaction terms (implicit cues × child and format characteristics)								
Voicing probability × Grade 6[b]					1.26	0.21	5.99	<0.001
Voicing probability × Reading					0.25	0.12	2.17	0.030
Voicing probability × Phoneme Awareness					0.28	0.10	2.67	0.008
Voicing probability × Auditory infinitive format[a]					−0.12	0.22	−0.59	0.554
Voicing probability × Written infinitive format[a]					−0.30	0.21	−1.39	0.164
Token frequency × Grade 6[b]					0.23	0.08	2.99	0.002
Token frequency × Reading					0.12	0.04	2.76	0.006
Token frequency × Phoneme Awareness					0.01	0.04	0.27	0.307
Token frequency × Auditory infinitive format[a]					0.03	0.08	0.42	0.676
Token frequency × Written infinitive format[a]					0.00	0.08	−0.02	0.983
Relative graphotactic bigram proportion × Grade 6[b]					0.08	0.19	0.41	0.685
Relative graphotactic bigram proportion × Reading					−0.17	0.11	−1.58	0.113
Relative graphotactic bigram proportion × Phoneme Awareness					0.18	0.10	1.81	0.070
Relative graphotactic bigram proportion × Auditory infinitive format[a]					−0.21	0.20	−1.02	0.307
Relative graphotactic bigram proportion × Written infinitive format[a]					0.10	0.20	0.49	0.628

Note. [a]Relative to dummy variable of picture+cloze format; [b]Relative to Grade 3

and task format. Dutch past tense spelling was assessed in children who had not (Grade 3) and had (Grade 6) been taught the orthographic rule. We used three task formats: a Written infinitive condition, an Auditory infinitive condition, and a Picture+cloze condition. Two multilevel logistic regression analyses were conducted to establish the main contributions of the implicit cues, the child-based factors and the task formats (Model 1), and the interactions between the implicit cues on the one hand and the child- and task-based factors (Model 2).

Model 1 confirmed that most predictors indeed affected allomorph errors. The model confirmed the importance of the implicit cue VP (de Bree et al., 2017; De Schryver et al., 2013). Token frequency also contributed to past tense spelling, in line with assumptions that frequency influences learning (Bybee, 1995; Norbury et al., 2001) and lexical quality (Perfetti & Hart, 2002). RGBP did not contribute independently to past tense spelling, not even when Model 1 was rerun without token frequency or VP. RGBP might only have an effect when it does not agree with VP values (de Bree et al., 2017) and thus not show an overall effect.

With respect to child characteristics, both grade level and reading ability positively affected Dutch past tense spelling accuracy. The latter result matches findings on literacy development, in which a general contribution of reading on spelling has been reported (Leppänen et al., 2006). PA, however, did not contribute independently. There might be two related reasons for this finding. The first is that PA is an important contributor to reading development (Melby-Lervåg, Lyster, & Hulme, 2012), which might subsume the effect of PA. Indeed, when Model 1 was rerun without word reading fluency, the effect of PA became marginally significant (B = 0.12, SE = 0.06, z = 1.93, p = 0.054). The second is that the ability to read on the whole word level might be more important than the ability to manipulate words at the local, phoneme level. This latter interpretation matches the findings on the implicit cues, with token frequency being a significant predictor, but not RGBP.

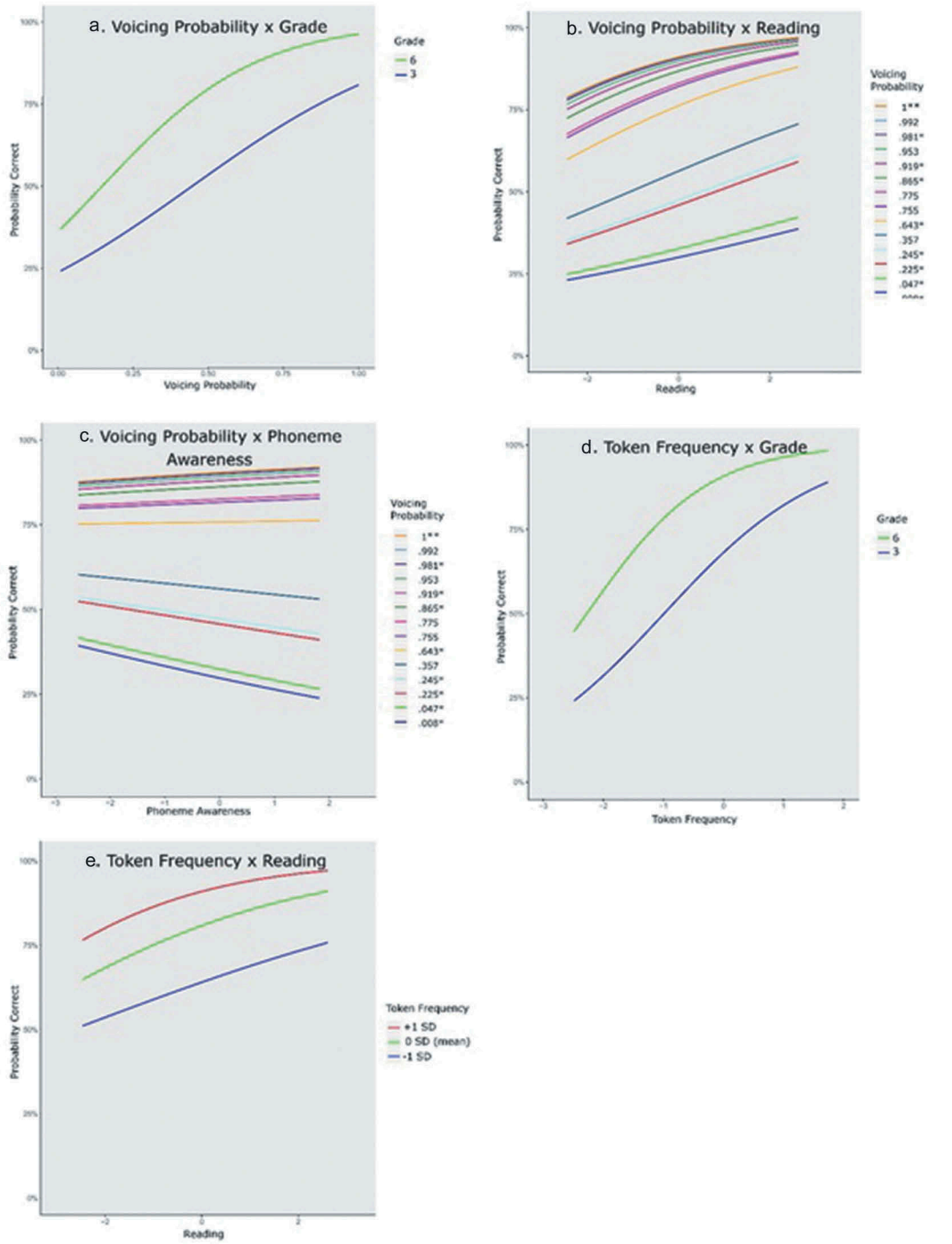

Figure 2. Graphical representation of the significant interaction effects of the multilevel logistic regression analysis (Model 2). *VP value for verbs taking <te>. The displayed VP value is the recoded value: (1 – actual VP value). **VP value for mixture of verbs taking <te> and verbs taking <de> that are all fully consistent: VP is 1 for verbs taking <de>, and recoded 1– 0 = also 1 for the verbs taking <te>.

Note: the top-to-bottom order of the lines in the graphs always corresponds to the order in the legend.

Task format did not matter for spelling accuracy. This finding partly agrees with previous findings. Puranik and Apel (2010) found no effect of spelling format (spelling-to-dictation, oral spelling, and using letter tiles to form the word) on kindergartners' spelling, once they could write 19 letters or more adequately. In contrast, Treiman and Bourassa (2000) found that first- and second-grade children had more difficulty in spelling words orally than on paper. It should be noted, however, that in the auditory infinitive condition the number of non-allomorph related errors, which were removed for the current analyses, was slightly higher (n = 462) than in the other two conditions (Picture+cloze: n = 411; Written infinitive: n = 418).

Model 2 showed that the effect of the implicit cues differed with child ability. Specifically, the effects of both VP and token frequency were stronger in Grade 6 and for children with better word reading ability. These findings mean that the importance of implicit cues is variable. We take the finding that older children and better readers rely more strongly on VP and token frequency to mean that exposure to spoken and orthographic information leads children to draw conclusions about past tense spelling patterns (Pacton, Borchardt, Treiman, Lété & Fayol, 2014). VP also interacted with PA: the higher the PA, the higher the accuracy for verbs with a consistent VP, but the *lower* the accuracy for inconsistent verbs. We assume that better PA relates to better segmentation ability (e.g., Boada & Pennington, 2006): better PA would thus mean that learners become more perceptive to the phonological environments such as the VP. This significant interaction effect with VP also forms an explanation for the lack of main effect in Model 1: a higher PA leads to an advantage for consistent verbs but a disadvantage for inconsistent verbs. These effects counter each other.

There was no significant interaction between the implicit cues and task format, contrary to our expectations based on Cognitive Load Theory (Sweller, 2010), which led us to expect larger effects of the implicit cues in the Picture+cloze condition. Possibly, our Picture+cloze condition was not sufficiently ecologically valid to attest such an effect. As the focus of the task was always on spelling, the cognitive load may not have been increased that much. A better, yet time-consuming approach would be to examine children's written assignments for other classes, such as creative writing. Only such a setting, in which cognitive load is not only caused by a correct application of the spelling rules but also by the other task at hand, is truly ecologically valid. Another explanation for the lack of effect is that possibly the dependence on implicit cues was already large in all conditions, meaning that there was not much room to switch the balance towards even more reliance on implicit cues. Future research could investigate older children who have more experience in applying the explicit rule. It should be noted, however, that the task formats used in the study, in which the required inflection for past tense was very explicit, should favor reliance on the explicit rules rather than implicit cues. Nevertheless, we found clear effects of VP, which were even stronger in Grade 6.

The outcomes of this study have implications for spelling models. Our findings suggest that even after instruction of the spelling rule, children still rely on implicit cues, not only on the rule. The results thus lend support to the lexical quality hypothesis (Perfetti & Hart, 2002): words that are higher in frequency were generally spelled better, accounting for the lower correct scores for surfen ("to surf" loanword, low in frequency, but also a conflicting VP) and for lower correct score of juichen ("to cheer") than lachen ("to laugh"). Words that are familiar might be easier to spell (see also de Bree et al., 2018). The findings also agree with the integration of multiple patterns view (Treiman & Kessler, 2014): there was an effect of token frequency (in line with both lexical quality and multiple patterns), as well as VP (in line with the implicit cue view). The effect of VP was stronger than the effect of token frequency. This is illustrated by the verbs in which VP and token frequency collide: e.g., <dansen> and <lachen>. Both are high-frequent words with a VP that is inconsistent with the actual past tense. We found a high number of allomorph errors <dansde> and <lachde>, in line with VP but not with token frequency. These findings thus lend the strongest support to the implicit cues view. The finding that the effect of implicit cues increased rather than decreased in older children underscores the need to develop more specified hypotheses. The findings confirm previous results that children and adults' spellings are influenced by different cues (e.g. Deacon, Leblanc, & Sabourin, 2011, 2018; de Bree et al., 2017; Pacton, Borchardt, Treiman, Lété, &

Fayol, 2014; Treiman & Boland, 2017). They also indicate that the sensitivity to such cues is influenced by other factors, such as word reading ability and phoneme awareness, and that assessment of spelling outcomes in tandem with cognitive child factors is needed (e.g. Kim, Petscher, & Park, 2016).

The study outcomes also have educational implications. Our findings suggest that it is desirable to tailor instruction of the mnemonic to the individual learner. Specifically, children of lower ability may need support in overall application of the rule, while children with more advanced abilities may need instruction specifically for the verbs in which the correct answer collides with the implicit cues, especially with VP.

Our study is qualified by a number of limitations. First, we did not check whether the children in Grade 6 actually knew the past tense spelling rule. However, the mnemonic is part of the national school curriculum. Also, De Schryver et al. (2013) found that Dutch university students were able to reproduce the rule, but nevertheless struggled in applying it. Thus, while we have no direct evidence that children in Grade 6 knew the mnemonic, we assume this to be the case.

Second, we presented tests with a limited number of verbs that needed to be turned into past tenses. However, the use of multilevel logistic regression enabled us to use the information from every past tense realization separately: the 7713 realizations established a high statistical power. Moreover, the verbs were carefully selected to represent the entire range of VPs and the pattern of findings is in line with a previous study with 227 verbs (de Bree et al., 2017). A related matter is that we only assessed past tense spelling and that no fillers or larger spelling test battery were presented. Children could thus have relied on a different spelling strategy than in a setting with assessment of different spelling rules. There is no way to establish the influence of this design. However, the study by de Bree et al. (2017) looked at past tense spelling in a Written infinitive format. These past tense spellings were mixed with other morphological spelling targets. The fact that the main pattern of findings of the present study is similar to that of de Bree et al. (2017) is reassuring. Nevertheless, a study containing past tense spelling as part of larger spelling assessment is required to truly assess the impact of the design on the past tense spellings.

A third limitation is that we did not assess whether the children were able to distinguish voiced and voiceless consonants, e.g., assessing how they spelled correctly inflected past tense (spell "surfte"). If they were not able to distinguish the consonants properly, this would complicate past tense spelling. Such an interpretation has been proposed, as the voicing distinction is fading in Dutch (De Schryver et al., 2013). However, there are indications that this is not the case. First, if children could not distinguish voiced and voiceless consonants, selection of <te> or <de> would be random. Instead, errors were tuned to VP, which is strongly related to voicing distinction. Furthermore, children in our sample came from two regions: the Randstad, the urbanized area in the Netherlands in which the voicing distinction in word-onset position (not necessarily syllable-onset) is fading and Brabant, a province in which the voicing distinction is still omnipresent. We conducted an additional analysis including an interaction effect of VP*region to test whether the effect of VP was stronger in the latter group, but this was not the case. Additionally, a study by Warner, Sereno, Jongman and Kemps (2004) found that in Dutch speech, even production of final devoiced consonants (e.g. <hand>, "hand" /hAnt/) is not neutralized entirely and thus differs subtly from consonants that are not underlyingly voiced (e.g. <krant> "newspaper" /krAnt/). This means that although the voicing contrast might be fading, there is still a voicing distinction.

A fourth limitation is that we tested children differing three years in age. Grade differences can thus be due to instruction, experience with implicit cues and to other age differences. This approach maximized potential differences and as such enabled us to identify clear grade differences. As a next step, targeted studies, including strategy assessment, could unravel how these differences arise (e.g. Rittle-Johnson & Siegler, 1999). Specifically, a microgenetic study, following children while they acquire the rule, can shed more light on the effect of learning the rule on the reliance on implicit cues and on child and format characteristics moderating these relationships. Such a study could also

look in more detail at the spoken and print exposure that children have had, to disentangle the contributions of both on past tense spelling outcomes.

In sum, this study confirms that the implicit cues of voicing probability and token frequency are important in Dutch past tense spelling, in line with different frameworks of spelling. The study adds that the importance of these cues is variable: the influence was stronger in older children and in children with better word reading abilities and better phoneme awareness. This finding needs to be incorporated in both research and teaching practice.

Conflict of Interest

Both authors report no conflict of interest.

References

Abrams, L., & White, K. (2011). Influences of word frequency, context and age on spelling. *Spelling Skills: Acquisition, Abilities, and Reading Connection*, 51–76.

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

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189–208. doi:10.1016/S0021-9924(03)00019-4

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi:10.18637/jss.v067.i01

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology*, *95*, 153–193. doi:10.1016/j.jecp.2006.04.003

Brus, B. T., & Voeten, M. J. M. (1999). *Eén-minuut-test* [One-minute-test]. Amsterdam, The Netherlands: Hartcourt Test Publishers.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*, 425–455. doi:10.1080/01690969508407111

Cassar, M. T., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, *89*, 631–644. doi:10.1037/0022-0663.89.4.631

Chambré, S., Ehri, L., & Ness, M. (2017). Orthographic facilitation of first graders' vocabulary learning with and without print referencing. *Reading & Writing: An Interdisciplinary Journal*, *30*, 1–20. doi:10.1007/s11145-016-9715-z

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*, *13*, 215–240. doi:10.1080/15475441.2016.1217777

de Bree, E. H., Geelhoed, J., & van Den Boer, M. (2018). Overruled! Implicit cues rather than an orthographic rule determine Dutch children's vowel spelling. *Learning & Instruction*, *56*, 30–41. doi:10.1016/j.learninstruc.2018.03.006

de Jong, P. F., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology*, *95*, 22–40. doi:10.1037/0022-0663.95.1.22

De Schryver, J., Neijt, A., Ghesquière, P., & Ernestus, M. (2013). Zij surfde, maar hij durfte niet. De spellingproblematiek van de zwakke verleden tijd in Nederland en Vlaanderen [She surfed, but he dared not. The spelling difficulties of the weak past tense in The Netherlands and Flanders]. *Dutch Journal of Applied Linguistics*, *2*, 133–151. doi:10.1075/dujal.2.2.01de

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. doi:10.1017/S0305000910000322

Ehri, L. (2005). Learning to read words: Theory, findings and issues. *Scientific Studies of Reading*, *9*, 167–188. doi:10.1207/s1532799xssr0902_4

Ehri, L. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*, 5–21. doi:10.1080/10888438.2013.819356

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. doi:10.1111/1467-8721.00160

Erçetin, G., & Alptekin, C. (2013). The explicit/implicit knowledge distinction and working memory: Implications for second-language reading comprehension. *Applied Psycholinguistics*, *34*, 727–753. doi:10.1017/s0142716411000932

Ernestus, M., & Baayen, H. (2001). Choosing between the Dutch past tense suffixes -te and -de. In T. Van der Woude & H. Broekhuis (Eds.), *Linguistics in the Netherlands* (pp. 77–88). Mahwah, NJ: Lawrence Erlbaum Associates.

Ernestus, M., & Baayen, H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, *79*, 5–38. doi:10.1353/lan.2003.0076

Evers, A., Egberink, I. J. L., Braak, M. S. L., Frima, R. M., Vermeulen, C. S. M., & Van Vliet-Mulder, J. C. (2009–2012). *COTAN documentatie* [COTAN documentation]. Amsterdam, The Netherlands: Boom Testuitgevers.

Fischer, K. W., Bullock, D., Rotenberg, E. J., & Raya, P. (1993). The dynamics of competence: How context contributes directly to skill. In R. H. Wozniak & K. W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments* (pp. 93–117). New York, NY: Psychology Press.

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. doi:10.1201/b10962-2

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

Harrison, G. L., Goegan, L. D., Jalbert, R., McManus, K., Sinclair, K., & Spurling, J. (2016). Predictors of spelling and writing skills in first- and second-language learners. *Reading and Writing*, *29*, 69–89. doi:10.1007/s11145-015-9580-1

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

Kim, Y.-S. G., Petscher, Y., & Park, Y. (2016). Examining word factors and child factors for acquisition of conditional sound-spelling consistencies: A longitudinal study. *Scientific Studies of Reading*, *20*, 265–282. doi:10.1080/10888438.2016.1162794

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*, 322–344. doi:10.1016/j.jecp.2005.04.005

Leppänen, U., Nieme, P., Aunola, K., & Nurmi, J.-E. (2006). Development of reading and spelling Finnish from preschool to Grade 1 and Grade 2. *Scientific Studies of Reading*, *10*, 3–30. doi:10.1207/s1532799xssr1001_2

Lervåg, A., & Hulme, C. (2010). Predicting the growth of early spelling skills: Are there heterogeneous developmental trajectories? *Scientific Studies of Reading*, *14*, 485–513. doi:10.1080/10888431003623488

Matthews, D. E., & Theakston, A. (2006). Errors of omission in English-speaking children's production of plurals and the past tense: The effects of frequency, phonology and competition. *Cognitive Science*, *30*, 1027–1052. doi:10.1207/s15516709cog0000_66

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*, 322–352. doi:10.1037/a0026744

Mitchell, P., Kemp, N., & Bryant, P. (2011). Variations among adults in their use of morphemic spelling rules and word-specific knowledge when spelling. *Reading Research Quarterly*, *46*, 119–133. doi:10.1598/RRQ.46.2.2

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*, 267–296. doi:10.1037/a0021890

Neijt, A., & Schreuder, R. (2007). Asymmetrical phoneme-grapheme mapping of coronal plosives in Dutch. *Written Language & Literacy*, *10*, 219–234. doi:10.1075/wll.10.2.04nei

Norbury, C. F., Bishop, D. V. M., & Briscoe, J. (2001). Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment. *Journal of Speech, Language and Hearing Research*, *44*, 165–178. doi:10.1044/1092-4388(2001/015)

Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review*, *22*, 115–121. doi:10.1007/s10648-010-9133-8

Pacton, S., Borchardt, G., Treiman, R., Lété, B., & Fayol, M. (2014). Learning to spell from reading: General knowledge about spelling patterns influences memory for specific words. *The Quarterly Journal of Experimental Psychology*, *67*, 1019–1036. doi:10.1080/17470218.2013.846392

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. T. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–214). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Puranik, C., & Apel, K. (2010). Effect of assessment task and letter writing ability on preschool children's spelling performance. *Assessment for Effective Intervention*, *36*, 46–56. doi:10.1177/1534508410380040

R Core Team. 2017. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rittle-Johnson, B., & Siegler, R. S. (1999). Learning to spell: Variability, choice, and change in children's strategy use. *Child Development*, *70*, 332–348.

Sarris, M., & Panagiotakopoulos, C. (2010). Word spelling assessment using ICT: The effect of presentation modality. *Themes in Science and Technology Education*, *3*(1–2), 93–118.

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*, 95–129. doi:10.1006/jecp.1998.2481

Share, D. L. (2008). Orthographic learning, phonology and the self-teaching hypothesis. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 36, pp. 31–82). Amsterdam, The Netherlands: Elsevier.

Share, D. L., & Shalev, C. (2004). Self-teaching in normal and disabled readers. *Reading and Writing*, *17*, 769–800. doi:10.1007/s11145-004-2658-9

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*, 123–138. doi:10.1007/s10648-010-9128-5

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

Treiman, R., & Bourassa, D. C. (2000). The development of spelling skill. *Topics in Language Disorders*, *20*, 1–18. doi:10.1097/00011363-200020030-00004

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

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

van Den Bos, K. P., Lutje Spelberg, H. C., Scheepstra, A. J. M., & de Vries, J. R. (1994). *De Klepel. Vorm A en B* [Nonword reading test]. Amsterdam, The Netherlands: Pearson.

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. doi:10.1207/s1532799xssr0304_3

Veber Nielsen, A.-M., & Juul, H. (2016). Predictors of early versus later spelling development in Danish. *Reading and Writing*, *29*, 245–266. doi:10.1007/s11145-015-9591-y

Verhaert, N. (2016). *Rules or regularities? The homophone dominance effect in spelling and reading regular Dutch verb forms* (Unpublished dissertation), University of Antwerp, Antwerp, Belgium.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, *32*, 251–276. doi:10.1016/S0095-4470(03)00032-9

Appendix

Verb stimuli with the Voicing Probability, Bigram Frequency, and Token Frequency Values and Outcomes per Grade and Task Format

Verb						Proportion Correct by Grade					
						Auditory infinitive format		Picture+cloze format		Written infinitive format	
Infinitive	Past tense	English translation	VP[a]	Bigram proportion	Token frequency	Grade 3	Grade 6	Grade 3	Grade 6	Grade 3	Grade 6
-te verbs											
blaffen	blafte	bark	0.081	0.154	218	0.27	0.65	0.26	0.78	0.31	0.73
blussen	bluste	extinguish	0.135	0.888	2	0.29	0.67	0.43	0.61	0.51	0.73
bukken	bukte	stoop	0	1	481	0.55	0.96	0.61	0.93	0.62	0.93
dansen	danste	dance	0.357	0.888	478	0.27	0.52	0.29	0.56	0.34	0.60
eisen	eiste	demand	0.755	0.888	522	0.19	0.42	0.37	0.52	0.40	0.62
juichen	juichte	cheer	0.992	1	121	0.09	0.36	0.21	0.50	0.20	0.44
koken	kookte	cook	0	1	177	0.38	1.00	0.49	0.98	0.43	0.98
lachen	lachte	laugh	0.953	1	4890	0.37	0.65	0.32	0.73	0.43	0.78
poetsen	poetste	brush	0.019	0.888	92	0.41	0.93	0.55	0.87	0.50	0.89
puffen	pufte	puff	0.081	0.154	31	0.38	0.73	0.43	0.78	0.41	0.80
scheppen	schepte	scoop	0.135	1	184	0.54	0.96	0.61	0.91	0.72	0.98
snoepen	snoepte	eat candy	0	1	15	0.63	1.00	0.71	0.93	0.63	0.93
surfen	surfte	surf	0.775	0.154	0	0.22	0.20	0.22	0.24	0.22	0.35
-de verbs											
durven	durfde	dare	0.755	0.846	1827	0.60	0.91	0.35	0.87	0.30	0.87
gooien	gooide	throw	1	1	1606	0.76	1.00	0.88	1.00	0.82	1.00
horen	hoorde	hear	1	0.995	9629	0.54	0.96	0.74	1.00	0.64	0.93
krabben	krabde	scratch	0.357	1	258	0.17	0.49	0.21	0.27	0.24	0.52
leven	leefde	live	0.992	0.846	1869	0.34	1.00	0.40	0.87	0.40	0.96
plonzen	plonsde	plash	0.357	0.112	25	0.54	0.73	0.28	0.65	0.26	0.71
reizen	reisde	travel	0.755	0.112	303	0.36	0.87	0.48	0.91	0.42	0.91
rennen	rende	run	1	0.994	1714	0.90	0.98	0.97	0.98	0.93	1.00
schrobben	schrobde	scrub	0.357	1	23	0.27	0.51	0.22	0.50	0.29	0.49
spugen	spuugde	spit	0.992	1	83	0.24	0.76	0.39	0.74	0.23	0.82
suizen	suisde	whizz	0.755	0.112	84	0.42	0.67	0.41	0.82	0.39	0.74
verbazen	verbaasde	surprise	0.755	0.112	207	0.26	0.82	0.31	0.82	0.33	0.80
volgen	volgde	follow	0.953	1	2903	0.70	0.87	0.82	1.00	0.78	0.96
zorgen	zorgde	take care	0.953	1	894	0.68	0.93	0.64	0.96	0.64	0.93

Note. [a]The table reports the actual VP value (with 0 referring to *-te* and 1 to *-de*). In the analyses, the score of 1 minus VP was used for the *-te* verbs.