Learning morphologically complex spoken words : Orthographic expectations of embedded stems are formed prior to print exposure

Beyersmann, Elisabeth, Wegener, Signy, Nation, Kate, Prokupzcuk, Ayako, Wang, Hua-Chen and Castles, Anne

Learning morphologically complex spoken words: Orthographic expectations of embedded stems are formed prior to print exposure

Elisabeth Beyersmann[1,2], Signy Wegener[1,2], Kate Nation[3], Ayako Prokupzcuk[4], Hua-Chen Wang[5,2] & Anne Castles[1,2]

[1]Department of Cognitive Science, Macquarie University, Sydney, Australia

[2]Macquarie University Centre for Reading, Macquarie University, Sydney, Australia

[3]Department of Experimental Psychology, University of Oxford, Oxford, UK

[4]Department of Linguistics, University of Leipzig, Leipzig, Germany

[5]Department of Educational Studies, Macquarie University, Sydney, Australia

Address for correspondence:

Elisabeth Beyersmann

Department of Cognitive Science

Australian Hearing Hub

16 University Avenue

Macquarie University NSW 2109

Tel: +61 406 249 916

Email: lisi.beyersmann@mq.edu.au

Author note: This is a pre-registered study: http://aspredicted.org/blind.php?x=zt7n8f.

Abstract

It is well known that information from spoken language is integrated into reading processes, but the nature of these links and how they are acquired is less well understood. Recent evidence suggests that predictions about the written form of newly learned spoken words are already generated prior to print exposure. We extend this work to morphologically complex words and ask if the information that is available in spoken words goes beyond the mappings between phonology and orthography. Adults were taught the oral form of a set of novel morphologically complex words (e.g., “neshing”, “neshed”, “neshes”), with a second set serving as untrained items. Following oral training, participants saw the printed form of the novel word stems for the first time (e.g., *nesh*), embedded in sentences, and their eye movements were monitored. Half of the stems were allocated a predictable and half an unpredictable spelling. Reading times were shorter for orally trained than untrained stems, and for stems with predictable than unpredictable spellings. Crucially, there was an interaction between spelling predictability and training. This suggests that orthographic expectations of embedded stems are formed during spoken word learning. Reading aloud and spelling tests complemented the eye movement data and findings are discussed in the context of theories of reading acquisition.

Keywords: spoken word learning, eye-tracking, morphological processing, reading acquisition

The question of how knowledge from spoken language is integrated within written language has concerned researchers for many years. When a child first learns to read, the ability to map letters onto sounds is the key to linking the written form with its spoken form and its meaning, and indeed, the correlation between children's spoken vocabulary and reading acquisition has been widely demonstrated (e.g., Duff, Reen, Plunkett, & Nation, 2015; Lee, 2011; Nation & Cocksey, 2009; Nation & Snowling, 2004). Even advanced readers sometimes encounter new words in text that they are already familiar with in oral form and have to make the link between print and spoken vocabulary (e.g., Johnston, McKague, & Pratt, 2004; McKague, Davis, Pratt, & Johnston, 2008). The extent to which this occurs varies with reading experience, but nonetheless the influence of spoken word knowledge on the acquisition of new written word forms remains an issue at all levels of literacy.

From a theoretical perspective, familiarity with the oral form of a word furnishes top-down support during the process of phonological decoding during reading. If the decoded phonological form does not match the phonology of any word in the lexicon, the partially decoded form can be modified until a phonologically similar word has been identified (e.g., Castles & Nation, 2006; Perfetti, 1992; Share, 1995). This mechanism provides one explanation for the association between spoken vocabulary and reading skills. An alternative explanation is that oral vocabulary assists the development of written word knowledge even before the word has been encountered in print (Johnston et al., 2004; McKague et al., 2008; Wegener et al., 2018). This does not falsify the partial decoding explanation, but instead suggests that predictions about written forms are already generated from spoken language; these can either facilitate reading, if the printed word matches the predicted form, or hinder it, if the printed form differs from which was predicted.

Wegener and colleagues (2018) provide evidence for the operation of this second mechanism in Grade 4 children, proposing that the children map a known phonological form onto an orthographic expectation, or “skeleton”. Participants were first trained on a list of novel spoken words (e.g., “nesh”, “bype”), and then participated in an eye-tracking experiment, in which they were exposed to the words in written form for the first time, together with a set of untrained words. The words were allocated spellings that were either predictable (*nesh*) or unpredictable (*bype*). It was found that words with predictable spellings were fixated for shorter periods of time than words with unpredictable spellings, but that this was particularly so when the words had been orally trained. Their study thus provides evidence that during oral word learning, children form a link between the phonological form of novel words and their expected orthography.

The words in Wegener et al.’s experiment were mono-morphemic. Yet, the vast majority of words in English are morphologically complex; that is, they comprise multiple morphemes such as a stem (*paint*) and an affix (*re* + *paint, paint* + *er*). Pre-school children show awareness of morphological structure in spoken language (Treiman & Cassar, 1997) and morphological awareness in primary school students is predictive of their reading comprehension (Carlisle, 2000; Deacon & Kirby, 2004; Kirby et al., 2012); it has also been linked with spelling development (Deacon & Bryant, 2006; Pacton & Deacon, 2008; Pacton, Foulin, Casalis, & Treiman, 2013). However, despite evidence for a relationship between morphological awareness, vocabulary and reading development, it is not known how children’s early oral morphological knowledge becomes integrated into their later written language knowledge. To address this, the present study extended Wegener et al.’s approach to morphologically complex words and asked if the information that is available in spoken words – the information that leads to the formation of an orthographic skeleton – goes beyond mappings between phonology and orthography to embody morphological structure.

Findings from studies of young children’s awareness of morphological structures suggest that the spoken language system sets up morphemic representations based on regularities between spoken complex words and their meanings (e.g., a *farmer* is someone who *farms*, a *painter* is someone who *paints*, etc.). Through repeated exposure to spoken form-meaning regularities, links are formed between morphologically complex whole words (*painter*) and their embedded stems (*paint*). Beginning readers are thus equipped with knowledge about the relationship between whole words and their embedded stems in the spoken domain, which is accessible from the early stages of reading development (Grainger & Beyersmann, 2017). This knowledge is what allows primary school children to assign meaning ("morphological analysis") and derive correct pronunciations ("morphological decoding") from morphologically complex words (e.g., Deacon, Tong, & Francis, 2017; Levesque, Kieffer, & Deacon, 2017), and is thus thought to be a "binding agent" in children's reading development (Kirby & Bowers, 2017). The important role of morphology is also evidenced by masked primed lexical decision tasks measuring children's online reading processes. For instance, second graders show significant priming from a complex prime to the embedded stem target in masked primed lexical decision (e.g., Beyersmann, Castles, & Coltheart, 2012; Beyersmann, Grainger, Casalis, & Ziegler, 2015; Beyersmann, Grainger, & Castles, 2019). This is evident with both affixed primes (*painter-paint*) and compound primes (*paintbrush-paint*), but typically absent for orthographically related prime-target pairs (*cashew-cash*; where *ew* is not an affix), suggesting that young children are sensitive to morphological structure when reading morphologically complex words.

What is less clear, however, is how exactly the links between morphemic representations and orthographic input are acquired. One possibility is that links between orthography and morphology are formed *after* exposure to the written form of the novel complex word, by attempting to map the written input word onto existing morphemic representations in the language system. This hypothesis is based on the idea that people are equipped with knowledge about the mappings between the spoken form of a complex word ([peɪntər]) and the semantic representations of both the whole word (*painter*) and its embedded stem (*paint*) that they can draw on when first seeing the word in print. Encounters with the orthographic unit (*painter*) will then lead to the addition of new links between the orthographic form and the phonological word form and the herewith associated semantic representations for both complex words and their embedded stems. An alternative, but not mutually exclusive option is that the links already begin to be formed *prior* to exposure to the written form. In line with Wegener et al.'s (2018) orthographic skeleton hypothesis, this second option predicts that when literate people learn novel complex words in the spoken domain, they generate an orthographic skeleton of any embedded morphemes before print exposure. That is, whenever a new spoken word form is encountered, the speech sounds of the embedded stem would be isolated and mapped onto a hypothesised orthographic form, such that links between phonological and visual word representations would already be acquired before having seen the word form in print.

The present study

The goal of the present study was to test the hypothesis that at least some orthographic knowledge about embedded stems is acquired prior to print exposure, by applying Wegener et al.'s (2018) novel word learning paradigm to the context of morphologically complex words. Adult participants were taught the oral forms of novel morphologically complex words ("neshing", "neshed", "neshes"), consisting of a novel word stem ("nesh") and an existing inflectional affix ("–s", "-ing", or "–ed"). During oral training, participants learned to associate pictures of inventions with the novel complex words forms (e.g. "This machine is used for neshing.") but never encountered the embedded morphemic units in isolation. Following training, participants read novel word stems (*nesh*), half trained and half untrained, for the first time. The words were embedded in sentences and eye movements were monitored as participants read silently. Spelling predictability was also manipulated, with half of the stems having a predictable spelling (*nesh*) and half an unpredictable spelling (*bype* – as opposed to the predictable counterpart *bipe*). Wegener et al.'s (2018) spelling predictability contrast was used as a marker effect for the generation of the expected orthographic forms of the embedded stem morphemes.

We hypothesised that if participants generate an orthographic skeleton, trained stems would have shorter looking times and would be less likely to be refixated than untrained stems, and stems with predictable spellings would have shorter looking times and be less likely to be refixated than unpredictable spellings. The key focus of our investigation was on the presence or absence of a training-by-spelling predictability interaction. If it is the case that links between printed complex words and their embedded stem morphemes are exclusively formed after print exposure, the effect of spelling predictability would not be modulated by oral vocabulary training. If, however, orthographic expectations of embedded stems already begin to form before print exposure, we would expect to see a significant training-by-spelling predictability interaction, showing that the difference between predictable and unpredictable spellings would be greater for trained than untrained items. We pre-registered these predictions along with our method, procedure and data analysis plans (http://aspredicted.org/blind.php?x=zt7n8f).

Following the sentence reading task, participants completed two additional pre-registered but more exploratory tasks on the trained and untrained stems: reading aloud and spelling dictation. As participants had already encountered – and potentially modified - the orthographic form of the novel word stems in the sentence reading task, it was less clear whether or not any effects of training or predictability would persist beyond first exposure.

Methods

Participants

Forty students from Macquarie University (30 female, 10 male; mean age: 21.7 years [SD: 5.0]), all English native speakers, participated for course credit. They were randomly split into two groups of 20. The first group was trained on Set 1 and the second group on Set 2. Participants were also assessed on a standardised reading fluency test (Test of Word Reading Efficiency [TOWRE]; Torgesen, Wagner, & Rashotte, 1999), Form A. Subtests for words and nonwords were administered, which measured the number of items named in 45 seconds. Raw scores were converted into age-based standard scores and percentiles, which revealed a mean standard score of 107 for word reading (SD: 11, percentile: 67) and 108 for nonword reading (SD: 12, percentile: 70).

Materials

Novel words. Materials consisted of a list morphologically complex novel spoken words (used during oral word training) and a list of morphologically simple novel written words (used during eye-tracking). Morphologically simple novel words were 32 three-phoneme monosyllabic nonwords, adapted from Wegener et al. (2018). Half of the items were assigned spellings that contained frequent phoneme to grapheme mappings and thus were highly predictable from their phonology (e.g., ‘g’ for /g/ as in *thog*). The other half were assigned spellings that were unpredictable as they contained less frequent mappings (e.g., ‘gg’ for /g/ as in *phegg*). To guide item development, the frequencies of phoneme to grapheme mappings were extracted from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993) and calculated based on the position-specific type and token count of individual phoneme-grapheme pairs. The spelling predictability manipulation was subsequently confirmed through a pre-experimental spelling dictation test (Wegener et al., 2018), showing that predictable items were spelled in the same way by all pilot participants (e.g., *nesh, coib*) while the unpredictable items were not spelled in that way by any participant (e.g., *veme, koyb*). The average logarithmic bigram frequency was slightly higher in the predictable condition (mean = 1.98 [SD = 0.77]) than in the unpredictable condition (mean = 1.73 [SD = 0.54]), but matched across training sets (Set 1 mean = 1.83 [SD = 0.71]; Set 2 mean = 1.87 [SD = 0.44]). Despite the variation in spelling predictability, all items were regular for reading in that they could be read aloud correctly using common grapheme-phoneme correspondences. Items were split into two different sets (Sets 1 and 2), which were matched for consonant/vowel structure. Set 1 served as trained items during oral exposure for half of our participants, and Set 2 for the other half (see Appendix A).

Morphologically complex words were created by combining novel stems ("nesh") with three different suffixes ("-ing", "-ed", "–s"), resulting in three complex forms for each stem ("neshing", "neshed", "neshes"). Inflectional rather than derivational suffixes were used to ensure that the addition of the affixes did not change the syntactic function of the target items and that syntactic structure and target word position were matched across sentences. Suffixes were represented by different allomorphs, depending on phonological context (e.g., the past-tense suffix /t/ was added to stems ending in voiceless consonants, and /d/ to stems ending in voiced consonants).

Oral training. Participants were trained in small groups of 2-4 individuals. They were told that they would be learning about 'Professor Parsnip's Inventions' and engaged in a range of activities to learn about the function and perceptual features of each invention. For example, they learned that "Professor Parsnip has invented a machine that neshes. It is used to take out the food you don't like from a meal. It has a tube and two open ends." Each invention was paired with a picture demonstrating its features (see Figure 1).

Eye-tracking sentences. Word stems (*nesh*) were embedded in a carrier sentence, and always occurred in mid-sentence position (Appendix B). For example: *Max put his food into the machine to nesh the green peas.* Following Wegener et al.'s procedure, carrier sentences were designed to be contextually rich, such that as participants read them, they would expect to see the word they had learned about during oral vocabulary training.

Procedure

Oral vocabulary training took place over three consecutive days (~30 minutes/session) to limit the learning load on participants at any one time. Upon completion of the last vocabulary training session, participants completed a range of additional tasks (see Table 1), which are summarised below.

Oral vocabulary training. This followed closely the procedure used by Wegener et al. (2018). Each participant was trained on one set of 16 complex novel words (either Set 1 or Set 2), with the other set constituting their untrained items. Sets were counterbalanced across groups. Eight items (four from each spelling predictability condition) were introduced and rehearsed twice on Day 1, the remaining eight were introduced and rehearsed twice on Day 2, and all 16 items were trained and rehearsed in the last session on Day 3. During rehearsal, participants were briefly reminded of each invention's meaning and then asked to repeat the associated novel word forms.

Picture-naming (post-training check). This was to check that participants had learned the spoken forms and their meanings. Participants were individually shown pictures of the inventions and asked what the invention did and what it was used for. Accuracy was recorded for remembering the novel word and its meaning.

Eye-tracking experiment (first orthographic exposure). In the eye-tracking experiment, participants encountered the novel stems for the first time, embedded in sentences. All sentences appeared on a single line. Eye movements were monitored as sentences were read silently; 16 sentences contained reference to inventions they had learned about (i.e., 'trained' stems) and 16 reference to inventions learned by the other group (i.e., 'untrained' stems). Half contained predictable spellings (*nesh*) and half unpredictable spellings (*bype*). An additional four filler sentences were included with novel words not learned by either group.

Eye movements were recorded using an Eyelink 1000 eye tracker (SR Research; Mississauga, Canada) sampling at 1000 Hz as participants read sentences on a computer monitor at a viewing distance of 85 cm. Each character covered 0.30˚ of horizontal visual angle. Sentences were presented in black, Courier New font on a white background. Participants read binocularly but only the movements of the right eye were monitored. A nine-point initial calibration of the eye tracker was performed (maximum average error of 0.30), followed by three practice trials, and then the experimental sentences. The experimenter triggered the beginning and end of each trial after the participants looked at a fixation cross to indicate their readiness. To promote attention to task, they were required to answer a (yes/no) question after each trial. Participants' average response accuracy was 95.4%.

Eye movement dependent variables were extracted, capturing reading behaviour on the target word: first fixation duration (duration of initial fixation on the target); gaze duration (sum of all fixations made on the target before the eyes move past the target to a subsequent word within the sentence); total reading time (sum of all fixations on the target, including any regressions back to it); and regressions in (probability of making a regression back to the target from a later portion in the sentence).

Reading aloud (second orthographic exposure). Participants read aloud all 16 trained and 16 untrained word stems, presented individually in the centre of a computer screen using DMDX software (Forster & Forster, 2003). Each trial consisted of a 800-ms fixation cross followed by the target, which remained until response or until 2 seconds had elapsed. Participants were instructed to name each word as quickly and accurately as possible, while reaction times and response accuracy were assessed.

Spelling dictation. The experimenter read out each novel stem and participants were instructed to spell them exactly as they were written in the sentence reading and reading aloud parts of the experiment.

Results

Assessment of oral vocabulary learning: picture naming

Participants correctly recalled 14.2 of the 16 orally trained invention verbs (*SD*=2.67). The difference in recall between participants who learned Set 1 (*M*=14.63, *SD*=2.19) and Set 2 (*M*=13.84, *SD*=3.02) was not significant (*t*(33)=0.86, *p*=.39), nor was the difference in recall for items allocated predictable (*M*=7.09, *SD*=1.12) and unpredictable (*M*=7.11, *SD*=1.75) spellings (*t*(34)=-0.14, *p*=.89). These results indicate that participants successfully learned the complex novel words and were able to match the spoken word forms with pictures of the inventions.

Eye movements

Very brief or very long (<80 milliseconds or >1200 milliseconds) eye fixations were deleted, as were trials that showed blinks or tracker loss on the target word (2.1% of the eye movement data). As set out in our pre-registration document, data were log transformed and analysed in the R computing environment (RDevelopmentCoreTeam, 2019). Linear mixed effects models were constructed using the lme4 package (Bates, Maechler, Bolker, & Walker, 2014) and *p* values were obtained using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). Models were run for each dependent variable: first fixation duration, gaze duration, total reading time and regressions in (see Figure 2). The area of interest was the trained or untrained stem. In all analyses, training (trained vs. untrained), spelling predictability (predictable vs. unpredictable) and their interaction were entered as fixed effects while participants and items were entered as random effects. Fixed effects and their interaction were centred using contrast coding. Fitting models with full random effects structure (random intercepts and slopes for subjects and items) produced singular fits, suggesting that models were over parameterised (Baayen, Davidson, & Bates, 2008). Reducing the random effects structure eliminated the problem of singularity. We selected the highest nonsingular converging models. For time data, overly influential observations were identified using the influence.ME package (Nieuwenhuis, te Grotenhuis, & Pelzer, 2012) and removed. Thirteen observations were removed for first fixation duration, twelve from gaze duration and six from total reading time. Removal of these observations did not change the direction or significance of the findings.When an interaction was significant, the phia package (Rosario-Martino & Fox, 2015) was employed to compute contrasts.

The model for first fixation duration revealed a significant effect of spelling predictability (β=-0.074, SE=0.025, t=-3.013, p=.006), showing that fixation durations were shorter for items with predictable than unpredictable spellings. The effect of training (β=-0.019, SE=0.023, t=-0.864, p=.395) and the interaction beween spelling predictability and training were not significant (β=-0.032, SE=0.045, t=-0.702, p=.488).

In the models for gaze duration and total reading time, each of the three effects was significant. The effect of spelling predictability (gaze duration: β=-0.192, SE=0.043, t=-4.425, p<.001; total reading time: β=-0.311, SE=0.050, t=-6.152, p<.001) indicated that fixation durations were shorter for items with predictable than unpredictable spellings. The effect of training (gaze duration: β=-0.058, SE=0.025, t=-2.318, p=.028; total reading time: β=-0.088, SE=0.036, t=-2.419, p=.022) indicated that fixation durations were shorter for trained than untrained items. Finally, there was an interaction between spelling predictability and training (gaze duration: β=-0.168, SE=0.050; t=-3.343, p=.002; total reading time: β=-0.318, SE=0.072, t=-4.397, p<.001). Interaction contrasts showed that items with predictable spellings benefited from training (gaze duration: $\chi^2 = -0.142$, $p < .001$; total reading time: $\chi^2 = -0.247$, $p < .001$), whereas items with unpredictable spellings did not (gaze duration: χ^2 =

0.057, $p = .450$; total reading time: $\chi^2 = 0.071$, $p = .153$). The effect of spelling predictability was present both for items that had received oral training (gaze duration: $\chi^2 = -0.276$, $p < .001$; total reading time: $\chi^2 = -0.470$, $p < .001$) and for items that had not (gaze duration: $\chi^2 = -0.108$, $p = .026$; total reading time: $\chi^2 = -0.151$, $p = .007$).

In the model reflecting the probability of regressions back to the target word, the effects of spelling predictability (β=-0.492, SE=0.279, z=-1.765, p=.078) and training (β=-0.143, SE=0.170, z=-0.862, p=.389) were not significant. The interaction between spelling predictability and training was significant (β=-0.908, SE=0.279, z=-2.868 , p=.004). Interaction contrasts showed that the effect of spelling predictability was present for items that had received oral training ($\chi^2 = 0.280$, $p = .007$), with predictable spellings being less likely to be reread than unpredictable spellings. There was no effect of spelling predictability when participants had not received oral training ($\chi^2 = 0.490$, $p = .904$). Items with predictable spellings benefited from training($\chi^2 = 0.354$, $p = .034$), whereas items with unpredictable spellings did not ($\chi^2 = 0.576$, $p = .144$).

If any of the three prespecified interest areas (target word, pre-target text, post-target text) was skipped during first pass reading, the trial was removed prior to analysis. There was a high rate of skipping during first pass reading (18.90% of targets). In view of the high skipping rate on first pass reading, we deviated from our preregistered analysis plan to run an additional exploratory model with skipping probability as the dependent variable. The model revealed a significant effect of spelling predictability (β=1.151, SE=0.266, z=4.328, p<.001), showing that skipping was more likely to occur for items with predicatable than unpredictable spellings, presumably because predictable words were easier to decode in the parafovea due to their slightly shorter length and more common spellings (Figure 3). The effect of training (β=0.083, *SE*=0.283, *z*=0.292, *p*=.770) and the interaction of spelling predictability and training were not significant (β=-0.118, *SE*=0.408, *z*=-0.298, *p*=.773).[1]

Reading aloud

Two participants were excluded, because error rates were above 40%. Incorrect responses were removed from the RT analysis (10.4% of all data). We used linear mixed-effect modelling to perform the main analyses, including factors spelling predictability (predictable, unpredictable), training (trained, untrained), the interaction between spelling predictability and training, and random intercepts for subjects and items. Inverse RTs (1/RT) were calculated for each participant to correct for RT distribution skew. The model was refitted after excluding data-points whose standardised residuals were larger than 2.5 in absolute value (1.6% of all data; Baayen, 2008). Post-hoc comparisons were carried out using cell means coding and single df contrasts with the glht function of the multcomp package (Version 1.4-8; Hothorn, Bretz, & Westfall, 2008) using the normal distribution to evaluate significance. RT analyses revealed a significant effect of spelling predictability ($X^2(1)$=35.36, *p*<.001), showing that predictable novel words were read faster than unpredictable novel words, and a significant effect of training ($X^2(1)$=22.61, *p*<.001), indicating that trained novel words were read faster than untrained novel words. There was also a significant interaction between spelling predictability and training ($X^2(1)$=6.21, *p*=.013), which went in the opposite direction of the interaction seen in the eye-tracking data, showing that the spelling predictability effect was reduced for trained compared to untrained items (Figure 4, left panel). Post-hoc contrasts showed a significant training effect in the predictable ($z = 4.76$, $p < .001$) and unpredictable condition ($z = 7.82$, $p < .001$), as well as a significant spelling

---

[1] We reran the analyses for the eye movement data excluding the two participants whose data was removed from the reading aloud analyses. These models additionally excluded the item "vaype" (because a reviewer noted that "vape" commonly known to describe smoking an e-cigarette) and also included a new covariate "allomorph" (to control for the realisation of the plural form of the words experienced during training). These additional analyses confirmed the direction and significance of our original findings.

predictability effect in the trained (z = 7.78, p < .001) and untrained condition (z = 5.95, p < .001).

Error rates were analysed by applying a binomial variance assumption to the trial-level binary data using the function glmer as part of the R-package lme4. As in the RT analyses, there was a significant interaction between spelling predictability and training ($X^2(1)=4.50$, $p=.034$), showing that participants made more errors reading items with unpredictable than predictable spellings, and that this effect was reduced for trained compared to untrained items (Figure 4, right panel). Post-hoc contrasts showed a significant training effect for items with unpredictable (z = 5.95, p < .001) but not predictable spellings (z = 1.41, p = .157). The effect of spelling predictability was significant in the trained condition (z = 4.98, p < .001) and marginally significant in the untained condition (z = 1.68, p = .093).

Spelling Dictation

Mean error rates across participants were calculated for each condition (Figure 5). A linear-mixed effects model was fitted including training (trained, untrained), spelling predictability (predictable, unpredictable), the interaction between training and spelling predictability, and random intercepts for subjects and items. The results revealed a significant effect of predictability ($X^2(1)=44.64$, $p<.001$). Neither the effect of training nor its interaction with spelling predictability was significant ($X^2(1)=0.06$, $p=.803$; $X^2(1)=0.25$, $p=.616$).

Discussion

We combined a spoken word learning paradigm with the recording of eye-movements to examine the mechanisms involved in integrating spoken word knowledge into online reading processes. For three consecutive days, participants were trained on a set of morphologically complex spoken novel words (“neshing”, “neshed”, “neshes”), half of which were later assigned to the predictable spelling condition and half to the unpredictable condition. On day three, participants read for the first time the written form of the embedded word stems (*nesh)* while their eye-movements were recorded. Each participant read trained and untrained words with both predictable and unpredictable spelling patterns. Planned analyses of the eye movement data revealed a main effect of training which was significant in the gaze duration and total reading time analyses, showing that participants spent overall less time fixating trained than untrained items. There was also a robust effect of spelling predictability on all four dependent variables (first fixation duration, gaze duration, total reading time, regressions in), showing that reading times were overall shorter for predictable than unpredictable stems and that items with predictable spellings were less likely to be reread than unpredictable spellings. The consistency of the spelling predictability effect throughout the analyses demonstrates that participants naturally relied on their prior knowledge about spelling regularities when decoding the novel word stems.

The most striking finding of our investigation concerns the interaction seen between training and spelling predictability, which significantly affected gaze duration (the sum of all fixations made on the target word before the eyes move past the target to a subsequent word within the sentence), total reading time (the sum of all fixations on the target word, including any regressions back to it), and regressions in (the probability of making a regression back to the target from a later portion in the sentence). Figure 2 highlights this important finding (top right panel and bottom left panel) and also shows that, although the interaction did not reach significance for first fixation duration (top left panel), the response pattern was comparable across all four dependent variables.[2] The interaction between training and spelling

---

[2] Gaze duration and first fixation duration are both first pass measures and therefore best capture the effects of initial orthographic exposure, with gaze duration reflecting a slightly more thorough analysis of the target letter strings. The fact that the interaction was not only significant in gaze duration, but also in total reading time and regressions in suggests that the orthographic skeleton effect persisted beyond the first pass, and continued to affect the later reading processes.

predictability suggests that participants developed an orthographic expectation (or skeleton) of the embedded novel stem *prior* to print exposure, consistent with Wegener et al.'s (2018) findings for mono-morphemic words. Whenever there was a match between the orthographic prediction and the printed word form (as was the case in the predictable trained condition) there was a clear benefit in reading times compared to when there was a mismatch between the orthographic prediction and the actual printed form (as was the case in the unpredictable trained condition).

More than that, our findings demonstrate that the orthographic skeleton hypothesis is not just specific to the words that participants were exposed to during oral vocabulary training (as in Wegener et al., 2018), but also applies to embedded words not heard or seen before. This suggests that the acquisition of oral vocabulary induces the setup of a complex set of orthographic predictions, not just for whole words, but also for stems embedded in morphologically complex words. For this to happen, two important processes must be at play during oral vocabulary acquisition: the process of decomposing complex novel words into embedded word stem and suffix, and the process of generating an orthographic prediction for the embedded novel word stems. One possibility is that participants used their existing knowledge about spoken form-meaning regularities from language to map the morphologically complex input words ("neshing", "neshed", "neshes") onto the embedded word stem ("nesh"), and then derive an orthographic prediction of *nesh* from the spoken form of its embedded stem. Another possibility is that the morphologically complex spoken words were initially mapped onto an orthographic skeleton of the whole item, which then induced an orthographic skeleton of the embedded stem. Either way, it is clear from the present findings that skilled readers are sensitive to both morphological structure and probabilistic orthographic patterns as they encounter new complex spoken words, resulting in an orthographic skeleton effect when the embedded stem is later encountered in print.

A second key finding is the significant interaction between training and spelling predictability that was observed in the reading aloud data, when participants encountered the written form of the embedded stem for the second time. The reading aloud task was administered following the eye-tracking experiment to test the robustness of the orthographic skeleton beyond initial print exposure. Interestingly, this training-by-spelling predictability interaction went in the exact opposite direction of the interaction observed in the eye-tracking data. As Figure 4 indicates, the effect of spelling predictability was significantly reduced in the trained compared to the untrained items. One way in which these results can be explained is that initial exposure to the orthographic form of the embedded stem within the sentence reading task led to a modification (or correction) of the previously predicted orthographic skeleton in the trained condition. The correction of the orthographic form would have been particularly relevant for items with low spelling predictability, because - as the eye movement data demonstrate - the mismatch between the orthographic skeleton and the actual printed item led to overall longer fixation times than for items with high spelling predictability. For instance, oral training of a spoken item assigned to an unpredictable spelling ([baɪp]) would have led to a mismatch between the predicted spelling (*bipe*) and the actual spelling (*bype*), slowing reading times in the eye movement data. Crucially, the reading aloud data demonstrate that the predicted printed form ([baɪp]-*bipe*) was quickly updated during initial print exposure by replacing the predicted printed form with the encountered printed form ([baɪp]-*bype*). This may be why the difference between the predictable and unpredictable items was significantly smaller in the trained compared to the untrained condition, in which participants had no prior knowledge of the associated spoken word form and therefore had greater difficulty in mapping the untrained unpredictable spellings onto sounds.

The combined eye movement and reading aloud data clearly support the idea that participants developed a lexical representation of the embedded word stems, despite never encountering them in isolation during oral vocabulary training. Knowledge about embedded stems from spoken language affected both silently reading words in text (as evidenced by the eye movement data), and reading aloud individual target words (as evidenced by the reading aloud data). Our study thus provides direct evidence for the hypothesis that orthographic learning of novel words is facilitated by prior morphological knowledge from oral language, and also highlights the speed by which orthographic representations are acquired and updated by skilled readers. One question that the current study does not address is whether the acquisition of embedded word skeletons is limited to embedded words occurring in a genuine morphological context (e.g. “neshed”), or also applies in situations where words are embedded in non-affixed novel words (e.g. “neshel”). An informative follow-up of our present results would therefore be a study examining novel words consisting of combinations of stems and non-morphemic endings, to more precisely determine the kind of mechanisms that are used to extract the stem during oral vocabulary exposure. Moreover, whether or not the present investigation of morphemic inflections generalises to novel word derivations (e.g. “nesher”), which can modify both meaning and syntactic category of the embedded stem, remains a key question for future research.

It is worth noting that the pattern of the reading aloud data was not mirrored in the results of the spelling dictation test. Although the spelling dictation data revealed a main effect of spelling predictability, a training effect was entirely absent in these data. Of course, the spelling dictation test involved the challenging task of recalling the spellings of items that participants had previously only encountered twice. As the results show, the mean error rates in this task were high, particularly in the unpredictable spelling condition (see Figure 5), suggesting that participants probably did not have sufficient exposure to the written forms to recall the correct spellings.

Although our current investigation was with skilled adult readers, our findings shed light onto the mechanisms involved in developing representations for novel word stems and hence make the prediction that children, who are still in the process of learning to read, should benefit from their prior knowledge from spoken language when learning to read morphologically complex words. This ties in with recent work showing that morphology forms a linkage among phonological, orthographic, and semantic information (Kirby & Bowers, 2017, 2018; Rastle, 2018; Ulicheva, Harvey, Aronoff, & Rastle, 2018) and therefore facilitates the acquisition of novel words beyond the single word level (see also J. S. Bowers & Bowers, 2018; P. N. Bowers, Kirby, & Deacon, 2010, who have been at the forefront of testing principles of explicit morphological instruction in the classroom). The current data also fit with the theoretical framework proposed by Grainger and Beyersmann (2017) suggesting that the activation of embedded stems in printed words acts as a bootstrapping mechanism for morphological parsing in the developing reading system. The authors argue that embedded stem representations provide an important pre-requisite for the development of an automatic morphological segmentation mechanism. This idea finds further support from the results of a recent masked primed lexical decision study (Beyersmann et al., 2019), showing evidence for the early, automatic activation of embedded stems in children as young as Grade 3. Significant embedded stem priming effects were observed with compound primes (e.g. *farmhouse-farm*) and pseudo-compound primes (e.g. *butterfly-butter*; where prime and target shared a pseudo-morphological, but not a semantic relationship), but not when preceded by an orthographic control prime (e.g. *sandwich-sand*, where ‘wich’ is not a morpheme). These robust embedded stem priming effects in children with still limited reading experiences demonstrate that the acquisition of embedded stem representations denotes a relatively early milestone in children’s reading acquisition. The present study provides an explanation as to why this might be the case, suggesting that the facility to induce orthographic skeletons for embedded stems during oral language exposure equips children (once literate) with knowledge about morphological reading units from the early stages of written language acquisition.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (CD-ROM).

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-5. http://CRAN.R-project.org/package=lme4.

Beyersmann, E., Castles, A., & Coltheart, M. (2012). Morphological processing during visual word recognition in developing readers: Evidence from masked priming. *The Quarterly Journal of Experimental Psychology, 65*(7), 1306–1326.

Beyersmann, E., Grainger, J., Casalis, S., & Ziegler, J. C. (2015). Effects of reading proficiency on embedded stem priming in primary school children. *Journal of Experimental Child Psychology, 139*, 115-126.

Beyersmann, E., Grainger, J., & Castles, A. (2019). Embedded stems as a bootstrapping mechanism for morphological parsing during reading development. *Journal of Experimental Child Psychology, 182*, 196-210. doi: 10.1016/j.jecp.2019.01.010

Bowers, J. S., & Bowers, P. N. (2018). The importance of correctly characterizing the English spelling system when devising and evaluating methods of reading instruction. Comment on Taylor, Davis, and Rastle (2017). *Quarterly Journal of Experimental Psychology*.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80*, 144-179.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal, 12*, 169–190.

Castles, A., & Nation, K. (2006). How does orthographic learning happen? In S. Andrews (Ed.), *From inkmarks to ideas: Challenges and controversies about word recognition and reading* (pp. 151-179). London, UK: Psychology Press.

Deacon, S. H., & Bryant, P. (2006). This turnip's not for turning: Children's morphological awareness and their use of root morphemes in spelling. *British Journal of Developmental Psychology, 24*, 567–575.

Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness: Just ‘‘more phonological”? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics, 25*, 223–238.

Deacon, S. H., Tong, X., & Francis, K. (2017). The relationship of morphological analysis and morphological decoding to reading comprehension. *Journal of Research in Reading, 40*(1), 1-16. doi: 10.1111/1467-9817.12056

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? . *Journal of Child Psychology and Psychiatry, 56*, 848–856.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116-124.

Grainger, J., & Beyersmann, E. (2017). Edge-aligned embedded word activation initiates morpho-orthographic segmentation. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 285–317).

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal: Biometrische Zeitschrift, 50*, 346–363.

Johnston, M. B., McKague, M., & Pratt, C. (2004). Evidence for an automatic code in the processing of visually novel word forms. *Language and Cognitive Processes, 19*, 273–317.

Kirby, J. R., & Bowers, P. N. (2017). Morphological instruction and literacy: Binding phonological, orthographic, and semantic features of words. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 437-462). Amsterdam, NL: John Benjamins Publishing Company.

Kirby, J. R., & Bowers, P. N. (2018). The effects of morphological instruction on vocabulary learning, reading, and spelling. . In R. Berthiaume, D. Daigle, & A. Desrochers (Eds.), *Morphological processing and literacy development: Current issues and research* (pp. 217-243). New York: Routledge.

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing, 25*(2), 389-410.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13), 1-26.

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics, 32*, 69–92.

Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2017). Morphological awareness and reading comprehension: Examining mediating factors. *Journal of Experimental Child Psychology, 160*, 1-20. doi: 10.1016/j.jecp.2017.02.015

McKague, M., Davis, C., Pratt, C., & Johnston, M. B. (2008). The role of feedback from phonology to orthography in orthographic learning: An extension of item-based accounts. *Journal of Research in Reading, 31*, 55–76.

Nation, K., & Cocksey, J. (2009). Beginning readers activate semantics from sub-word orthography. *Cognition, 110*(2), 273-278. doi: 10.1016/j.cognition.2008.11.004

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*, 342–356.

Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B. (2012). influence.ME: Tools for Detecting Influential Data in Mixed Effects Models. *The R Journal, 4*(2), 38-47.

Pacton, S., & Deacon, S. H. (2008). The timing and mechanisms of children's use of morphological information in spelling: A review of evidence from English and French. *Cognitive Development, 23*, 339-359.

Pacton, S., Foulin, J. N., Casalis, S., & Treiman, R. (2013). Children benefit from morphological relatedness when they learn to spell new words. *Frontiers in Psychology, 4*.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. Gough (Ed.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Erlbaum.

Rastle, K. (2018). The place of morphology in learning to read in English. *Cortex*. doi: https://doi.org/10.1016/j.cortex.2018.02.008

RDevelopmentCoreTeam. (2019). *R: A language and environment for statistical computing*: R version 3.6.0.

Rosario-Martino, H., & Fox, J. (2015). Package 'phia'.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151-218.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency*. Austin, TX: PRO-ED Publishing, Inc. .

Treiman, R., & Cassar, M. (1997). Spelling Acquisition in English. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell : research, theory and practice*. Mahwah, NJ: Erlbaum.

Ulicheva, A., Harvey, H., Aronoff, M., & Rastle, K. (2018). Skilled readers' sensitivity to meaningful regularities in English writing. *Cognition*. doi: 10.1016/j.cognition.2018.09.013

Wegener, S., Wang, H. C., de Lissa, P., Robidoux, S., Nation, K., & Castles, A. (2018). Children reading spoken words: interactions between vocabulary and orthographic expectancy. *Developmental Science, 21*(3), 1-9. doi: 10.1111/desc.12577

Table 1

Table 1. Testing procedure involving three consecutive days of oral vocabulary training.

Test components		Materials	Time
Day 1	Oral vocabulary training Rehearsal 1 Rehearsal 2	8 items	30 min.
Day 2	Oral vocabulary training Rehearsal 1 Rehearsal 2	8 items	30 min.
Day 3	Oral vocabulary training Rehearsal Reading proficiency test Picture-naming test Eye tracking experiment Reading aloud Spelling dictation	all 16 items	90 min.

Figure 1

Figure 1. Example of a picture used during oral vocabulary training. A machine that is used to ‘nesh’ the food you do not like from a meal.

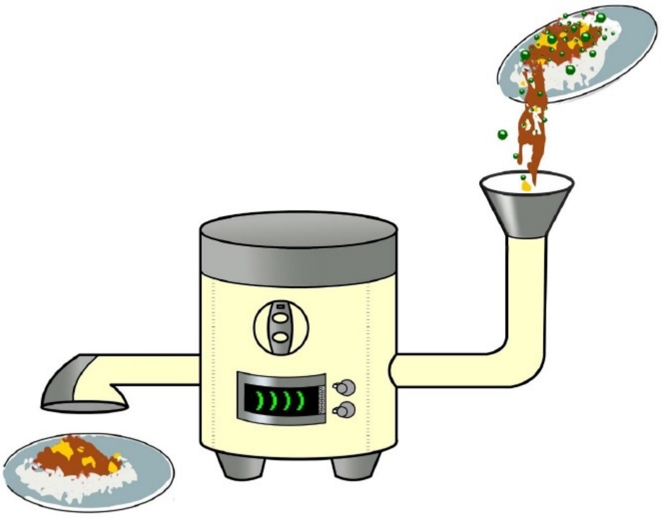


Figure 2

Figure 2. Arithmetic (untransformed) means and standard errors of eye movements in the target word interest area. First fixation duration, gaze duration and total reading time are expressed in milliseconds while probability of regressions reflects the likelihood of occurrence.

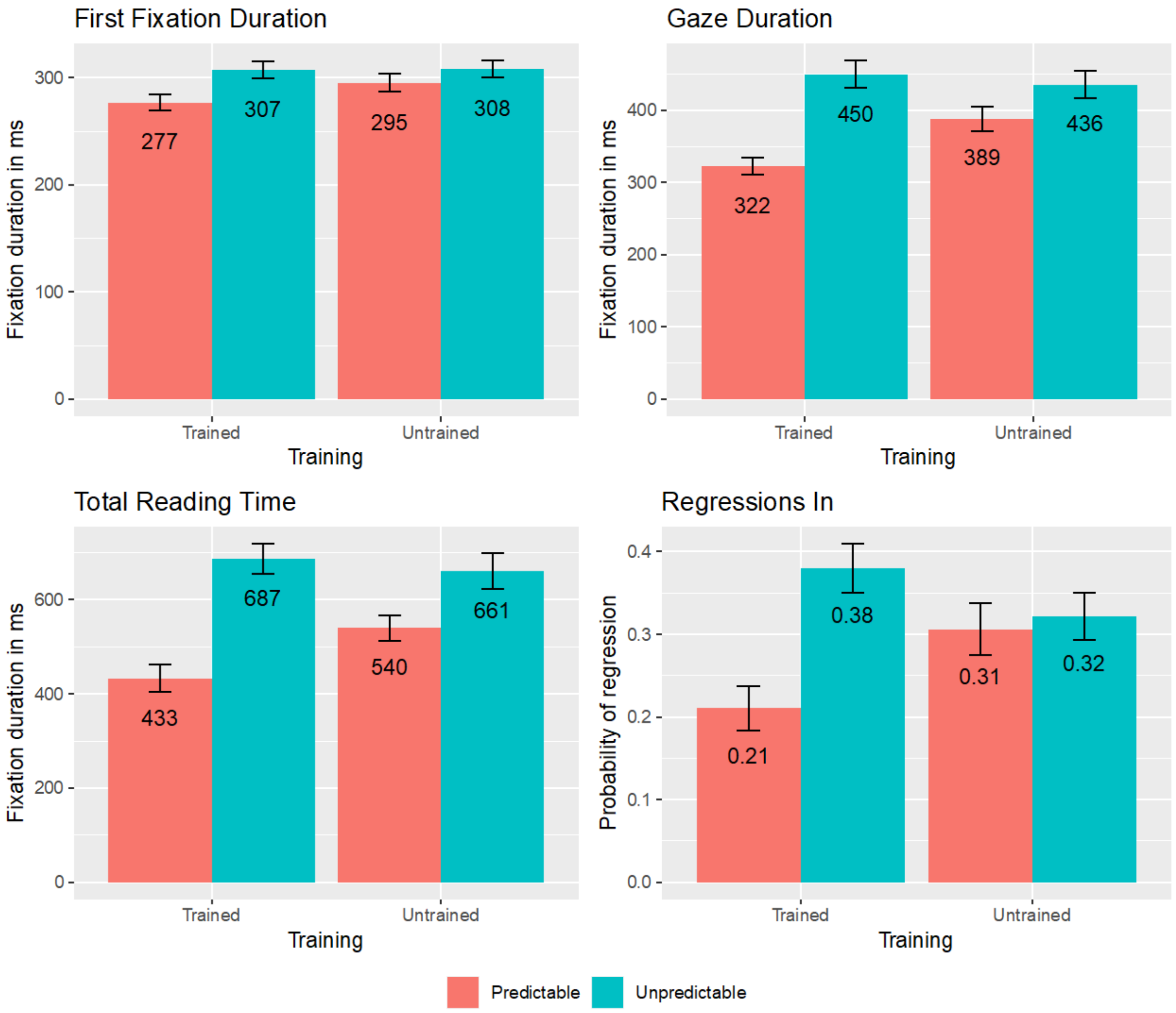


Figure 3

Figure 3. Means and standard errors of skipping of the target word on first pass reading, reflecting the probability of occurrence.

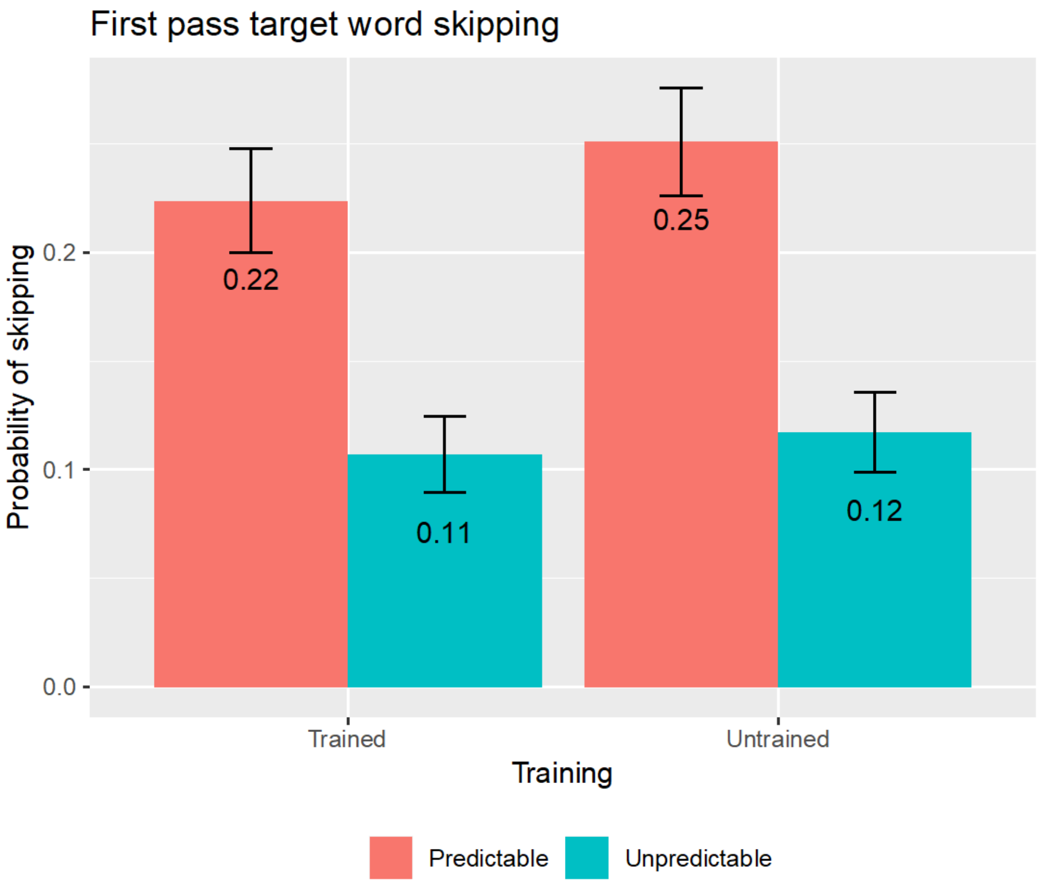


Figure 4

Figure 4. Results of the reading aloud task. Mean reaction times (ms) for each condition are presented in the left panel and mean error rates (%) in the right panel.

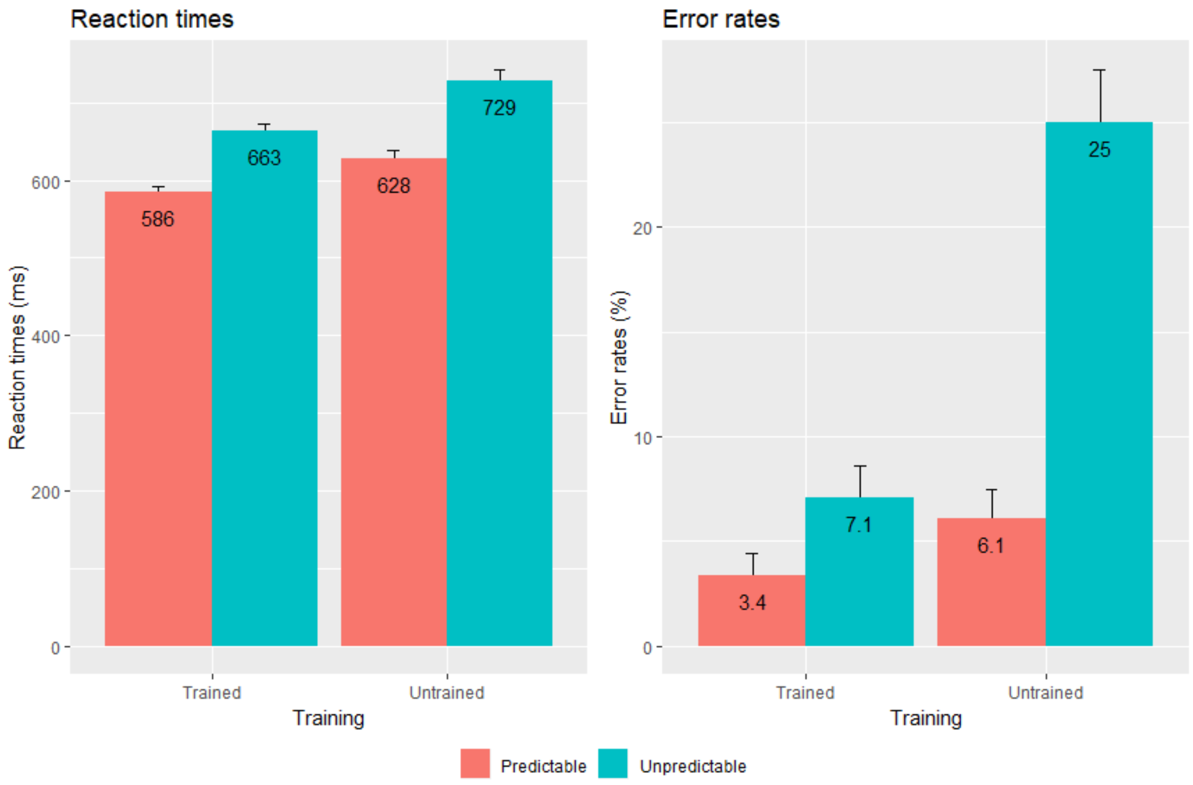
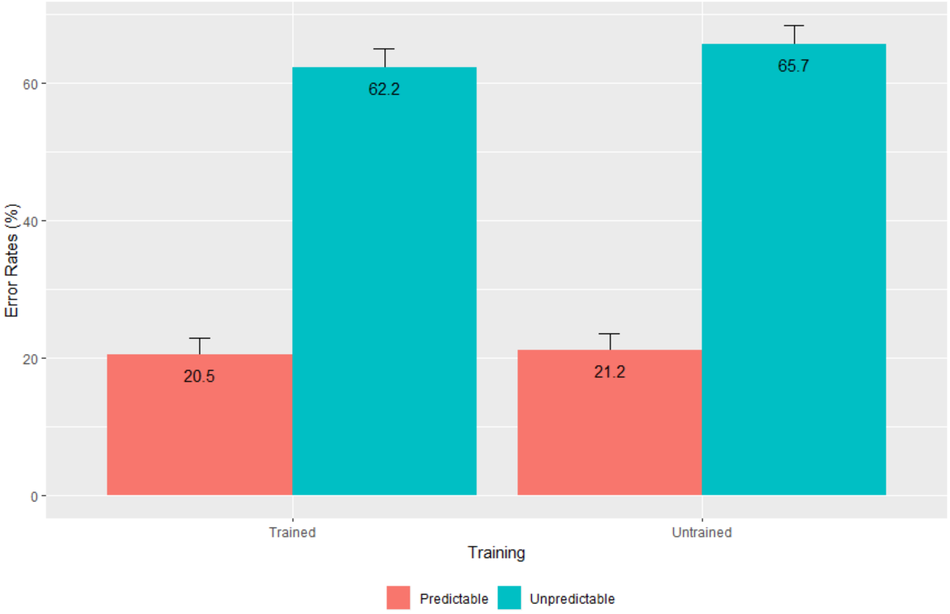


Figure 5

Figure 5. Results of the spelling dictation test.



Appendix A

	Set 1		**Set 2**	
	complex words	stem morphemes	complex words	stem morphemes
Predictable	/dʒevɪŋ/ /dʒevd/ /dʒevz/	jev	/temɪŋ/ /temd/ /temz/	tem
	/jægɪŋ/ /jægd/ /jægz/	yag	/nɪdɪŋ/ /nɪdəd/ /nɪdz/	nid
	/vɪbɪŋ/ /vɪbd/ /vɪbz/	vib	/dʒɪtɪŋ/ /dʒɪtɪd/ /dʒɪts/	jit
	/tʌpɪŋ/ /tʌpt/ /tʌps/	tup	/jæbɪŋ/ /jæbd/ /jæbz/	yab
	/neʃɪŋ/ /neʃt/ /neʃɪz/	nesh	/vɪʃɪŋ/ /vɪʃt/ /vɪʃɪz/	vish
	/tʃɒbɪŋ/ /tʃɒbd/ /tʃɒbz/	chob	/ʃepɪŋ/ /ʃept/ /ʃeps/	shep
	/ʃʌgɪŋ/ /ʃʌgd/ /ʃʌgz/	shug	/θɒgɪŋ/ /θɒgd/ /θɒgz/	thog
	/θʌbɪŋ/ /θʌbd/ /θʌbz/	thub	/tʃɪgɪŋ/ /tʃɪgd/ /tʃɪgz/	chig
Unpredictable	/viːmɪŋ/ /vɪːmd/ /vɪːmz/	veme	/juːnɪŋ/ /juːnd/ /juːnz/	yune
	/baɪpɪŋ/ /baɪpt/ /baɪps/	bype	/kaɪvɪŋ/ /kaɪvd/ /kaɪvz/	kyve
	/jɜːpɪŋ/ /jɜːpt/ /jɜːps/	yirp	/bɜːvɪŋ/ /bɜːvd/ /bɜːvz/	birv
	/kɔɪbɪŋ/ /kɔɪbd/ /kɔɪbz/	koyb	/dʒaɪfɪŋ/ /dʒaɪft/ /dʒaɪfs/	jayf
	/dʒiːbɪŋ/ /dʒiːbd/ /dʒiːbz/	jeabb	/miːfɪŋ/ /miːft/ /miːfs/	meaph
	/fɜːfɪŋ/ /fɜːft/ /fɜːfs/	phirf	/gʌzIŋ/ /gʌzd/ /gʌzɪz/	ghuzz
	/gækɪŋ/ /gækt/ /gæks/	ghakk	/fegɪŋ/ /fegd/ /fegz/	phegg
	/mɜːbɪŋ/ /mɜːbd/ /mɜːbz/	mirbe	/veɪpɪŋ/ /veɪpt/ /veɪps/	vaype

Appendix B

Set 1	
1	Rick put his dirty socks into the machine to jev them clean.
2	Diana put the best orange on the machine to veme the juice.
3	Pam put the dirty flowers under the machine to yag them shiny.
4	Max put his food into the machine to bype the yucky peas.
5	Sara put her soaking wet hat on the machine to vib it dry.
6	Lucy loaded all the rubbish into the machine to yirp it for recycling.
7	Lucas put his sore tummy beside the machine to tup it better again.
8	Jennifer put all her soggy chips under the machine to koyb them crispy.
9	Nick put the playing cards into the machine to nesh before starting the game.
10	Rex put the tennis balls into the machine to jeabb as he played fetch.
11	James put the picture of the girl into the machine to chob her name.
12	Jane put her cold and sore feet into the machine to phirf them warm.
13	Matt put his feet into the machine to shug quickly up the wall.
14	Sam saw a black bird and then made the machine ghakk to hear it sing.
15	Ben put the machine into the fish tank to thub the glass clean again.
16	Pip waited for the brushes on the machine to mirbe the sand from his body.
Set 2	
1	Rick put his dirty socks into the machine to tem them clean.
2	Diana put the best orange on the machine to yune the juice.
3	Pam put the dirty flowers under the machine to nid them shiny.
4	Max put his food into the machine to kyve the yucky peas.
5	Sara put her soaking wet hat on the machine to jit it dry.
6	Lucy loaded all the rubbish into the machine to birv it for recycling.
7	Lucas put his sore tummy beside the machine to yab it better again.
8	Jennifer put all her soggy chips under the machine to jayf them crispy.
9	Nick put the playing cards into the machine to vish before starting the game.
10	Rex put the tennis balls into the machine to meaph as he played fetch.
11	James put the picture of the girl into the machine to shep her name.
12	Jane put her cold and sore feet into the machine to ghuzz them warm.
13	Matt put his feet into the machine to thog quickly up the wall.
14	Sam saw a black bird and then made the machine phegg to hear it sing.
15	Ben put the machine into the fish tank to chig the glass clean again.
16	Pip waited for the brushes on the machine to vaype the sand from his body.