Scientific Studies of Reading

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hssr20

Orthographic Expectancies in the Absence of Contextual Support

Signy Wegener, Hua-Chen Wang, Elisabeth Beyersmann, Kate Nation, Danielle Colenbrander & Anne Castles

SCIENTIFIC STUDIES OF READING
https://doi.org/10.1080/10888438.2022.2127356

Routledge
Taylor & Francis Group

Check for updates

Orthographic Expectancies in the Absence of Contextual Support

Signy Wegener[a,b], Hua-Chen Wang[b,c], Elisabeth Beyersmann[a,b], Kate Nation[d], Danielle Colenbrander[a,b], and Anne Castles[a,b]

[a]School of Psychological Sciences, Macquarie University, Sydney, Australia; [b]Macquarie University Centre for Reading, Macquarie University, Sydney, Australia; [c]School of Education, Macquarie University, Sydney, Australia; [d]Department of Experimental Psychology, University of Oxford, Oxford, UK

ABSTRACT

Purpose: Readers can draw on their knowledge of sound-to-letter mappings to form expectations about the spellings of known spoken words prior to seeing them in written sentences. The current study asked whether such orthographic expectancies are observed in the absence of contextual support at the point of reading.

Method: Seventy-eight adults received oral vocabulary training on 16 novel words over two days, while another set of 16 items was untrained. Following training, participants saw both trained and untrained novel words in print for the first time within a lexical recognition task. Half of the items had spellings that were predictable from their pronunciations (e.g., *nesh*), while the remainder had spellings that were less predictable from their pronunciations (e.g., *koyb*).

Results: Participants were able to recognize newly learned words, and lexical recognition latencies displayed clear evidence of orthographic expectancies, as evidenced by a larger effect of spelling predictability for orally trained than untrained items.

Conclusion: These data are consistent with the emergence of orthographic expectancies even when written words are first encountered in isolation.

Oral vocabulary, comprising both the pronunciation and meaning of words, plays a role in reading (Taylor et al., 2015). Training studies (Duff & Hulme, 2012; Hogaboam & Perfetti, 1978; McKague et al., 2001; Taylor et al., 2011) suggest that the relationship is likely causal: knowing a spoken word places the reader at an advantage when that word is first seen in print, relative to words that are not familiar orally. Recent work with children (Wegener et al., 2018, 2020) and adults (Beyersmann et al., 2021) has identified one mechanism that might underpin this association: readers can draw on their knowledge of how spoken sounds map onto written letters to generate *orthographic skeletons* or expectations about the spellings of known spoken words prior to seeing them in print for the first time. Using a training study design in combination with a lexical recognition task at test, we build on this work to ask whether orthographic expectancies are generated in the absence of support from the sentence context at the point of reading.

The *orthographic skeleton hypothesis* provides an account of how spoken word knowledge might benefit children's written word learning (Wegener et al., 2018). It proposes that oral vocabulary, combined with knowledge of phoneme-to-grapheme correspondences, could allow the reader to develop an expectation of the likely spelling of a word prior to seeing it in print for the first time. Two experiments with children (Wegener et al., 2018, 2020) tested this idea directly. Children were first taught the pronunciations and meanings of a group of spoken words. Later, they read both trained and untrained novel words embedded in contextual sentences, while their eye movements were

CONTACT Signy Wegener signy.wegener@mq.edu.au Department of Cognitive Science and Macquarie University Centre for Reading, Macquarie University, New South Wales, Sydney 2109, Australia

recorded. When orally trained novel words were later shown with predictable spellings (e.g., the spoken word /neʃ/ was written as *nesh*), processing times were shorter than for orally trained words with unpredictable spellings (e.g., the spoken word /kɔɪb/ was written as *koyb*). Critically, the difference in fixation durations between items with predictable and unpredictable spellings was larger for orally trained novel words than orally untrained items. These findings were interpreted as evidence for a mechanism causally implicated in word reading that begins prior to visual exposure. Subsequent work with adults has supported these findings (Beyersmann et al., 2021; Jevtović et al., 2022), thereby suggesting that orthographic expectancies may operate across the spectrum of reading skill.

Previous investigations of the orthographic skeleton effect have embedded orally trained and untrained novel words within meaningful sentences at the point of first reading them. Sentential context may well play a role in the generation of orthographic expectancies. For instance, a considerable body of work has shown that the time required to process a written word is influenced by the context that precedes it. Typically, the more predictable a word is given the surrounding context, the more rapidly it is processed (Ehrlich & Rayner, 1981; Fischler & Bloom, 1979; Rayner & Well, 1996; Stanovich & West, 1979, 1983; for a review, see Staub, 2015). This effect holds across methods and indicates that people make predictions about upcoming words as they read sentences (Frisson et al., 2005; Rayner et al., 2004). Extending this reasoning to orthographic expectancies, if readers draw on context in this way, then the orthographic skeleton of a known spoken word might be brought to mind as a constraining sentence fragment is read, potentially pre-activating it. What happens when sentential context is entirely absent is not known. If orthographic skeletons are pre-activated as meaningful sentences are read, then it is possible that orthographic expectancies may either be absent or too weak to be observed without contextual support.

A closely related theoretical issue concerns the timing of the establishment of orthographic expectancies. Orthographic expectancies might be formed "on the fly" during reading, when supportive sentence contexts lead the reader to expect to see a particular word. If orthographic expectancies are entirely absent when words are presented in isolation, this would imply that spelling expectations could be formed during the reading process itself. An alternative possibility is that orthographic expectancies can be accessed at the lexical level, independently of sentence context. On this view, orthographic expectancies are likely formed when a novel spoken word is acquired – a word is heard and, via interactions between spoken and written language codes, a nascent orthographic form is generated. If this is the case, orthographic expectancies should be observed even when written words are presented in isolation. Prior work has demonstrated interactivity between spoken and written word processing at the level of single words (Bakker et al., 2014; Pattamadilok et al., 2009, 2007; Taft, 2011; Taft et al., 2008; Ventura et al., 2008, 2007; Ziegler & Ferrand, 1998). It should therefore be possible to observe evidence of orthographic expectancies when orally known but visually novel written words are presented in isolation. Determining whether orthographic expectancies are observed in an isolated word recognition task will provide important theoretical clarifications regarding their nature and the timing of their emergence.

The current experiment investigated whether skilled readers generate orthographic expectancies for orally known words presented in isolation, and therefore in the absence of contextual support at the point of the first print exposure. Adults were taught the pronunciations and meanings of one set of novel spoken words over two sessions on consecutive days, while another set of items was untrained. Immediately following oral training on the second day, participants' learning of the novel spoken words was checked using a picture naming task. Finally, participants completed a lexical recognition task in which they saw the trained and untrained novel words in print for the first time. In line with previous work investigating the orthographic skeleton, the spelling predictability of the novel words was manipulated such that half of the items in each set had spellings that were highly predictable from their pronunciations (e.g., *nesh*) and therefore likely consistent with participants' expectations, whereas the other items had spellings that were unpredictable from their pronunciations (e.g., *koyb*) and therefore likely inconsistent with participants' expectations. It was expected that oral training and predictable spellings would confer an advantage on lexical recognition accuracy and latency. Further, and critically, if skilled readers generate orthographic expectancies strong enough to be evident in isolated word recognition, this would be supported by the presence of a larger effect of spelling predictability for orally trained compared to untrained items in the lexical recognition task. These hypotheses, together with the experimental design and analysis plan, were preregistered (https://aspredicted.org/pc47v.pdf).

Method

The design was two (trained vs. untrained) by two (predictable vs. unpredictable spelling) within-participants. Dependent variables were lexical recognition latency and accuracy. The experiment was conducted over two days using the Gorilla online platform (Anwyl-Irvine et al., 2020).

Participants

Complete datasets were obtained from 108 undergraduates in an urban university setting, who participated for course credit. This study was approved by the Macquarie University Human Research Ethics Committee and adhered to the standards of the Australian Code for the Responsible Conduct of Research (2007). All participants gave their informed consent to participate in this study and were informed of their right to withdraw at any time. No information about ethnicity was collected. Exclusion criteria were preregistered as follows: the participant (1) provided fewer than 80% of the required novel word productions during spoken word learning; (2) reported having written down the novel spoken words during training; (3) indicated that their data should not be used; (4) demonstrated very poor learning of the novel spoken words (defined as fewer than four of sixteen correct on the learning check). An additional exclusion criterion was not preregistered: participants were excluded if their accuracy on the final lexical recognition task was poor (defined as ≤50% accuracy in one or more conditions). Following these exclusions, data from 78 participants remained (61 females, 16 males, 1 other; mean age 23.5 years, SD = 8.25 years) resulting in a sample size that was slightly smaller than pre-registered (N = 80). The proportion of participants excluded from the current experiment on the basis of the preregistered exclusion criteria, while high compared to some lab-based studies, is comparable to that reported in other online memory experiments (e.g., James et al., 2020), suggesting that the exclusion criteria were successful in identifying participants who did not complete the task as instructed.

Materials

There were two sets of 16 three-phoneme monosyllabic nonwords matched for consonant–vowel structure, taken from Wegener et al. (2018). All items were regular for reading, that is, they used the most common grapheme-to-phoneme correspondences (type frequency M = 94.23%, SD = 10.78). Within each set, half of the items had spellings that were highly predictable from phonology because they contained frequent phoneme-to-grapheme mappings (e.g., “f” for /f/). The remaining items had spellings that were unpredictable from phonology because they contained less frequent phoneme-to-grapheme mappings (e.g., “ph” for /f/). Full item sets appear in Appendix A.

Procedure

Oral vocabulary training

Using a training procedure adapted from Wang et al. (2011), adults were randomly allocated to learn one set of 16 novel spoken words over two days, while the other item set was untrained. Thirty-one participants learned Set 1 and 47 participants learned Set 2. Day 1 consisted of oral vocabulary training and Day 2 began with a top-up of this training. The oral vocabulary training components and experimental procedure are presented in Figure 1.

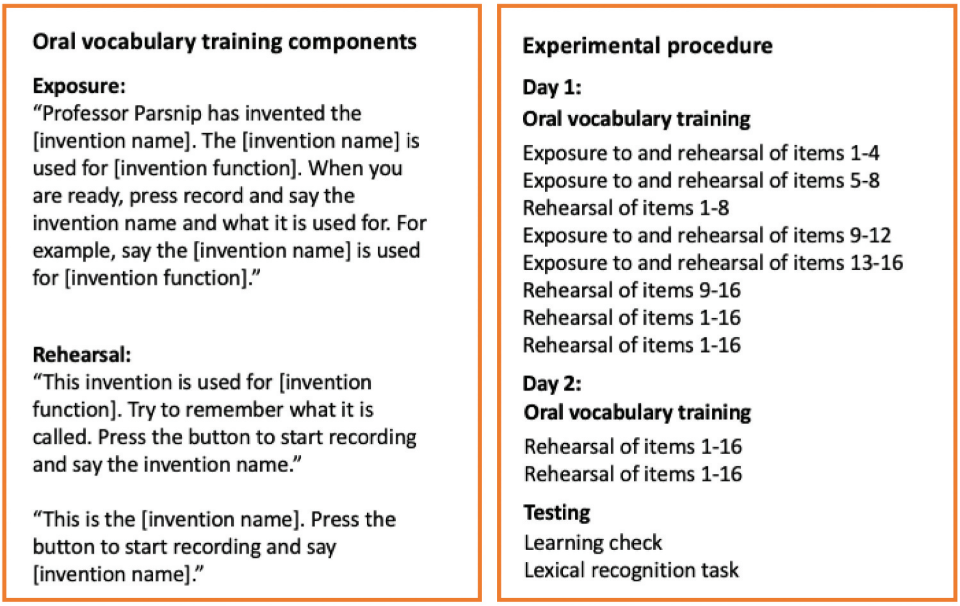

Figure 1. Oral vocabulary training components and experimental procedure.

Learning check

Following oral vocabulary training on Day 2, participants were shown pictures of the inventions one at a time and asked to provide both the name of the invention and its function. Spoken responses were recorded. Feedback was provided regardless of accuracy.

Lexical recognition task: Participants saw the orthography of the orally trained and untrained words for the first time in this task. Trials began with a central fixation cross for 500 ms, which was then replaced with a written word. Participants were instructed to press a button as quickly as possible to indicate whether the word was one they had learned about during oral vocabulary training. The "Y" key was pressed to indicate that the word had been learned orally, and the "N" key was pressed to indicate that it had not. Following three practice trials, participants made decisions to 16 orally trained items and 16 untrained items, with trial order randomized for each participant. Two scores were extracted from this task: recognition accuracy (correct responses coded as 1 and incorrect responses coded as 0) and recognition latency (recorded in milliseconds from the onset of the screen that displayed each written word to the button press response).

Data analysis

Data were analyzed in R (R Core Team, 2021). Participants' responses on the learning check were summed to give a total score out of 16 orally trained items. Independent samples t-tests compared the performance of participants who learned each item set, and paired samples t-tests were used to compare participants' performance on items with predictable and unpredictable spellings. For latency, only correct responses were included. Latencies greater than three standard deviations above or below each condition mean were removed. A generalized linear mixed effects model with lexical recognition accuracy as the dependent variable was computed, along with a linear mixed effects model with log lexical recognition latency as the dependent variable. Both models were constructed using the *lme4* package (Bates et al., 2020) and *p* values were obtained using the *lmerTest* package (Kuznetsova et al., 2017). The fixed effects of training (trained vs. untrained) and spelling predictability (predictable vs. unpredictable) were deviation coded (−0.5, +0.5). As described in the preregistration, we employed a data-driven approach to model selection. Following Barr et al. (2013), models were computed with the maximal random effects structure but these were over-fitted (Baayen et al., 2008). Next, the random intercept model was computed and random slopes were added incrementally. As per Matuschek et al. (2017), a random slope was retained if a likelihood ratio test showed that its addition improved model fit ($p < .2$). The random effects structure for each reported model appears in the Appendix B. Significant interactions were unpacked using the *phia* package (Rosario-Martinez, 2015).

Results

Learning check

On average, participants correctly recalled 13.45 (*SD* = 2.52) of the 16 orally trained invention names. The was no difference in recall between participants who learned Set 1 (*M* = 13.74, *SD* = 2.76) and Set 2 (*M* = 13.26, *SD* = 2.36), $t(76) = 0.83$, $p = .408$. The difference in recall for items with predictable spellings (*M* = 6.79, *SD* = 1.52) and unpredictable spellings (*M* = 6.65, *SD* = 1.28) was not significant, $t(77) = 1.00$, $p = .320$.

Lexical recognition task

Descriptive statistics are represented in Figure 2, while means and variance statistics are reported in Table 1. Model outputs for accuracy and latency are presented in Table 2.

There was no effect of training or predictability for *lexical recognition accuracy*. There was, however, a significant interaction between training and predictability. Interaction contrasts revealed

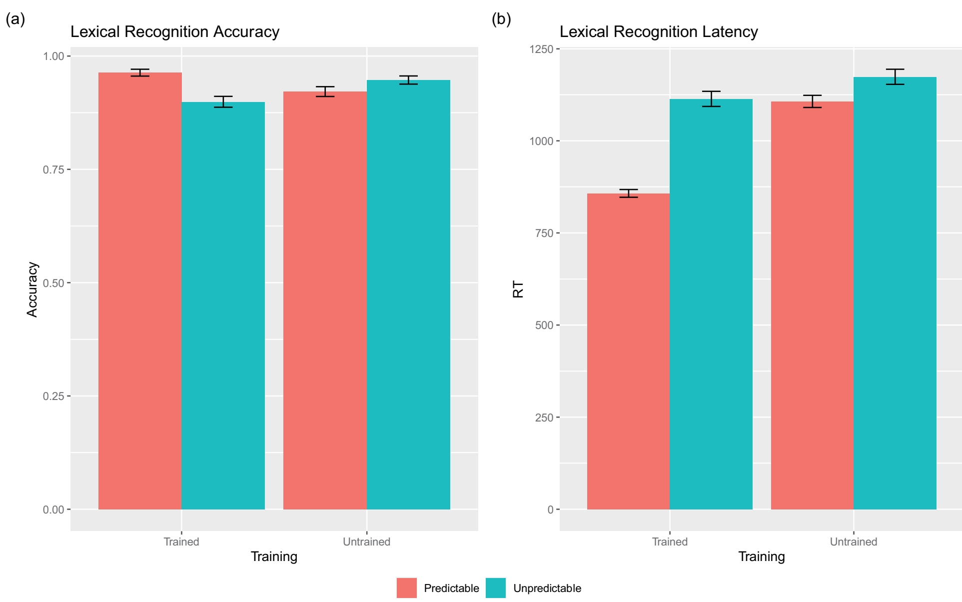

Figure 2. Means and standard errors of lexical recognition accuracy and latency. Accuracy reflects proportion correct, latency is expressed in milliseconds.

Table 1. Means and standard errors of lexical recognition accuracy and latency by condition.

	Lexical recognition accuracy	Lexical recognition latency
Trained predictable	0.96 (0.01)	857 (10.61)
Trained unpredictable	0.90 (0.01)	1114 (20.41)
Untrained predictable	0.92 (0.01)	1107 (16.48)
Untrained unpredictable	0.95 (0.01)	1174 (20.50)

Table 2. Output of the models for lexical recognition accuracy and latency.

		Training	Predictability	Interaction
Lexical recognition accuracy				
	b	0.40	−0.32	1.64
	SE	0.35	0.32	0.65
	z	1.14	−1.02	2.54
	p	.26	.31	.011
Lexical recognition latency				
	b	0.15	0.14	−0.18
	SE	0.02	0.03	0.04
	t	6.55	4.47	−4.35
	p	< .001	< .001	< .001

a significant effect of spelling predictability for trained items (χ^2 = 0.36; p = .001) with a higher probability of errors on unpredictable items. There was no effect of predictability for untrained items (χ^2 = 0.84; p = .359). Further, there was no effect of training for items with predictable spellings (χ^2 = 0.45; p = .390), but untrained items with unpredictable spellings were associated with a higher probability of being recognized than trained items with unpredictable spellings (χ^2 = 0.65; p = .020).

Turning to *lexical recognition latency*, trained items were recognized more quickly than untrained items. Items with predictable spellings were recognized more quickly than those with unpredictable spellings. Critically, training and spelling predictability interacted. Interaction contrasts showed that there was a significant effect of spelling predictability for trained items (χ^2 = 0.12; p <.001) but not for untrained items (χ^2 = 0.03; p = .178). Further, items with predictable spellings benefited from training (χ^2 = 0.12; p < .001) but items with unpredictable spellings did not (χ^2 = 0.03; p = .056).[1]

Discussion

Previous work on orthographic expectancies has always tested reading on the first visual exposure with recently learned words embedded in meaningful sentences. This experiment asked whether orthographic expectancies are evident in the absence of such contextual support. The results for lexical recognition latency clearly showed the predicted pattern of results: trained items and those with predictable spellings were recognized more rapidly than untrained items and those with unpredictable spellings. These two factors interacted, showing that there was a larger effect of spelling predictability for orally trained compared to untrained items. This is consistent with prior findings (Beyersmann et al., 2021; Wegener et al., 2018, 2020), and supports the idea that readers generate expectations of the spellings of orally known words prior to seeing them in print for the first time. The specific pattern of results suggests that when orthographic expectancy and form are congruent, processing is facilitated. In contrast, incongruence between expectancy and form is surprising to participants, and reduces the benefit of oral familiarity on processing times.

These findings extend prior work by showing that the characteristic *orthographic skeleton* effect is also evident when novel written words are seen for the first time in isolation. As such, contextual support at the point of reading is not a necessary precondition for the emergence of orthographic expectations. This is consistent with prior work suggesting that orthography is automatically activated during spoken word processing (Pattamadilok et al., 2009, 2007; Taft, 2011; Taft et al., 2008; Ventura et al., 2008, 2007; Ziegler & Ferrand, 1998). Although our findings show that contextual support at the point of reading is not required to observe orthographic expectancies at the first visual exposure, it cannot rule out the possibility that context might still exert an influence on the generation of orthographic expectancies. For instance, it might be that context, perhaps via pre-activation of the orthographic expectancy, may give rise to stronger orthographic expectancies than those arising when the context surrounding a target word is uninformative. To address this question, future work should directly compare the magnitude of the orthographic expectancy effect when participants read trained and untrained target words within contextually supportive and neutral sentences at test.

Contrary to predictions, neither oral training nor spelling predictability influenced accuracy on the lexical recognition task. The absence of these effects is perhaps not surprising: participants discriminated trained and untrained orthographic forms with ease and overall, accuracy was high (range 90–96% correct). An interaction emerged due to a small recognition advantage for words with an unpredictable spelling in the untrained condition compared to words with unpredictable spellings that had been learned in oral form (≈ 5%). One interpretation of this finding is that unpredictable spellings are easier to reject as unknown when participants have no expectations of their spellings than to accept words as known when participants' spelling expectations are violated. This implies that, even though all items were regular for reading, unexpected spellings might induce uncertainty regarding familiarity for a small proportion of items.

We employed a single outcome measure because the key test of whether orthographic expectancies can be observed following oral vocabulary training is the first visual exposure. Future work might consider using other outcome measures at the first and subsequent exposures to investigate the emergence of expectancies, and how they influence written word learning over the longer term, especially the acquisition of new spellings. This is important in view of prior work demonstrating the facilitatory role of spelling when learning the meaning of new words (Ricketts et al., 2009; Rosenthal & Ehri, 2008). Another question to ask is whether orthographic expectancies can be observed among poor readers whose processing of phonology-orthography links is less strong.

An anonymous reviewer raised an alternative interpretation of the latency findings, proposing that because the test phase immediately followed the second oral vocabulary training session, the novel words remained active within working memory during the test phase, thus possibly driving the observed effect. If the short delay between training and testing accounted for the findings of the current study, then increasing the delay between these two phases should remove the effect. However, prior studies have shown that this is not the case; the effect remains present even when there is a delay between training and testing (range 1 to 7 days: Wegener et al., 2018, 2020), suggesting that this alternative explanation does not account for the current findings.

We have concentrated here on orthographic expectancies as one potential mechanism via which oral vocabulary knowledge may support written word processing. This is not to suggest that orthographic expectancies are the only plausible cognitive mechanism; other potential causal mechanisms have been proposed in the literature. For instance, according to the lexical restructuring model (Walley & Metsala, 2003), vocabulary knowledge exerts an indirect effect on reading acquisition. As one's vocabulary grows, phonological representations must become more fine-grained in order that each remains distinct. This in turn promotes the development of phonological representations that are segmental rather than holistic, thereby supporting emerging phonemic awareness skills that promote reading acquisition. However, the results of several training studies (Duff & Hulme, 2012; Hogaboam & Perfetti, 1978; McKague et al., 2001), including this one, suggest that oral vocabulary likely influences orthographic processing directly and at the item level, such that prior knowledge of a given spoken word supports processing of that specific written word. An alternative cognitive mechanism, termed set for variability (Venezky, 1999), views the relationship between vocabulary and written word processing as being direct, but situates the influence of vocabulary at a later time point. Specifically, this mechanism is thought to operate from the point of visual exposure: when a decoded pronunciation does not match any existing entries in the spoken lexicon, this prompts the reader to adjust their pronunciation until they find one that matches a known spoken word and makes sense within the context. This mechanism is particularly expected to influence irregular word processing, where the need for adjustments to decoded pronunciations is salient. The current experiment employed only regular written words, with the manipulation being on the predictability of the spelling, limiting by design the requirement to consciously adjust mispronunciations and removing entirely any influence of context. Although orthographic expectancies and set for variability are distinct causal accounts, they need not be mutually exclusive (Wegener et al., 2022).

In summary, our findings suggest that learning new words in the oral domain leads skilled readers to form orthographic expectancies that influence reading behavior on the first print exposure. These expectancies were indexed by a speed of processing advantage in lexical recognition when a word's spelling aligned with expectation, in combination with disruption of processing when orally known words were spelled unexpectedly. Orthographic expectancies occurred when words were presented in isolation, demonstrating that they did not rely on contextual pre-activation during sentence reading. Instead, our findings provide further support for the orthographic skeleton hypothesis (Wegener et al., 2018) and indicate that learning a new word in oral vocabulary establishes a nascent orthographic form such that orthographic expectancies are later observed at the point of reading, even when processing is assessed via a single word recognition task and in the absence of contextual cues.

Note

1. Please see the supplementary material for an additional exploratory analysis.

Acknowledgments

The authors thank Erik D. Reichle for his comments on an earlier version of this work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by a grant from the Australian Research Council [grant number DP200100311] to A. C., K. N. & E. B.

ORCID

Signy Wegener http://orcid.org/0000-0001-5364-0202
Hua-Chen Wang http://orcid.org/0000-0002-9845-147X
Elisabeth Beyersmann http://orcid.org/0000-0001-9653-6106
Danielle Colenbrander http://orcid.org/0000-0001-5577-7501
Anne Castles http://orcid.org/0000-0001-8228-8260

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

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*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language*, *73*, 116–130. https://doi.org/10.1016/j.jml.2014.03.002

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1),* 1-48 doi:10.18637/jss.v067.i01

Beyersmann, E., Wegener, S., Nation, K., Prokupzcuk, A., Wang, H.-C., & Castles, A. (2021). Learning morphologically complex spoken words: Orthographic expectations of embedded stems are formed prior to print exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(1), 87–98. doi:10.1037/xlm0000808.

Duff, F. J., & Hulme, C. (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading*, *16*(6), 504–525. https://doi.org/10.1080/10888438.2011.598199

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 641–655. https://doi.org/10.1016/S0022-5371(81)90220-6

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior*, *18*(1), 1–20. https://doi.org/10.1016/S0022-5371(79)90534-6

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 862–877.

Hogaboam, T. W., & Perfetti, C. A. (1978). Reading skill and the role of verbal experience in decoding. *Journal of Educational Psychology*, *70*(5), 717–729. https://doi.org/10.1037/0022-0663.70.5.717

James, E., Ong, G., Henderson, L., & Horner, A. J. (2020). *Make or break it: Boundary conditions for integrating multiple elements in episodic memory* Preprint. PsyArXiv. https://doi.org/10.31234/osf.io/pd9us

Jevtović, M., Antzaka, A., & Martin, C. D. (2022). Gepo with a G, or Jepo with a J? Skilled readers generate orthographic expectations for novel spoken words even when spelling is uncertain. *Cognitive Science*, *46*(3), 1-26. https://doi.org/10.1111/cogs.13118

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. doi:10.18637/jss.v082.i13.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McKague, M., Pratt, C., & Johnston, M. B. (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition*, *80*(3), 231–262. https://doi.org/10.1016/S0010-0277(00)00150-5

Pattamadilok, C., Morais, J., De Vylder, O., Ventura, P., & Kolinsky, R. (2009). The orthographic consistency effect in the recognition of French spoken words: An early developmental shift from sublexical to lexical orthographic activation. *Applied Psycholinguistics*, *30*(3), 441–462. https://doi.org/10.1017/S0142716409090225

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes*, *22*(5), 700–726. https://doi.org/10.1080/01690960601049628

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 720–732. doi:10.1037/0096-1523.30.4.720.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*(4), 504–509. https://doi.org/10.3758/BF03214555

R Core Team. (2021). *R: The R project for statistical computing*. https://www.r-project.org/

Ricketts, J., Bishop, D. V. M., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Quarterly Journal of Experimental Psychology*, *62*(10), 1948–1966. https://CRAN.R-project.org/package=phia

Rosario-Martinez, H. (2015). Phia: Post-Hoc Interaction Analysis. R package version 0.2-1. CRAN. https://cran.r-project.org/web/packages/phia/phia.pdf

Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology*, *100*(1), 175–191. doi:10.1037/0022-0663.100.1.175.

Stanovich, K. E., & West, R. F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition*, *7*(2), 77–85. https://doi.org/10.3758/BF03197588

Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, *112*(1), 1–36. doi:10.1037/0096-3445.112.1.1.

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass*, *9*(8), 311–327. https://doi.org/10.1111/lnc3.12151

Taft, M. (2011). Orthographic influences when processing spoken pseudowords: Theoretical implications. *Frontiers in Psychology*, *2*, 1–7. https://doi.org/10.3389/fpsyg.2011.00140

Taft, M., Castles, A., Davis, C., Lazendic, G., & Nguyen-Hoan, M. (2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. *Journal of Memory and Language*, *58*(2), 366–379. https://doi.org/10.1016/j.jml.2007.11.002

Taylor, J. S. H., Duff, F. J., Woollams, A. M., Monaghan, P., & Ricketts, J. (2015). How word meaning influences word reading. *Current Directions in Psychological Science*, *24*(4), 322–328. https://doi.org/10.1177/0963721415574980

Taylor, J. S. H., Plunkett, K., & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: An artificial orthography paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 60–76. doi:10.1037/a0020126.

Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography*. Guilford Press.

Ventura, P., Kolinsky, R., Pattamadilok, C., & Morais, J. (2008). The developmental turnpoint of orthographic consistency effects in speech recognition. *Journal of Experimental Child Psychology*, *100*(2), 135–145. https://doi.org/10.1016/j.jecp.2008.01.003

Ventura, P., Morais, J., & Kolinsky, R. (2007). The development of the orthographic consistency effect in speech recognition: From sublexical to lexical involvement. *Cognition*, *105*(3), 547–576. https://doi.org/10.1016/j.cognition.2006.12.005

Walley, A. C., & Metsala, J. L. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing*, *16*(1), 5–20. doi:10.1023/A:1021789804977.

Wang, H.-C., Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology*, *109*(1), 39–57. https://doi.org/10.1016/j.jecp.2010.11.005

Wegener, S., Beyersmann, E., Wang, H.-C., & Castles, A. (2022). Oral vocabulary knowledge and learning to read new words: A theoretical review. *Australian Journal of Learning Difficulties*, 1–26. https://doi.org/10.1080/19404158.2022.2097717

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), e12577. https://doi.org/10.1111/desc.12577

Wegener, S., Wang, H.-C., Nation, K., & Castles, A. (2020). Tracking the evolution of orthographic expectancies over building visual experience. *Journal of Experimental Child Psychology*, *199*, 104912. https://doi.org/10.1016/j.jecp.2020.104912

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, *5*(4), 683–689. https://doi.org/10.3758/BF03208845

Appendix A

Experimental items

	Set 1		Set 2	
	Phonology	Orthography	Phonology	Orthography
Predictable	/d ev/	jev	/tem/	tem
	/jæg/	yag	/nld/	nid
	/vlb/	vib	/d lt/	jit
	/tʌp/	tup	/jæb/	yab
	/neʃ/	nesh	/vlʃ/	vish
	/t b/	chob	/ʃep/	shep
	/ʃʌg/	shug	/θ g/	thog
	/θʌb/	thub	/t lg/	chig
Unpredictable	/viːm/	veme	/juːn/	yune
	/baɪp/	bype	/kaɪv/	kyve
	/j ːp/	yirp	/b ːv/	birv
	/kɔɪb/	koyb	/d aɪf/	jayf
	/d iːb/	jeabb	/miːf/	meaph
	/f ːf/	phirf	/gʌz/	ghuzz
	/gæk/	ghakk	/feg/	phegg
	/m ːb/	mirbe	/veɪp/	vaype

Appendix B

Random effects structure of reported models

Model	Random effects structure
Accuracy	correct ~ training*predict + (1 | Participant) + (1 + training| item)
Latency	log(RT) ~ training*predict + (1 + training + predict | Participant) + (1 + training + predict | item)