Applied Psycholinguistics (2026), **47**, e61, 1–23
doi:10.1017/S0142716426100800

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

How many syllables do you hear in [skis]: One or two? It depends on whether you have learned to read or not

Samantha Ruvoletto[1], Teng Guo[2], Katarina Bartkova[1], Marie Flesch[1] and Daniel Zagar[1]

[1]University of Lorraine, France and [2]Royal Holloway University of London, UK
Corresponding author: Teng Guo; Emails: teng.guo@unibo.it; guoteng931013@gmail.com

(Received 7 February 2025 UTC; revised 19 March 2026 UTC; accepted 1 June 2026 UTC)

Abstract

This study investigates how statistical orthographic knowledge—how frequently a phonological unit is written in a language—influences phonological awareness tasks. In Experiment 1, 144 children with varying reading levels and 61 undergraduate French-speaking students completed a syllable-counting task for consonant-vowel-consonant (CVC) and consonant-consonant-vowel-consonant (CCVC) acoustic monosyllabic pseudowords. The number of syllables of the pseudowords was ambiguous (one or two) since some of their ending rhymes could be frequently spelled with an "e" (e.g., [-uʃ]), whereas others are frequently spelled without an "e" (e.g., [-is]). The results showed that (1) pseudowords tended to be more frequently counted as one-syllable sequences when they had a high probability of not be spelled with an "e" rather than when they had a low probability of not be spelled with an "e, " and that (2) this effect was more pronounced in literate adults as compared to children with lower reading levels. Experiment 2, involving prereaders, confirmed that children without formal literacy instruction showed little influence of orthographic information on syllable counts. These findings suggest that phonological responses are shaped by statistical relationships between phonological and orthographic representations, which emerge gradually through learning to read.

Keywords: literacy acquisition; orthography; phonology; phonological awareness task

Introduction

How many syllables do you hear in [skis]? If the judgment relies solely on phonological information, the answer is one syllable. However, knowledge of spelling can lead to a different response. In French, many words with this rime are spelled *-isse*, that is, with a final "e," which corresponds to a bisyllabic structure (e.g., /i.sə/). This example illustrates how orthographic knowledge may influence syllable perception. More generally, in the present study, we investigate this interaction between orthography and phonology during literacy acquisition.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Neuroimaging studies have demonstrated a functional and anatomical link between phonemic and graphemic representations, which is strengthened by literacy acquisition (Dehaene et al., 2015; Dehaene et al., 2010). Converging behavioral evidence indicates that orthographic information can influence a range of tasks that require an analysis of acoustic forms. These tasks range from rime judgment (Donnenwerth-Nolan et al., 1981; Seidenberg & Tanenhaus, 1979) and phoneme monitoring (Hallé et al., 2000; Frauenfelder et al., 1990; Perre et al., 2009; Ziegler et al., 2004), to syllable monitoring (Taft & Hambly, 1985), click detection (Bertelson, 1972), and spoken word recognition (Jakimik et al., 1985; Taft et al., 2008). Importantly, these tasks differ in the extent to which they require explicit reflection on phonological structure, ranging from online speech-processing paradigms to more explicit metalinguistic tasks.

Beyond these tasks, the influence of orthographic representations has also been examined in explicit phonological awareness tasks (Liberman et al., 1974; Melby-Lervåg et al., 2012). Phonological awareness refers to the intentional ability to analyze and manipulate units of speech (words, onsets-rimes, syllables, and phonemes). Many studies have established that children's success in reading and spelling correlates positively with their ability to recognize and manipulate phonemes—a skill often termed phonemic awareness (Adams, 1998; Brady & Shankweiler, 1991; Bradley & Bryant, 1983; Goswami & Bryant, 2016). Importantly, these findings suggest that phonological awareness does not develop independently from orthographic knowledge, but it is progressively shaped by literacy acquisition. As children's orthographic knowledge grows, their phonological representations become more fine-grained, suggesting that literacy acquisition exerts a bidirectional influence on the cognitive mechanisms underlying speech perception and production (Brady & Shankweiler, 1991).

Although the studies discussed above establish a close relationship between phonological awareness and literacy acquisition, they do not directly address how orthographic knowledge influences performance in phonological awareness tasks. The following studies provide more direct evidence of this influence.

To our knowledge, Ehri and Wilce (1980) were the first to provide evidence that orthography actively shapes phoneme identification in phonemic awareness tasks. In their Experiment 1, they measured the phonemic counting and spelling skills of 24 native English-speaking fourth graders. Participants had to blend phonemically and then spell a list of word pairs (e.g., "rich" [ɹɪč] and "pitch"[pɪč][1]). For each pair, one of the words included an extra letter in the spelling corresponding to a potential extra sound (e.g., both "rich" [ɹɪč] and "pitch"[pɪč] had three phonemes, but the "t" of "pitch" was the extra letter associated to a potential extra sound; therefore, subjects could count [pɪtč]). They found that the extra letter (the "t" of "pitch") was often counted as an extra sound (e.g., four phonemes for "pitch" [pɪč] vs. three phonemes for the word "rich" [ɹɪč]). Moreover, when this extra sound was not identified, it was also not correctly spelled (e.g., "pich" instead of "pitch"). To explain their results, they proposed that orthography "does leave its mark on the reader's conceptualization of the sound structures of words" (Ehri & Wilce, 1980, p. 380). Literates had an expert knowledge of word forms and spelling, which shaped the organization and the nature of phonological representations, changing the counting of phonemes.

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Subsequent studies have further explored the mechanisms underlying this influence. Stuart (1990) investigated the process and the factors underlying the use of orthographic representation in phonemic awareness tasks. She tested 9-year-old good and poor spellers with a phoneme deletion task consisting of words and nonwords. Results showed that good spellers performed better than poor spellers on both types of items. However good spellers produced more orthographical responses (removing the letter "l" from a word like "cold" [koʊld] and changing the vowel into /ɒ/, resulting in "cod" [kɒd]) than phonological responses (deleting the sound /l/ and preserving the vowels /oʊ/ producing "code" [koʊd]) with word stimuli. To account for these results, she suggested that in general all children employed a phonological strategy to delete phonemes, but only good spellers used an additional orthographic strategy to perform the task correctly, especially with words.

Importantly, later work questioned whether this influence reflects strategic or automatic processes. Castles et al. (2003) confirmed the influence of orthographic information on phonological awareness tasks and also suggested that the underlying process is not strategic but automatic. They proposed a deletion task which compared orthographically transparent and opaque words (e.g., deleting the /d/ from "dentist" [dent̪ɪst] vs. deleting the /s/ from "fox" [fɑ:ks]), presented either in a block that mixed them randomly or in two pure blocks. Results showed that participants performed better on transparent items than on opaque items regardless of the manner of presentation (mixed or pure blocks). The authors considered these results as evidence that literates cannot voluntarily turn off orthographic information to perform phonological tasks even when it is disadvantageous to them.

Tyler and Burnham (2006) provided robust evidence that the influence of orthography on phonological tasks is not under conscious control. They recorded reaction times in a phoneme deletion task testing alphabetically literate individuals specifically instructed not to use orthography for responding. Data highlighted an automatic use of orthography, since participants performed the task more slowly with orthographically mismatched stimulus-response pairs (e.g., "worth"- "earth") than with matched pairs (e.g., "wage"- "age"), despite the instructions provided.

Collectively, these studies underscore the robust interaction between orthography and phonology (see Morais, 2021, for a comprehensive review), demonstrating that when performing phonological awareness tasks, literates inevitably engage orthographic representations. In addition, some studies have employed nonlexical stimuli (e.g., pseudowords) to investigate orthographic influences while reducing reliance on stored lexical representations. Nonlexical items have served different methodological purposes, including controlling for lexical influences while preserving phonological structure (e.g., Dijkstra et al., 1995), manipulating orthographic similarity independently of lexical representations (e.g., Taft et al., 2008), or testing whether orthographic effects generalize beyond real words, typically in terms of grapheme–phoneme consistency (e.g., Ziegler & Ferrand, 1998). Across these paradigms, orthographic effects are typically interpreted in terms of grapheme–phoneme correspondences or item-specific representations, rather than broader properties of the writing system.

However, as individuals acquire reading and spelling skills in a writing system, they become sensitive to the broader statistical properties of their orthography,

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

reflecting distributional patterns across the lexicon rather than properties of individual words. In the current study, statistical orthographic knowledge refers to the frequency with which a phonological unit is written in a given language. For example, English speakers may discover that "bn" cannot start a word while "sn" can. This requires an understanding of how specific letter combinations correspond to sounds.

Studies have provided evidence of the influence of statistical orthographic properties on orthographic tasks such as orthographic judgment, dictation, etc. (e.g., Cassar & Treiman, 1997; Chetail, 2015; Pacton et al., 2001; Samara & Caravolas, 2014; Treiman & Boland, 2017; Treiman & Kessler, 2006). However, it remains unclear whether such statistical orthographic knowledge—independent of specific lexical items—can influence phonological representations. The present study aims to fill this gap by investigating whether statistical orthographic knowledge can influence performance in a phonological awareness task.

Relatedly, Ventura et al. (2001) showed that orthographic knowledge can influence speakers' segmentation of syllables in phonological awareness tasks using both words and pseudowords. However, their interpretation focused on the influence of orthography on postperceptual phonological representations. The present study builds on this work by specifically examining whether statistical orthographic knowledge—defined in terms of distributional spelling patterns—can bias phonological judgments.

To address this question, we designed a task that isolates statistical orthographic knowledge from lexical influences. In the present study, statistical orthographic knowledge reflects the frequency with which a phonological unit is written in a given language.

Our manipulation focuses on a specific property of French orthography: the probability that a phonological rime (vowel-consonant, VC) is spelled with a final "e." Importantly, different rimes vary in this probability across the lexicon. For example, the rime [is] is more often spelled without a final "e" (e.g., "dix" [dis], "six" [sis], "fils" [fis], "jadis" [ʒa.dis], but also "hisse" [is], "lisse" [lis]), whereas the rime [uʃ] is typically spelled with a final "e" (e.g., "bouche" [buʃ], "touche" [tuʃ], "douche" [duʃ], "couche" [kuʃ]). This variation allows us to operationalize statistical orthographic knowledge as a probabilistic property of rime spelling. Importantly, this measure captures graded distributional properties of orthography (i.e., frequency across the lexicon), rather than categorical notions of orthographic consistency, which are often defined in binary terms.

To test whether such knowledge influences phonological processing, we used a syllable-counting task. This task is particularly suitable in French because syllable perception can be ambiguous due to the presence of word-final schwa (Laks, 1995). The schwa (/ə/) is an optional vowel that may or may not be realized (e.g., [ʃə.val] vs. [ʃval] for "cheval"), and can also be epenthetically added after consonant-final words (e.g., [stop] vs. [sto.pə]). Corpus data indicate that such schwa realizations in speech production occur in 13.40% of instances where words ended with a consonantal coda (Hutin et al., 2021), leading to variability in perceived syllable structure (see also Racine & Grosjean, 2005).

Critically, this ambiguity allows orthographic knowledge to bias phonological judgments. In the present design, all stimuli were acoustically monosyllabic

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

consonant-(consonant)-vowel-consonant (C(C)VC) pseudowords (e.g., [skis]). However, rimes with a higher probability of being spelled with a final "e" are more likely to be associated with a bisyllabic representation (C(C)V.Cə, i.e., [ski.sə]), whereas rimes with a lower probability are more likely to be treated as monosyllabic. Thus, the probability of final-"e" spelling serves as the key predictor of participants' responses. This effect of statistical orthographic knowledge was expected to vary across groups with different reading levels, being weaker or absent in beginning readers and stronger in more experienced readers.

Experiment 1 had two main aims. First, it tested whether statistical orthographic knowledge influences performance in a phonological awareness task. Second, it examined how this influence varies across groups differing in reading experience. We predicted that the effect would increase with literacy level, being strongest in literate adults and weaker in less experienced readers. To further characterize the developmental trajectory of this effect, Experiment 2 tested French-speaking prereaders.

Experiment 1

Design

Experiment 1 involved a syllable-counting task using two types of test stimuli: a fricative-coda set and a plosive-coda set. The ambiguity inherent to this task arises from the fact that C(C)VC sequences can be interpreted either as monosyllabic (VC) or disyllabic (V.CV). Importantly, this ambiguity may have two distinct potential sources: (i) acoustic–articulatory factors, related to the possible realization of a final schwa in speech, and (ii) orthographic factors, related to whether a given phonological rime is typically associated with a final "e."

The fricative-coda set constituted the critical condition, designed to test whether statistical orthographic knowledge influences syllable-counting responses and whether this influence varies with reading level. These pseudowords were acoustically monosyllabic C(C)VC sequences ending in fricative consonants (/f/, /s/, /ʃ/). They were selected not only because fricative codas have a very low probability of being followed by a schwa in speech production (6.83%–8.03%, according to Hutin et al., 2021), thereby minimising acoustic ambiguity, but also because they exhibit substantial orthographic variability across rimes. Specifically, some fricative-coda rimes are almost always spelled without a final "e," whereas others are typically associated with a final "e." This variability makes it possible to operationalize statistical orthographic knowledge as a graded property of rime spelling.

The plosive-coda set served as a control condition to assess the potential influence of acoustic factors. These items ended in plosive consonants and had a higher and more variable probability of schwa realization in speech (11.31%–23.91% according to Hutin et al., 2021). This set was included to determine whether participants' responses might reflect sensitivity to schwa-production probability in spoken French. Critically, if participants do not show sensitivity to this phonetic factor in the plosive-coda set, it becomes unlikely that such factors explain the effects observed in the fricative-coda set.

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

The two stimulus types were presented together and mixed within the same experimental blocks.

Ethics

The experiments were approved by the Ethics Committee of University of Lorraine (CER-UL, number 2023/262).

Participants

Participants were 144 children attending primary school (73 first graders: 37 girls, 36 boys, $M_{age} = 7.08$, $SD = 0.25$; 71 third graders: 37 girls, 34 boys, $M_{age} = 10.92$, $SD = 0.25$) and 61 undergraduate students (56 females, 5 males, $M_{age} = 25.41$, $SD = 1.88$). They were native French speakers with no reported hearing problem and normal or corrected vision. Verbal consent for participation in the study was obtained from the adult participants. The child participants' school supervisors and parents or legal guardians provided written informed consent prior to inclusion in the study. Adults read an informed consent form and agreed to participate in the study.

All participants completed a phonological tapping task to assess basic syllable awareness. In this task, participants heard five pseudowords spoken by the experimenter and indicated the number of syllables by tapping. Nine first -grade children who failed more than one item were excluded from the analysis. Six children receiving speech therapy were also excluded.

Children's literacy levels were assessed using the "Alouette" test (Lefavrais, 2005), a standardized French assessment of reading level. In this task, children were instructed to read aloud 265 words as quickly and accurately as possible within 3 minutes. In this test, error rates and speed were used to assign reading levels regardless of the actual school-grade level. According to this assessment, eight reading level groups were generated. From these eight reading level groups, three reading level groups with a sufficient number of subjects were selected: first-grade reading level group with 36 children (all were first graders), second-grade reading level group with 33 children (27 of whom were first graders and 6 of whom were third graders), fifth-grade reading level group with 26 children (all were third graders). Five reading level groups were excluded from the analysis because they had a small number of participants (third grade = 5, fourth grade = 7, sixth grade = 11, seventh grade = 9, and eighth grade = 2, for a total of 34 children). The characteristics of the three retained reading level groups and of adult participants are summarized in Table 1.

Stimuli

The stimuli were designed to test whether phonological judgments are influenced by orthographic knowledge while minimising the influence of lexical orthographic representations tied to specific words. For this reason, pseudowords were used rather than real words. Real words can strongly activate stored orthographic forms and their associated syllabic structure, whereas pseudowords make it possible to isolate subtler effects of statistical orthographic knowledge. In addition, the use of pseudowords allowed us to systematically control and balance the distribution of

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Table 1. Characteristics of reading level groups

Reading level	First grade	Second grade	Fifth grade	Literate adults
Number	36	33	26	61
Males; females	21; 15	15; 18	13; 13	5; 56
Mean chronological age (SD)	7.06 (0.37)	7.61 (1.45)	10.82 (0.54)	25.41 (1.88)

rimes across conditions, which would have been difficult to achieve using real words due to lexical constraints.

All stimuli were recorded by a native speaker (a young female native speaker majoring in speech therapy in Nancy, France). The speaker was instructed to utter the pseudowords in a neutral way, keeping the same speech rate and a neutral intonation pattern. The pronunciation of the pseudowords was carried out in a list-like way. The recorded speech signal was annotated into API symbols (with Transcriber software) and segmented into sounds using an automatic forced speech alignment between the speech signal and the sounds that it contained (with Astali software, https://astali.loria.fr/). Finally, the segmentation of the speech signal into separate auditory files, each containing a single stimulus, was done automatically, using a signal-processing script.

Materials also contained 6 training items (see Appendix A) and 40 fillers (see Appendix B). Fillers were open unambiguous monosyllables and bisyllables. Six French mother-tongue speakers verified that all stimuli in the list had no semantic meaning.

Fricative-coda set

The fricative-coda set included 40 acoustically monosyllabic C(C)VC pseudowords (Appendix C) ending in fricative codas /f/, /s/, and /ʃ/. In this set, the relevant unit was the rime because the orthographic manipulation concerned the probability that a given rime is spelled with a final e in French.

The frequencies of words ending in the same rime were extracted from the French lexical database Lexique3 (New et al., 2004). Ten rimes varying in the probability that their spelling does not end in "e" were selected (Table 2). Each rime was combined with four onsets, either a single consonant (C-) or a two-consonant cluster (CC-), to form the pseudowords.

During the recording of this set, the speaker was instructed not to produce the schwa at the end of the items. We controlled the absence of schwa for each item using PRAAT software (Boersma & Weenink, 2023).

Plosive-coda set

The plosive-coda set included 40 C(C)VC pseudowords (Appendix D) ending in voiced plosive codas /d/, /g/, and /b/ (Table 3). In this set, the relevant unit was the coda because this set was not designed to manipulate orthographic probability at the rime level, but rather to evaluate whether responses might generally be influenced by the probability of schwa realization associated with the final consonant in speech.

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Table 2. Fricative-coda set

Rime	Probability that spelling does not end in "e"	Coda	% Schwa[a]	Example
[us]	0.785	/s/	8.03	[slus]
[ys]	0.455	/s/	8.03	[fʁys]
[is]	0.396	/s/	8.03	[slis]
[ɑ̃s]	0.048	/s/	8.03	[glɑ̃s]
[ɛf]	0.982	/f/	6.83	[blɛf]
[uf]	0.292	/f/	6.83	[kluf]
[iʃ]	0.018	/ʃ/	7.25	[fliʃ]
[uʃ]	0	/ʃ/	7.25	[spuʃ]
[ɑ̃ʃ]	0	/ʃ/	7.25	[stɑ̃ʃ]
[yʃ]	0	/ʃ/	7.25	[slyʃ]

[a]% Schwa represents the proportion of the realization of final-schwa in words ending in the same coda, according to the corpus-based study of the distribution of the word-final schwa in speech production in Standard French (Hutin et al., 2021).

Each coda was combined with consonantal onsets (C- or CC-) and a vowel nucleus to form pseudowords. These items also contained no final schwa in the acoustic signal, or only a residual schwa-like trace when complete deletion was not possible. The signal was checked in PRAAT software (Boersma & Weenink, 2023).

Procedure

The experiment was conducted on a Dell Windows SO computer running the open-source software *OpenSesame* (Mathôt et al., 2012). All participants listened to stimuli through headphones (Sennheiser hd 202 II), on volume level 16 out of 100.

Child and adult participants were tested one by one in a quiet room. Test stimuli from both the fricative- and plosive-coda sets, along with fillers, were presented to the participants in a fixed pseudorandom order in a single session. Participants were instructed to respond spontaneously if the stimulus was composed of one syllable, pressing the "q" key (labeled with a "1"), or of two syllables, pressing "m" key (labeled with a "2") on AZERTY keyboards.

Participants started with a training block composed of six trials, followed by two testing blocks, each one made up of 60 trials. Each block contained 40 test stimuli, with 20 from the fricative-coda set and 20 from the plosive-coda set, along with 20 fillers consisting of 10 monosyllables and 10 bisyllables. The session lasted about 30 min. No feedback was provided. All participants achieved more than 80% correct answers on the filler items, which were nonambiguous.

Results

Fricative-coda set

Table 4 reports the average proportion of one-syllable responses in the fricative-coda set across reading levels, as a function of the probability that the rime is spelled

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Table 3. Plosive-coda set

Coda	Example	Syllable structure	% Schwa [a]
/d/	[flud], [pud]	CCVC, CVC	23.91
/g/	[klɛ̆g], [fag]	CCVC, CVC	14.27
/b/	[glab], [fɔb]	CCVC, CVC	11.32

[a]% Schwa represents the proportion of the realization of final-schwa in words ending in the same coda, according to the corpus-based study of the distribution of the word-final schwa in speech production in Standard French (Hutin et al., 2021).

without a final "e." As observed, the proportion of one-syllable responses decreases as the probability of the spelling rime not ending in "e" decreases. This trend is more pronounced among literate adults and children at fifth grade reading level than it is among children at the first grade and second grade reading level.

To explore the potential interaction, we employed a Bayesian linear mixed-effects model with the "brms" package (see Bürkner, 2017) in R (R Core Team, 2023). We employed the default priors of the "brms" package for each parameter. The model employed a Bernoulli distribution appropriate for binary outcomes. Bayesian inference was carried out through Markov Chain Monte Carlo sampling, using four chains with 10,000 iterations each, which includes a warm-up phase of 1,000 iterations to ensure chain convergence. Our statistical inferences are based on the posterior distributions, for which we obtained estimates of the mean (i.e., the coefficient *b*), its standard error (SE), and its 95% credible interval (CrI). Evidence in favor of an effect is taken when its corresponding 95% CrI did not cross zero.

The fixed factors in the model were reading level ("RL") and probability that spellings do not end in "e" ("proba_without_e (spelling)"), as well as their interaction. We fitted random intercepts and slopes for participants and random intercepts for test items. This structure allows the model to account for individual differences among participants in their baseline likelihood of providing one-syllable responses and in their sensitivity to statistical orthographic knowledge. It also accounts for intrinsic differences among the test items in providing one-syllable responses. The models converged successfully and the values of $\hat{R}$ were 1.00 for all parameters. The formula of the model is as follows:

Dependent variable (one- or two-syllable responses) ~ RL × proba_without_e + (1 + proba_without_e | participant) + (1 | test_item)

The model predicted one-syllable responses, with "literate adults" serving as the reference level. The choice of "literate adults" as the baseline is motivated by two main reasons. First, the research question focuses on the evolution of the effect of statistical orthographic knowledge as the reading level changes. We predicted that the effect of statistical orthographic knowledge is more pronounced among literate adults. Choosing "literate adults" as the baseline highlights the absence or reduced presence of the effect of statistical orthographic knowledge among children with a lower reading level in comparison with literate adults. Second, literate adults have the most advanced reading skills and thus can provide a stable baseline for comparisons.

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Table 4. Average proportion (%) of one-syllable responses (standard deviation) across reading level by rimes probability of non-"e" ending spellings in pseudowords

Rime	Probability that spelling does not end in "e"	Literate adults	Fifth grade	Second grade	First grade
[ɛf]	0.982	73.0 (44.5)	68.3 (46.8)	40.2 (49.2)	34.0 (47.5)
[us]	0.785	77.0 (42.1)	63.5 (48.4)	44.7 (49.9)	36.1 (48.2)
[ys]	0.455	80.7 (39.5)	61.5 (48.9)	43.2 (49.7)	30.6 (46.2)
[is]	0.396	77.5 (41.9)	67.3 (47.1)	48.5 (50.2)	33.3 (47.3)
[uf]	0.292	76.2 (42.7)	60.6 (49.1)	40.2 (49.2)	31.3 (46.5)
[ɑ̃s]	0.048	66.4 (47.3)	59.6 (49.3)	41.7 (49.5)	35.4 (48.0)
[iʃ]	0.018	65.2 (47.7)	61.5 (48.9)	41.7 (49.5)	31.3 (46.5)
[yʃ]	0	64.8 (47.9)	52.9 (50.2)	40.9 (49.4)	35.4 (48.0)
[uʃ]	0	64.3 (48.0)	49.0 (50.2)	37.1 (48.5)	31.3 (46.5)
[ɑ̃ʃ]	0	61.9 (48.7)	46.2 (50.1)	34.8 (47.8)	29.9 (45.9)

The model revealed that the intercept was estimated at 1.30 (SE = 0.30, 95% CrI [0.73, 1.91]). There were negative effects for the first grade reading level ($b = -2.57$, SE = 0.48, 95% CrI = [−3.53, −1.64]), the second grade reading level ($b = -1.86$, SE = 0.48, 95% CrI = [−2.81, −0.91]). However, there was no effect for the fifth grade reading level ($b = -0.99$, SE = 0.52, 95% CrI = [−2.04, 0.03]), when compared with the baseline level. This indicates that the log-odds of providing one-syllable responses decrease as the reading level descends to lower school grades compared with literate adults. Moreover, an interaction effect was observed between reading level and spelling (Figure 1). We changed the reference level of the model to examine the effect of spelling in each reading level group. There was an effect of spelling on the log odds of providing one-syllable responses for literate adults ($b = 0.01$, SE = 0.00, 95% CrI = [0.01, 0.02]) and a small effect of spelling for fifth grade reading level ($b = 0.01$, SE = 0.00, 95% CrI = [0.00, 0.01]). There was no effect for second grade reading level ($b = 0.00$, SE = 0.00, 95% CrI = [−0.01, 0.01]) or first grade reading level ($b = 0.00$, SE = 0.00, 95% CrI = [−0.01, 0.01]).

Plosive-coda set

For the plosive-coda set, 47.6% were counted as one-syllable sequences and 52.4% were counted as two-syllable sequences, regardless of the reading levels. As shown in

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

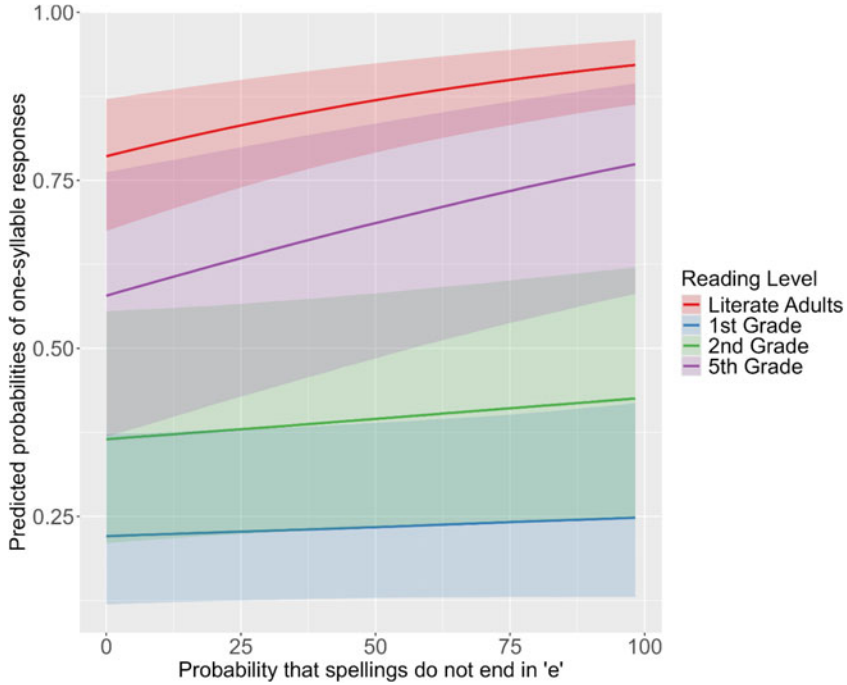


Figure 1. Regression graph showing the effect of the interaction of reading level and probability of spelling the rime without "e" on the predicted probability of one-syllable responses.

Table 5, there is no relationship between the probability of being followed by a schwa and the proportion of one-syllable responses in the syllable counting of pseudowords ending in a plosive coda across the four reading level groups.

To further evaluate this observation, we conducted a Bayesian mixed-effects analysis including schwa probability, reading level, and their interaction as predictors. This analysis revealed no reliable effect of schwa probability ($\beta = 0.08$, 95% CrI [−0.06, 0.22]), nor any interaction with reading level, confirming that acoustic factors related to schwa realization do not account for participants' responses.

Discussion

In Experiment 1, we examined whether statistical orthographic knowledge influences syllable-counting responses by testing two types of stimuli: a fricative-coda set, used to assess the influence of orthographic probability, and a plosive-coda set, used as a control for possible effects of schwa probability in speech production.

The results of the plosive-coda set revealed that one- or two-syllable responses were unrelated to the probability of being followed by a perceptual final schwa in speech as determined by the oral corpus analysis data by Hutin et al. (2021), indicating that the speech environment was not the source of the participants' counting.

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Table 5. Proportion (%)of one-syllable responses for pseudowords ending in plosives codas across reading level

Plosive codas	% Schwa	Literate adults	Fifth grade	Second grade	First grade
/b/	11.32	65.08	58.33	37.12	37.04
/g/	14.27	59.82	56.01	35.23	34.90
/d/	23.91	64.48	54.81	38.13	31.94

For fricative-coda set, data showed that the effect of statistical orthographic knowledge was observed in participants who had a wide knowledge of orthography (i.e., fifth grade and adult-like reading levels) and not in beginning readers (first- and second-grade reading levels). As familiarity with reading increases, we observe a parallel increase in monosyllabic responses (see Table 4) and in the influence of orthographic forms (see Figure 1). The results revealed that there was no effect of statistical orthographic knowledge for the first- and second-grade reading levels, but an effect for the fifth-grade reading level and literate adults (see Chetail & Content, 2017 for similar developmental tendency).

Before discussing the implications of these important results, we will first try to understand why participants at the first- and second-grade reading levels were more inclined to respond that pseudowords were bisyllabic, whereas stimuli were generally monosyllabic according to linguistic and acoustic criteria. To further explore these results, we conducted a second experiment (Experiment 2) in kindergarten, using only the fricative-coda set.

Experiment 2

The results of two-syllable responses in Experiment 1 for the fricative-coda set observed in first- and second-grade reading levels might depend on teaching methods in early school years. Indeed, in the first years of French primary school, many spelling errors are usually observed, especially in word endings (Sénéchal, Gingras & L'Heureux, 2016). To facilitate the learning process and reduce this type of errors, teachers emphasize final codas by usually adding a final schwa. For example, for a word as *carte* ([kart], en. "card") ending with the bigram "te" where the "t" is pronounced but not the "e," teachers emphasize the word ending adding in the pronunciation a schwa vowel (e.g., [kartə]) to ensure that children write the final "e." Moreover, to support the encoding of this orthographic final "e," they repeat the word, syllabifying it in two syllables [kar]-[tə], even if it is a monosyllable in speech. All these explicit instructions provided at the beginning of literacy acquisition could entail a two-syllable counting for monosyllables ending with a rime that could be spelled with an "e." To verify this hypothesis, we proposed the same experiment to a group of prereaders in kindergarten. We supposed that kindergarteners, who have not already received explicit instructions of writing, should count one syllable for our acoustic pseudowords.

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Ethics

The experiment was approved by the Ethics Committee of University of Lorraine (CER-UL, number 2023/262).

Participants

Twenty-four children in the final year of kindergarten participated in this study, comprising 12 girls and 12 boys, with an average age of 5 years and 2 months. All participants were native French-speaking children with no reported hearing problems and normal or corrected vision. Prior to participation, participants' parents or legal guardians were provided with a letter detailing the study and a consent form, which were subsequently sent home with the students. Two participants did not complete the entire study due to a lack of willingness to continue.

We assessed participants' letter knowledge and syllable reading ability. Specifically, they were asked to name eight letters ("A", "I", "O", "U", "B", "F", "T", and "S"), to produce the sounds of four consonant letters ("B", "F", "T", and "S"), and to read 16 consonant-vowel syllables ("BA", "BI", "BO", "BU", "FA", "FI", "FO", "FU", "TA", "TI", "TO", "TU," "SA," "SI," "SO," and "SU"). The results showed that they were prereaders with high letter knowledge (the average score was 7.6 over 8), low letter sound knowledge (the average score was 0.4 over 4), and low syllable-reading knowledge (the average score was 1.2 over 16).

Stimuli and procedure

The stimuli were the same as in Experiment 1. We proposed to participants the fricative-coda set and 20 fillers (see Appendix E). The procedure was the same as in Experiment 1, except that participants were asked to respond verbally and clap their hands if the stimulus was composed of one or two syllables. The experiment duration was 15 minutes in a single testing session.

Results and discussion

Firstly, four participants who did not achieve more than 80% correct answers on the filler items were excluded from the analysis. The proportion of correct responses was 98.4% for monosyllabic fillers and 90.6% for bisyllabic fillers.

As expected, prereader participants counted 98% of pseudowords as one syllable, compared with only 32.6% among participants at the first-grade reading level. The results of Experiment 2 in kindergarten showed that before schooling and the explicit teaching of written French, children considered all the stimuli to be monosyllables.

Table 6 shows the developmental evolution of the proportion of one-syllable responses, including the results of the fricative-coda set of Experiment 1. For prereaders, the stimuli were not ambiguous; they were considered simple monosyllabic sequences. At the beginning of reading acquisition, however, children showed a marked increase in two-syllable responses, suggesting that these stimuli became open to an alternative interpretation once explicit instruction in writing began and schwa was more likely to be emphasized in teaching contexts. As reading

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Table 6. Mean proportion (%) of one-syllable responses in prereaders (Experiment 2), first-, second-, and fifth-grade reading levels and literate adults (Experiment 1, fricative-coda set)

Literate adults	Fifth grade	Second grade	First grade	Prereaders
69.1	59.0	41.3	32.6	97.9

levels subsequently increased, the proportion of one-syllable responses rose again, consistent with a reduced reliance on explicit instruction provided in first-grade-driven interpretations.

According to our data, before primary school, children's responses were based on acoustic criteria, resulting mainly in one-syllable responses. Only when starting primary school, the stimuli become ambiguous, allowing a C(C)V.CV two-syllable counting. This ambiguity of the stimuli did not disappear with literacy acquisition since it was observed in second-grade, fifth-grade reading levels and even in literate adults.

In some way, the difference in syllable-counting observed between kindergarten and first-grade reading level involved literacy acquisition, since the explicit instructions provided by teachers during the first years of primary school oriented the phonological counting. It is also possible to suppose that these results depend on a potential new syllable structure with a different number of elements that become the referential phonological structure to accomplish the task. The explicit instructions emphasizing ending schwa provided at the beginning of literacy acquisition (i.e., for "porte" teachers would say [pɔʁ.tə] instead of [pɔʁt]) could entail a phonological representation that anchors the ending vowel [ə] to a C(C)V.Cə structure resulting in a two-syllable counting for C(C)VC monosyllables. This explanation is consistent with Ehri and Wilce's (1980) explanation which evokes the mark of writing in the conceptualization of phonological structure.

General discussion

The main goal of the experiments reported in this article was to examine whether an influence of statistical orthographic knowledge could be observed on a phonological awareness task and whether this effect varies across different reading levels.

In Experiment 1, the fricative-coda set results indicated that there was an effect of statistical orthographic knowledge—defined as the frequency with which a phonological unit is written in a language—on the phonological awareness task. This effect was observed at the fifth-grade and literate adult reading levels, which presumably have wide orthographic knowledge and can generalize their statistical properties. However, this effect was not observed in beginning readers. Moreover, the plosive-coda set results revealed that the proportion of one-syllable responses was unrelated to the probability of phonological codas being followed by a final schwa in speech, as determined by the corpus analysis data by Hutin et al. (2021). This justified the exclusion of this variable from our study, as the probability of a final schwa is likely related to other linguistic factors (such as prosody or the speaker's speech style) not present in our experimental stimuli.

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Experiment 2 further clarified the nature of the task by showing that prereaders overwhelmingly identified the same C(C)VC pseudowords as monosyllabic. Taken together, these findings indicate that the stimuli are phonologically unambiguous in the absence of orthographic knowledge, and that the observed variability emerges over the course of literacy acquisition.

Importantly, the developmental trajectory revealed a nonlinear pattern. Although prereaders relied on acoustic–phonological information and consistently produced monosyllabic responses, beginning readers showed a marked increase in two-syllable responses. We propose that this shift reflects the influence of instructional practices during early literacy acquisition. In French classrooms, children are often encouraged to align orthography and phonology by explicitly pronouncing the final "e" (schwa) in words that are phonologically monosyllabic. For example, a word such as /lis/ ("lisse") may be articulated as /li.sə/ to make its orthographic structure more transparent. This may lead to an overgeneralization of an orthographic syllabification strategy to novel C(C)VC sequences, thereby introducing ambiguity at a metalinguistic level.

As reading experience increases, this instructional influence diminishes, and responses gradually return toward phonological structure. However, even in literate adults, responses do not fully converge on the phonological pattern, with only about 70% of responses being monosyllabic. At this stage, performance becomes sensitive to the distributional properties of orthography, as reflected in the influence of rime-based spelling probabilities. This suggests that phonological judgments are shaped by a competition between phonological representations and statistical orthographic knowledge.

One possible alternative explanation is that the distribution of filler items may have induced a response bias, leading participants to produce a higher proportion of two-syllable responses. However, given that the same filler structure was used across all participant groups, such a bias cannot explain the developmental evolution of responses, nor the systematic influence of statistical orthographic knowledge observed in the present study.

Previous studies have shown the influence of orthographic word forms on phonological awareness tasks (Ehri & Wilce, 1980; Castles et al., 2003; Stuart, 1990; Tyler & Burnham, 2006) and the impact of statistical orthographic knowledge on orthographic tasks (Chetail, 2015; Pacton et al., 2001; Samara & Caravolas, 2014; Treiman & Boland, 2017; Treiman & Kessler, 2006). Our findings uniquely show that phonological responses are significantly shaped by the statistical properties of associations between phonological units and specific letter sequences across various orthographic forms. The simple influence of individual lexical properties cannot account for our data. First, the stimuli were acoustic pseudowords with no defined spelling pattern. Furthermore, the orthographic information did not concern a particular lexical representation, but the frequency with which the same phonological rime is spelled in different orthographic forms across the lexicon. For example, the rime [uʃ] is always followed by an orthographic "e" in the corresponding written form as "bouche" [buʃ], "touche" [tuʃ], "douche" [duʃ], or "couche" [kuʃ], whereas [is] is occasionally followed by an orthographic "e" in word forms as "dix" [dis], "six" [sis], "fils" [fis], or "jadis" [ʒa.dis] and also "hisse" [is] or "lisse" [lis]. This indicates that phonological responses are sensitive to the statistical properties of associations between phonological units and specific letter sequences

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

across various orthographies, which people might generalize progressively as they learn to read and spell.

An additional issue concerns the mechanism underlying orthographic influences observed in phonological tasks. Because syllable-counting is a metalinguistic task, participants may strategically recruit orthographic knowledge when making their judgments. For instance, Ventura et al. (2001) interpreted orthographic effects observed in a syllable-blending task as reflecting a postperceptual influence on conscious phonological representations.

At the same time, a growing body of evidence suggests that orthographic information can also be activated automatically during spoken language processing (e.g., Taft et al., 2008; Ziegler & Ferrand, 1998). In the present study, the task was relatively simple and participants were instructed to respond spontaneously, which may favor a more automatic influence of orthographic knowledge. Similar accounts have been proposed in other contexts involving verbal responses, such as the word superiority effect (e.g., McClelland & Rumelhart, 1981). Nevertheless, the present data do not allow us to fully disentangle strategic from automatic mechanisms.

Rather, our findings suggest a more complex interaction between phonological and orthographic mental representations than direct connections between mental lexical orthographic and phonological representations described by Tyler and Burnham (2006) and Castles et al. (2003). Such a complex interaction was also observed in a symmetric task using orthographic inputs and appears to be sensitive to phonological patterns: the illusory conjunction task (Prinzmetal et al., 1986). It consists in very quickly presenting a word whose letters are written in two different inks. Participants are instructed to detect a target letter and to report its color. In fact, several studies used this paradigm to examine what are the orthographic sublexical units in visual word recognition and whether such units are influenced by phonology (Prinzmetal et al., 1986; Prinzmetal et al., 1991; Rapp, 1992; Seidenberg, 1987). Doignon and Zagar (2005) used this paradigm with French speakers to understand the respective influence of bigram boundaries (defined by their frequency in written French) and syllable boundaries on reading units. Data showed that syllable boundaries produced a bias on illusory conjunction. Moreover, the analysis of the illusory conjunction errors revealed that the distributional properties of written language amplified the syllable effect, and that bigram boundaries only had an impact when they coincided with syllable boundaries. These results were replicated in the study by Doignon-Camus et al. (2009) using monosyllabic words and pseudowords. These findings suggest that phonological syllables had a main influence on the perception of orthographic sublexical units, which are simultaneously modulated by orthographic redundancy. Overall, the illusory conjunction study reflects how individuals reorganize an "ambiguous" visual stimulus, and the present study reflects how individuals reorganize an "ambiguous" auditory stimulus. These findings provide a complex theoretical account of the relationship between orthography and phonology, rather than a simple direct association between them.

Theoretical implications

Both findings—from the studies using phonological inputs (i.e., the present study) and orthographic inputs (i.e., illusory conjunction studies)—suggest that responses are sensitive to statistical properties progressively built around the connections

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

between phonological mental representations and their orthographic counterparts. The formation of this statistical knowledge is a progressive process resulting from learning to read, specifically, the cumulative learning of orthographic-phonological correspondences.

The current findings are consistent with a theoretical framework that conceptualizes reading acquisition as the construction of an orthographic-phonological system, centerd around three levels: orthography, phonology, and a set of associative units mediating between them (Guo et al., 2023). This conception aligns with the family of connectionist models, notably the first such model capable of accounting for the generalization and use of knowledge concerning orthographic redundancy and orthographic-phonological correspondences, known as the Parallel Distributed Processing (PDP) model (Seidenberg & McClelland, 1989). The PDP model posits that a single network facilitates the reading of a quasi-regular system such as English. In their model, the relationships between orthography and phonology are not directly connected but are mediated by hidden units, which "encode what the model knows about written English—specifically, orthographic redundancy (i.e., the frequency and distribution of letter patterns in the lexicon) and the correspondences between orthography and phonology" (Seidenberg, 1989, p. 26).

Building on this, Guo et al.'s theoretical framework explores the composition of the "hidden units," termed "associative units." It uniquely claims that this level of associative units results from a progressive reorganization of associations between sequences of letters and phonological representations. This level starts to emerge once individuals are exposed to any orthographic-phonological associations. As individuals learn to read and encounter more orthographic-phonological pairs, the influence of these associative units intensifies, enabling children to extract grapheme–phoneme correspondence regularities (see also Vazeux et al., 2020). With increasing reading proficiency, the role of these associative units solidifies, particularly in advanced readers. This cognitive architecture can account for the emerging statistical orthographic knowledge that is potentially developed at the level of associative units through literacy (for a similar view see Steacy et al., 2019). In this way, this theoretical framework could explain our results in the phonological awareness tasks. In our study with prereaders (Experiment 2), who have not yet been sufficiently exposed to orthographic representations, there was no significant development at the level of associative units. Consequently, their responses to the syllable-counting task depend on the acoustic criteria, limited to the phonological level. As individuals learn to read and spell words (Experiment 1), the interlevel level of associative units progressively forms to generalize statistical orthographic knowledge, influencing their phonological responses.

Conclusion

The current study shows that responses on the phonological awareness task are influenced by statistical orthographic knowledge. The findings suggest a nondirect association between orthography and phonology, assuming that statistical orthographic knowledge is generalized at the level of associative units. These units are progressively formed through the interaction between orthographic and phonological representations as literacy levels evolve. There is no doubt that more

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

experiments are needed to explore this issue. In particular, it would be interesting to examine whether statistical orthographic knowledge can be observed in languages other than French and how it might implicitly influence syllable counting.

Replication package. Stimulus materials, data, analysis scripts and experimental programs are available at Open Science Framework: https://osf.io/qgmhp/.

Acknowledgments. Teng Guo was funded by the Chinese Scholarship Council, P.R. China.

Competing interests. The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Note

1 Here, we used the phonetic transcriptions provided in the reference rather than the International Phonetic Alphabet.

References

Adams, M. J. (1998). *Beginning to read: Thinking and learning about print*. MIT Press.

Bertelson, P. (1972). Listening from left to right versus right to left. *Perception*, **1**(2), 161–165. https://doi.org/10.1068/p010161

Boersma, P. & Weenink, D. (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.09, retrieved 2 March 2023 from https://www.praat.org/

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—A causal connection. *Nature*, **301**(5899), 419–421. https://doi.org/10.1038/301419a0

Brady, S. A., & Shankweiler, D. P. (Eds.). (1991). *Phonological processes in literacy: A tribute to Isabelle Y. Liberman*. Lawrence Erlbaum Associates, Inc. https://doi.org/10.4324/9781315044439

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, **80**, 1–28. https://doi.org/10.18637/jss.v080.i01

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, **89**(4), 631–644. https://doi.org/10.1037/0022-0663.89.4.631

Castles, A., Holmes, V. M., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *The Quarterly Journal of Experimental Psychology: Section A*, **56**(3), 445–467. https://doi.org/10.1080/02724980244000486

Cavalli, E., Duncan, L. G., Elbro, C., El Ahmadi, A., & Colé, P. (2017). Phonemic—Morphemic dissociation in university students with dyslexia: An index of reading compensation? *Annals of Dyslexia*, **67**(1), 63–84. https://doi.org/10.1007/s11881-016-0138-y

Chetail, F. (2015). Reconsidering the role of orthographic redundancy in visual word recognition. *Frontiers in Psychology*, **6**, 645. https://doi.org/10.3389/fpsyg.2015.00645

Chetail, F., & Content, A. (2017). The perceptual structure of printed words: The case of silent E words in French. *Journal of Memory and Language*, **97**, 121–134. https://doi.org/10.1016/j.jml.2017.07.007

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, **16**(4), 234–244. https://doi.org/10.1038/nrn3924

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, **330**(6009), 1359–1364. https://doi.org/10.1126/science.1194140

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, **49**(2), 264–271. https://doi.org/10.1037/1196-1961.49.2.264

Doignon, N., & Zagar, D. (2005). Illusory conjunctions in French: The nature of sublexical units in visual word recognition. *Language and Cognitive Processes*, **20**(3), 443–464. https://doi.org/10.1080/01690960444000269

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Doignon-Camus, N., Zagar, D., & Mathey, S. (2009). Can we see syllables in monosyllabic words? A study with illusory conjunctions. *European Journal of Cognitive Psychology*, **21**(4), 599–614. https://doi.org/10.1080/09541440802052543

Donnenwerth-Nolan, S., Tanenhaus, M. K., & Seidenberg, M. S. (1981). Multiple code activation in word recognition: Evidence from rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, **7**(3), 170–180. https://doi.org/10.1037/0278-7393.7.3.170

Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, **1**(4), 371–385. https://doi.org/10.1017/S0142716400009802

Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory?*Journal of Experimental Psychology: Human Perception and Performance*, **16**(1), 77–91. https://doi.org/10.1037/0096-1523.16.1.77

Goswami, U., & Bryant, P. (2016). *Phonological Skills and Learning to Read.* Routledge.

Guo, T., Vazeux, M., Doignon-Camus, N., Bosse, M.-L., Mahé, G. & Zagar, D. (2023). Before learning the code: A commentary on sargiani, Ehri, and Maluf (RRQ, 2022). *Reading Research Quarterly*, **58**, 103–112. https://doi.org/10.1002/rrq.487

Hallé, P. A., Chéreau, C., & Segui, J. (2000). Where is the /b/in "absurde" [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language*, **43**(4), 618–639. https://doi.org/10.1006/jmla.2000.2718

Hutin, M., Jatteau, A., Vasilescu, I., Lamel, L., Wu, Y., & Adda-Decker, M. (2021). A corpus-based study of the distribution of word-final schwa in Standard French and what it teaches us about its phonological status. *Isogloss. Open Journal of Romance Linguistics*, **7**, 1–27. https://doi.org/10.5565/rev/isogloss.152

Jakimik, J., Cole, R. A., & Rudnicky, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language*, **24**(2), 165–178. https://doi.org/10.1016/0749-596X(85)90022-1

Laks, B. (1995). A connectionist account of French syllabification. *Lingua*, **95**(1-3), 51–76. https://doi.org/10.1016/0024-3841(95)90101-9

Lefavrais, J. (2005). *Alouette-R: Test d'analyse de la lecture et de la dyslexie*. Paris: Editions du Centrede Psychologie Appliquée.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, **18**(2), 201–212. https://doi.org/10.1016/0022-0965(74)90101-5

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, **44**(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, **88**(5), 375. https://doi.org/10.1037/0033-295X.88.5.375

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, **138**(2), 322–352. https://doi.org/10.1037/a0026744

Morais, J. (2021). The phoneme: A conceptual heritage from alphabetic literacy, *Cognition*, **213**, 104740. https://doi.org/10.1016/j.cognition.2021.104740

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, **36**(3), 516–524. https://doi.org/10.3758/BF03195598

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology, General*, **130**, 401–426. https://doi.org/10.1037/0096-3445.130.3.401

Perre, L., Midgley, K., & Ziegler, J. C. (2009). When *beef* primes *reef* more than *leaf*: Orthographic information affects phonological priming in spoken word recognition. *Psychophysiology*, **46**(4), 739–746. https://doi.org/10.1111/j.1469-8986.2009.00813.x

Prinzmetal, W., Hoffman, H., & Vest, K. (1991). Automatic processes in word perception: An analysis from illusory conjunctions. *Journal of Experimental Psychology: Human Perception and Performance*, **17**(4), 902–923. https://doi.org/10.1037/0096-1523.17.4.902

Prinzmetal, W., Treiman, R., & Rho, S. H. (1986). How to see a reading unit. *Journal of Memory and Language*, **25**(4), 461–475. https://doi.org/10.1016/0749-596X(86)90038-0

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

R Core Team (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Racine, I., & Grosjean, F. (2005). Le coût de l'effacement du schwa lors de la reconnaissance des mots en français [The cost of the obliteration of the schwa at the time of the recognition of French words]. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, **59**(4), 240–254. https://doi.org/10.1037/h0088052

Rapp, B. (1992). The nature of sublexical orthographic organization: The bigram trough hypothesis examined. *Journal of Memory and Language*, **31**(1), 33–53. https://doi.org/10.1016/0749-596X(92)90004-H

Samara, A., & Caravolas, M. (2014). Statistical learning of novel graphotactic constraints in children and adults. *Journal of Experimental Child Psychology*, **121**, 137–155. https://doi.org/10.1016/j.jecp.2013.11.009

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (pp. 245–263). Lawrence Erlbaum Associates, Inc.

Seidenberg, M. S. (1989). Visual word recognition and pronunciations: A computational model and its implications. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 25–74). MIT Press.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**(4), 523–568. https://doi.org/10.1037/0033-295X.96.4.523

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, **5**(6), 546–554. https://doi.org/10.1037/0278-7393.5.6.546

Sénéchal, M., Gingras, M., & L'Heureux, L. (2016). Modeling spelling acquisition: The effect of orthographic regularities on silent-letter representations. *Scientific Studies of Reading*, **20**(2), 155–162. https://doi.org/10.1080/10888438.2015.1098650

Steacy, L. M., Compton, D. L., Petscher, Y., Elliott, J. D., Smith, K., Rueckl, J. G., Sawi, O., Frost, S. J., & Pugh, K. R. (2019). Development and prediction of context-dependent vowel pronunciation in elementary readers. *Scientific Studies of Reading*, **23**(1), 49–63. https://doi.org/10.1080/10888438.2018.1466303

Stuart, M. (1990). Processing strategies in a phoneme deletion task. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, **42**(2–A), 305–327. https://doi.org/10.1080/14640749008401224

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**, 366–379. https://doi.org/10.1016/j.jml.2007.11.002

Taft, M., & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language*, **24**(3), 320–335. https://doi.org/10.1016/0749-596X(85)90031-2

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

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology*, **98**(3), 642–652. https://doi.org/10.1037/0022-0663.98.3.642

Tyler, M. D., & Burnham, D. K. (2006). Orthographic influences on phoneme deletion response times. *Quarterly Journal of Experimental Psychology*, **59**(11), 2010–2031. https://doi.org/10.1080/17470210500521828

Vazeux, M., Doignon-Camus, N., Bosse, M. L., Mahé, G., Guo, T., & Zagar, D. (2020). Syllable-first rather than letter-first to improve phonemic awareness. *Scientific Reports*, **10**(1), 22130. https://doi.org/10.1038/s41598-020-79240-y

Ventura, P., Kolinsky, R., Brito-Mendes, C., & Morais, J. (2001). Mental representations of the syllable internal structure are influenced by orthography. *Language and Cognitive Processes*, **16**(4), 393–418. https://doi.org/10.1080/01690960042000184

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

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition*, **32**(5), 732–741. https://doi.org/10.3758/BF03195863

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Appendix A: Six Trials in the Training Block.

Phonetic forms
[mlo]
[ʃla]
[blo]
[zoti]
[valto]
[kʁigu]

Appendix B: Forty Fillers Used in the Experiment 1 (20 Bisyllables, 20 Monosyllables)

Bisyllables	Monosyllables
[pito]	[mʁi]
[paty]	[kli]
[tipa]	[spo]
[katy]	[sfi]
[klɔ̃pi]	[vʁo]
[vuke]	[zba]
[medu]	[ʒɛ̃]
[miʁlɛ̃]	[tly]
[bidɑ̃]	[sto]
[bɔʁly]	[dlo]
[bylni]	[ksa]
[dopiʁ]	[ʃʁu]
[dygɑ̃]	[sly]
[gʁazo]	[zlɔ̃]
[guʃĩ]	[mʁa]
[niʃy]	[ɥaj]
[nefo]	[tli]
[fozy]	[ɥij]
[sopy]	[sla]
[sakɔ̃]	[vzo]

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Appendix C: Forty Phonetic Forms of the Pseudowords of the Fricative-coda Set

[ys]	[ɛf]	[us]	[is]	[uf]
[klys]	[blɛf]	[klus]	[blis]	[bʁuf]
[fʁys]	[kʁɛf]	[dʁus]	[gʁis]	[kluf]
[skʁys]	[glɛf]	[plus]	[skis]	[gʁuf]
[snys]	[psɛf]	[slus]	[vʁis]	[tʁuf]
[ɑ̃ʃ]	[ɑ̃s]	[yʃ]	[iʃ]	[uʃ]
[kɑ̃ʃ]	[dʁɑ̃s]	[gʁyʃ]	[fliʃ]	[ʃuʃ]
[fɑ̃ʃ]	[glɑ̃s]	[pʁyʃ]	[gʁiʃ]	[spuʃ]
[stɑ̃ʃ]]	[psɑ̃s]	[slyʃ]	[stiʃ]	[tʁuʃ]
[dʁɑ̃ʃ]	[spɑ̃s]	[tlyʃ]	[vʁiʃ]	[vuʃ]

Appendix D: Forty Phonetic Forms of the Pseudowords of the Plosive-coda Set

Phonetic forms ending in /b/	Phonetic forms ending in /d/	Phonetic forms ending in /g/
[bɑ̃b]	[flud]	[ʃɛ̃g]
[bʁab]	[fɔd]	[klɛ̃g]
[fʁɑ̃b]	[kʁyd]	[fag]
[glab]	[dʁɔd]	[flɔg]
[klɔb]	[pud]	[glig]
[fɔb]	[gʁud]	[nag]
[psab]	[plɔd]	[plɛ̃g]
[skab]	[plyd]	[slɔg]
[skɑ̃b]	[skɔd]	[spig]
[plɑ̃b]	[stud]	[spɛ̃g]
[spɔb]	[styd]	[stag]
[tʁɔb]	[vyd]	[tʁag]
		[spɔg]
		[tig]
		[vʁig]
		[zɔg]

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press

Appendix E: Twenty Fillers Used in the Experiment 2 (10 Bisyllables, 10 Monosyllables)

Bisyllables	Monosyllables
[paty]	[kli]
[katy]	[spo]
[klɔ̃pi]	[vʁo]
[medu]	[sto]
[miʁlɛ̃]	[dlo]
[bidɑ̃]	[ksa]
[bɔʁly]	[ʃʁu]
[bylni]	[tli]
[niʃy]	[qij]
[nefo]	[sla]

Cite this article: Ruvoletto, S., Guo, T., Bartkova, K., Flesch, M., & Zagar, D. (2026). How many syllables do you hear in [skis]: One or two? It depends on whether you have learned to read or not. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716426100800

https://doi.org/10.1017/S0142716426100800 Published online by Cambridge University Press