Published in final edited form as:

Sci Stud Read. 2018 ; 22(3): 191–208. doi:10.1080/10888438.2017.1414219.

Statistical and Cooperative Learning in Reading: An Artificial Orthography Learning Study

Jingjing Zhao^{1,2}, Tong Li², Mark A. Elliott³, and Jay G. Rueckl²

¹ School of Psychology, Shaanxi Normal University and Key Laboratory for Behavior and Cognitive Neuroscience of Shaanxi Province, Xi'an, Shaanxi, China ² Department of Psychological Sciences, University of Connecticut, Storrs, CT, USA and Haskins Laboratories, New Haven, CT, USA ³ School of Psychology, National University of Ireland, Galway, Ireland.

Abstract

This paper reports two experiments in which the artificial orthography paradigm was used to investigate the mechanisms underlying learning to read. In each experiment, participants were taught the meanings and pronunciations of words written in an unfamiliar orthography and the statistical structure of the mapping between written and spoken forms (O-P) was manipulated independently of the mapping between written forms and their meanings (O-S). Our results support three main conclusions. First, the statistical structure of O-P and O-S mappings determined how easily each of those mappings was learned, suggesting that the learning of both mappings engages a common statistical learning mechanism. Second, learning to read is a cooperative process, in that learning in any particular component of the reading system is influenced by knowledge stored in the rest of the system. Finally, knowledge of sublexical regularities can be acquired as the result of exposure to words embodying those regularities.

Keywords

reading; statistical learning; cooperative learning; artificial orthography

Introduction

Studies of reading acquisition often involve tracking changes in the reading skills of beginning readers over months or even years (Backman, Bruck, Hebert, & Seidenberg, 1984; Bogaerts, Szmalec, De Maeyer, Page, & Duyck, 2016; Connor et al., 2013; Goswami, Gombert, & de Barrera, 1998; Linkersdörfer et al., 2015; Pan et al., 2016; Seymour, Aro, & Erskine, 2003; Yeatman, Dougherty, Ben-Shachar, & Wandell, 2012). While the results of such studies can provide important insights about the impact of cognitive (Backman et al., 1984; Bogaerts et al., 2016; Pan et al., 2016), biological (Linkersdörfer et al., 2015; Yeatman et al., 2012), environmental (Goswami et al., 1998; Seymour et al., 2003), and instructional

Address correspondence to: Jingjing Zhao, School of Psychology, Shaanxi Normal University, Xi'an, Shaanxi, China, jingjing.zhao@snnu.edu.cn, Phone/Fax: 86-29-85303395, Jay G. Rueckl, Department of Psychological Sciences, University of Connecticut, Storrs, CT, USA, jay.rueckl@uconn.edu, Phone/Fax: 1-860-486-5502.

(Connor et al., 2013) factors on literacy acquisition, they are generally not well-suited for testing mechanistic theories of the computational processes underlying learning to read (Harm & Seidenberg, 2004; Share, 1995; Ziegler & Goswami, 2005) or their neural bases (Pugh et al., 2008). In this paper we report the results of two experiments employing a complementary approach—a laboratory method that provides tighter control of key experiential factors than can be afforded by naturalistic approaches and thus readily accommodates a variety of theoretically informative experimental manipulations.

The experiments we report test hypotheses derived from the *triangle model* (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989), a theoretical framework that has been used to account for a variety of findings concerning skilled reading (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989), reading acquisition in typically developing and learning-disabled children (Harm & Seidenberg, 1999; Seidenberg & McClelland, 1989), acquired dyslexia (Plaut, McClelland, Seidenberg, & Patterson, 1996), cross-language differences (Plaut & Gonnerman, 2000; Yang, Shu, McCandliss, & Zevin, 2009, 2013), and the impact of reading intervention (Harm, McCandliss, & Seidenberg, 2003). A core principle of the triangle model is that the processes underlying word reading are shaped by statistical regularities in the mappings between the orthographic, phonological, and semantic representations of written words. These regularities are primarily at the sublexical level. Thus, for example, words with an orthographic body that is reliably paired with a particular phonological rime (e.g., “consistent” words such as *pill*, *mill*, and *still*) are read aloud faster and more accurately than words with inconsistent body-rime correspondences (e.g., *pint*) (Glushko, 1979; Jared, 2002). Similarly, effects of morphological structure on word reading (Burani & Caramazza, 1987; Fowler, Napps, & Feldman, 1985; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Taft & Forster, 1975) can be understood as the influence of regularities in the mappings between (orthographic and phonological) form and meaning (see Feldman, 1995). From the triangle model perspective, effects such as these are consequences of the computational processes that underlie reading acquisition. According to the model, learning to read is an “exercise in statistical learning” (Harm & Seidenberg, 1999) in the sense that learning attunes the organization of the reading system to the statistical structure of the writing system—including in particular the regularities embodied in the relationships between the orthographic, phonological, and semantic properties of the words in one’s language.¹

Developmental studies of reading acquisition (e.g., Backman et al., 1984; Deacon & Leung, 2013; Ricketts, Davies, Masterson, Stuart & Duff, 2016; Treiman & Kessler, 2006; Wang, Nickels, Nation, & Castles, 2013) provide a glimpse of these underlying learning processes “in the wild”, but these processes can be probed more directly using laboratory methods. For example, a growing body of studies (e.g., Bitan & Karni, 2003, 2004; Deng, Booth, Chou, Ding, & Peng, 2008; Deng, Chou, Ding, Peng, & Booth, 2011; Havas, Waris, Vaquero, Rodríguez-Fornells, & Laine, 2015; Merks, Rastle, & Davis, 2011; Moore, Brendel, & Fiez, 2014; Rueckl & Dror, 1994; Sandak et al., 2004; Taylor, Plunkett, & Nation, 2011; Trudeau,

¹The term “statistical learning” has been used to refer to the detection of spatial and temporal regularities in variety of linguistic and non-linguistic domains. There is evidence that learning to read is related to statistical learning in this broader sense (e.g. Arciuli & Simpson, 2012, Frost et al., 2013), although the precise nature of this relationship is not yet well understood

2006; Xue, Chen, Jin & Dong, 2006; Yoncheva, Wise, & McCandliss, 2015) have employed an experimental paradigm in which participants are taught the meanings and/or pronunciations of a set of made-up words. In some of these *artificial lexicon* studies, the words are printed in the script of the participants' native language and are given novel meanings or pronunciations that conform more-or-less to the orthography-to-semantics (O-S) or orthography-to-phonology (O-P) regularities of that writing system (e.g., Havas et al., 2015; Merkx et al., 2011; Rueckl & Dror, 1994; Trudeau, 2006). In others, the words are printed in an *artificial orthography*, a script that is either entirely made up (e.g., Bitan & Karni, 2003, 2004; Moore et al., 2014; Taylor et al., 2011; Yoncheva et al., 2015) or based on an unfamiliar writing system (e.g., for readers of English, a script based on Chinese characters (Deng, Booth, et al., 2008; Deng, Chou, et al., 2011), and for readers of Chinese, a script based on Korean characters (Xue et al., 2006)). There are, of course, important differences between these laboratory tasks and the circumstances associated with learning to read a natural language. However, laboratory results concerning the behavioral effects of lexical properties (e.g., Merkx et al., 2011; Rueckl & Dror, 1994; Taylor et al., 2011; Trudeau, 2006), instructional manipulations (e.g., Bitan & Karni, 2003, 2004; Yoncheva et al., 2015), and individual differences (e.g., Xue et al., 2006), as well as their neural correlates (e.g., Havas, Laine & Fornells, 2017; Sandak et al., 2004; Taylor, Davis, & Rastle, 2017), suggest that these experimental manipulations provide a theoretically informative probe of the processes involved in reading acquisition.

In the experiments reported below, we used the artificial orthography paradigm to investigate the role of statistical structure in learning to read. As explained above, because the learning process is understood to be fundamentally statistical, the model predicts that more structured mappings will be easier to learn. Moreover, because the theory posits that the same mechanisms underlie the computation of a written word's pronunciation and meaning, it predicts that statistical structure should facilitate the learning of O-P and O-S mappings equally. (For relevant computer simulations, see Harm & Seidenberg, 1999; Plaut & Gonnerman, 2000; Plaut et al., 1996; Rueckl & Raveh, 1999; and Seidenberg & McClelland, 1989.) Alternatively, other models assume O-S mappings and O-P mappings might be computed by different learning/memory mechanisms. For example, O-P learning might be more relevant to procedural learning and involve cortical-striatal-thalamic circuits, whereas O-S learning might be mediated by a declarative learning process and governed mainly by the medial temporal lobe (MTL), particularly the hippocampus (Ullman & Pierpont, 2005).

Previous laboratory studies provide some support for this hypothesis. A number of studies have demonstrated that statistical structure facilitates the learning of the mapping from orthography to phonology (e.g., Brooks, 1977; Brooks & Miller, 1979; Byrne, 1984; Deng et al., 2011; Trudeau, 2006; Taylor et al., 2011). For example, in an early study, Brooks (1977) taught English speakers the pronunciations of six new words by combinations of novel nonalphabetic symbols and timed their naming responses as they read repeatedly through a 400-trial word list. Participants who were assigned the systematic symbol-phoneme correspondences were found to have faster reading times over the last 200 trials than those who had been assigned an arbitrary symbol-phoneme relationship. Similarly, a number of studies have also demonstrated that statistical structure facilitates the learning of the mapping for orthography to semantics (e.g., Deng et al., 2008; Havas et al., 2015; Merkx et

al., 2011; Rueckl & Dror, 1994). For example, Rueckl and Dror (1994) trained English speakers over a 5-week period on the meanings of a set of English pseudowords belonging to different semantic categories (e.g., animal, clothes, furniture, etc.) and manipulated consistency between spelling and meaning of the pseudowords during training. Their results indicated that pseudowords with systematic spelling-to-meaning mappings (e.g., *durch-dog*, *hurch-cat*, *kurch-bear*) were easier to learn and more accurately identified than pseudowords without a systematic relationship between spelling and meaning (e.g., *durch-dog*, *hurch-shirt*, *kurch-table*).

Results such as these suggest that a statistical learning process underlies both O-P and O-S learning. However, because most of these studies investigated *either* O-P *or* O-S learning, they do not allow for a direct comparison of processes underlying the learning of these mappings. Moreover, studies that investigated both O-P and O-S learning simultaneously generally embody the same asymmetry present in the writing systems of most (if not all) natural languages (e.g., Sandak et al., 2004). In particular, even in a ‘deep’ orthography such as English, orthography and phonology are more highly correlated than orthography in semantics. That is, similarly spelled words are typically pronounced similarly as well, whereas orthographic similarity carries little information about meaning. (Consider, for example, the words *lake*, *take*, *make*, and *wake*.) Thus, comparisons of O-P and O-S learning (both in the wild and in laboratory experiments) typically confound differences in the domain to which written forms are mapped (phonology or semantics) with differences in the degree of statistical structure in those mappings.

In the present experiments we took advantage of the opportunities afforded by the artificial orthography paradigm to equate the statistical structure of the O-P and O-S mappings. The artificial orthographies used in our experiments were modeled on the Chinese writing system, and in particular, the structure of Chinese compound characters (*phonograms*). Approximately 80%–90% of modern Chinese characters are composed of two components called *radicals*, one that provides information about how that character is pronounced and the other providing information about its meaning. In real Chinese, the degree to which a radical is a reliable indicator of phonology or semantics varies (Shu, Chen, Anderson, Wu, & Xuan, 2003), and indeed readers are sensitive to this statistical structure (Lee et al., 2004; Ho, Ng, & Ng, 2003; Shu, Anderson, & Wu, 2000). Like Chinese phonograms, the words in our artificial orthographies were also composed of two components. In some conditions, these components were highly reliable indicators of either phonology or meaning. In other conditions these components carried minimal information about either domain. Critically, this scheme allowed us to independently manipulate the O-P and O-S mappings and to impose comparable amounts of structure on each.

Although the primary purpose of this study was to investigate the role of statistical structure in O-P and O-S learning, the structure of our experiments also allowed us to investigate two other hypotheses derived from the triangle model. One of these concerns the *cooperative* nature of the learning process. According to the model, learning does not occur independently in different parts of the reading system. Rather, learning in any particular component of the system is influenced by knowledge stored in the rest of the system (Harm & Seidenberg, 2004; Yang et al., 2013). Thus, because learning is cooperative in this sense,

the learning of a given mapping (e.g., the O-P mapping) is predicted to depend not only on the characteristics of that mapping, but also on the characteristics of the other mapping (e.g., O-S).

Although relatively few studies have investigated the cooperative nature of learning, the results from a couple of artificial orthography learning studies that have been reported are consistent with the cooperative hypothesis. For example, in Trudeau's (2006) artificial lexicon study, novel words (e.g., *bint*) were assigned meanings that were either highly imageable (e.g., *the blade part of an ice skate*) or not (e.g., *the status of being protected from dismissal*) and pronunciations that were either consistent (e.g., /bɪnt/) or inconsistent (e.g., /bɑrnt/) with English O-P regularities. O-P consistent words were learned more quickly than O-P inconsistent words and, critically, this consistency effect was modulated by semantic imageability—the consistency effect was smaller for novel words with highly imageable meanings than for words with more abstract meanings. This pattern mirrors what is observed in the real-word reading behavior of skilled adult readers of English (Strain, Patterson, & Seidenberg, 1995) and is taken a signature of cooperative learning (Harm & Seidenberg, 2004). Similar results were reported by Taylor et al. (2011), who investigated O-P learning using an artificial orthography and who also found that O-P learning was enhanced when participants learned the meanings of the novel words prior to O-P training. However, these two studies both examined the semantic contribution to O-P learning. None of previous studies have explored the phonological influences on O-S learning. In the present experiments we investigate cooperative learning influences of O-P on O-S and vice versa under conditions where the to-be-learned mappings are of equivalent structure.

Finally, we also used these experiments to investigate the impact of the circumstances of training on the detection of sublexical regularities. Learning in the triangle model benefits from sublexical regularities even when the model is only trained on whole words (Seidenberg & McClelland, 1989). Thus, while the model does not deny that training focused on sublexical regularities can enhance a reader's sensitivity to those regularities (Harm et al., 2003), the model predicts that whole-word training is a sufficient basis for the acquisition of sublexical knowledge. While several laboratory learning studies demonstrate that the acquisition of sublexical regularities is promoted by training tasks that direct learners' attention to the sublexical level (Bitan & Karni, 2003, 2004; Brooks & Miller, 1979; Yoncheva et al., 2015), these and other studies also suggest that sublexical knowledge can often be acquired via exposure to novel words in tasks that focus attention on the lexical level (e.g., Deng, Booth, et al., 2008; Deng, Chou, et al., 2011; Havas et al., 2015; Merks et al., 2011; Rueckl & Dror, 1994; Taylor et al., 2011; Trudeau, 2006). It is possible that in some of these cases meta-knowledge about the learner's native language could facilitate the abstraction of sublexical knowledge. For example, familiarity with an alphabetic writing system such as English might facilitate the learning of an artificial orthography that also embodies the alphabetic principle (e.g., Havas et al., 2015; Trudeau, 2006). In the present experiments the statistical regularities in the O-P and O-S mappings are relatively dissimilar to those of our participants' native writing system (English) and more like those of Chinese phonograms. Experiment 1 included training tasks focused at both the lexical and sublexical levels. In Experiment 2, all training was focused at the lexical level.

Experiment 1

To investigate the role of statistical structure in O-P and O-S learning, we constructed three artificial orthographies (based on Chinese phonograms) that shared the same written forms but differed in how those forms were mapped to phonology and semantics. To construct these orthographies we selected two non-overlapping sets of radicals; written words were composed of two radicals, one from each set. To create a systematic mapping from orthography to semantics, the definitions assigned to the words were systematically related to the radicals from one of the two sets. For example, the meanings assigned to all the words containing a particular radical were kinds of animals, whereas the words containing another semantic radical were defined as types of furniture. Thus, in the consistent O-S condition, the semantic radicals were systematically associated with different semantic categories. In contrast, in the inconsistent condition, the words and meanings were re-paired such that no two words that shared a particular ‘semantic’ radical referred to members of the same semantic category. Consistent and inconsistent O-P mappings were constructed in the same way. (See Figure 1.) For example, in the consistent O-P condition of Experiment 1, words were assigned two-syllable pronunciations and words that shared a particular phonological radical also shared the same second syllable. In the inconsistent condition, if two words shared a ‘phonological’ radical they had different second syllables.

Because the manipulations of O-P and O-S consistency involved different sets of radicals, the consistent and inconsistent O-P mappings and the consistent and inconsistent O-S mappings can be combined in various ways. For Experiments 1 and 2, we constructed three orthographies: consistent phonology/consistent semantics (CP/CS), consistent phonology/inconsistent semantics (CP/IS), and inconsistent phonology/consistent semantics (IP/CS). Participants learned these mappings by performing training tasks that entailed judgements about the phonological or semantic properties of the written words. Note that the contrast between the CP/CS and IP/CS conditions (on phonological tasks) provides evidence about the effect of structure on the learning of the O-P mapping (with the structure of the O-S mapping held constant). Similarly, the comparison of the CP/CS and CP/IS on semantic tasks provides evidence about the effect of statistical regularities on the learning of the O-S mapping. Note too that the comparison of CP/CS and CP/IS on phonological tasks indexes cooperative effects of the O-S knowledge on O-P learning, whereas the comparison of CP/CS and IP/CS on semantic tasks provides a measure of the cooperative effect of O-P knowledge on the learning of the O-S mapping.

Methods

Participants—Forty-eight University of Connecticut undergraduate students with normal hearing and normal or corrected-to-normal vision participated in this experiment. Five were bilingual speakers with English as their dominant language. None of either the monolingual or bilingual participants spoke Chinese. Participants were randomly assigned to the three orthography conditions, with sixteen in each condition.

Materials—Ten Chinese simple (one component) characters were selected (“士”, “中”, “下”, “刀”, “山”, “小”, “大”, “儿”, “广”, and “父”) to construct the written forms of the

artificial orthographies. Half of them were designated as phonetic radicals (meaning that their relation to pronunciation was manipulated) and the other half were designated as semantic radicals (meaning that their relation to meaning was manipulated) counterbalanced across subjects. Semantic and phonetic radicals were combined to generate 25 semantic-phonetic compound “characters”, with position (left or right) counterbalanced across subjects (e.g., “士小” vs. “小士”). Half the participants saw the characters with semantic radicals on the left and phonetic radicals on the right while the other half saw the opposite position order.

Each of the 25 semantic-phonetic compound characters was associated with a unique pronunciation and meaning. Twenty-five novel disyllables were generated by pairing 25 unique first syllables with five second syllables, each of which was used in five words (e.g., /faɪbi/, /gUbi/, /næbi/, /sɒbi/, and /pobi/). The object definitions were chosen by selecting five high-rank exemplars from each of five semantic categories (animal, body part, fruit, furniture, and clothing) using the Battig and Montague (1969) norms. As described above, in the consistent-phonology (CP) conditions, characters that shared the same phonological radical (e.g., “小” in “士小”, “中小”, “下小”, “刀小”, and “山小”) also shared the same second syllable (e.g., /bi/ in /faɪbi/, /gubi/, /næbi/, /sɒbi/, and /pobi/). Similarly, in the consistent-semantic (CS) conditions, characters that shared the same semantic radical (e.g., “士”) belonged to the same semantic category (e.g., animal). (For example, “士小”, “士大”, “士儿”, “士广”, and “士父” were paired with definitions of dog, cow, cat, horse, and lion.) In the conditions with an inconsistent O-P mapping (IP), the pairings of written forms and pronunciations were shuffled such that no two characters that shared the same phonetic radical also shared second syllables, and likewise, in the conditions with an inconsistent O-S mapping (IS), the pairings of written forms and meanings were shuffled such that no two characters that shared the same semantic radical referred to members of the same semantic category. Finally, in all conditions, the semantic radicals were never systematically related to the phonological properties of the words, nor were the phonological radicals ever systematically related to word meaning. Similarly, there were no regularities in the mapping from phonology to semantics (P-S) in any of the three orthographies. Thus, all three were P-S inconsistent.

Finally, in order to test participants’ knowledge of sublexical regularities, two sets of 25 transfer “pseudo-characters” were generated by combining the five phonetic/semantic radicals with five new radicals (“子”, “爪”, “手”, “尸”, and “戈”). Phonological transfer pseudo-characters shared the same phonetic radicals with the critical characters, but differed from the critical characters on the other part of the word forms. (E.g., the phonological transfer character “子小” shared the phonological radical “小” with some of the trained characters, but had a new semantic radical “子” that never appeared in any critical characters.) Semantic transfer pseudo-characters shared the same semantic radicals with the critical characters, but differed from the critical characters in the other part of the written forms. (E.g., the semantic transfer character “士爪” shared the semantic radical “士” with some of the trained characters, but had a new phonological radical “爪” that never appeared in any critical characters.)

Procedure—Participants were trained and tested in three phases. First, participants were pre-exposed to the pronunciations and meanings of all the 25 characters. Each character was presented in the middle of the computer screen for five seconds with a picture below it indicating its “meaning”, during which the auditory pronunciation of the character was presented through headphone three times. The pictures indicating the meaning of the characters were black-and-white line drawings from Snodgrass and Vanderwart (1980)’s picture database. Participants were asked to learn the pronunciation and meaning of each character by associating the written form of each character with its pronunciation and meaning. The 25 characters were randomly presented with one presentation per character.

Following this, participants were trained in four tasks that crossed mapping (O-P, O-S) with focus (lexical, sublexical). In the lexical phonology task, participants judged whether the pronunciation of the target character presented on the screen matched a spoken novel word presented through headphones. In the sublexical phonology task, participants judged whether the target character had the same second syllable as a comparison item (a real English word, e.g., “RUBY”), also displayed visually. In the lexical semantic task, participants decided whether the meaning of the target character matched a picture displayed visually. In the sublexical semantic task, participants decided whether the meaning of the target character belonged to the semantic category indicated by a real English word (e.g., ANIMAL). In all four tasks, subjects responded by pressing one of two marked buttons on the keyboard. Feedback about the accuracy of the response was provided by showing “Correct” or “Incorrect” on the computer screen, along with each target character’s pronunciation heard through headphones (in the phonological tasks) or each target character’s meaning indicated by a picture (in the semantic tasks).

Training trials were administered in a series of 24 blocks (6 blocks of each task). Training alternated between phonological and semantic tasks every block, and alternated between lexical and sublexical training tasks every other block. The initial training condition was counterbalanced across participants. Each block included 25 trials with each character presented once per block in a randomly determined order. The number of trials with correct yes and no responses were equated within blocks and balanced (for each character) across blocks. Participants were instructed to respond as quickly as possible or responses were timed out after five seconds.

In the third phase of the protocol, participants completed two judgment tasks and two recall tasks. In the judgment tasks, both trained characters and untrained transfer pseudo-characters were tested using a procedure that was the same as in the sublexical training tasks except without feedback. Thus, in the phonological judgement the participants judged whether each character (both trained and pseudo-characters) shared a final syllable with an English word presented visually, while in the semantic judgement task the participants decided whether each character referred to a member of a particular semantic category. In the recall tasks, participants were asked to verbalize the pronunciations or meanings for the 25 trained characters, with one repetition of each character in one trial presented randomly. The judgment tasks were tested first, and for both the judgment and the recall tasks the order of the semantic and phonological blocks was counterbalanced across subjects. The duration of

the whole experimental session was approximately 1.5 hours, with five minutes for pre-exposure learning, one hour for training, and 20 minutes for testing.

Data Analyses—The data for each task were normalized using the arcsine transform of the square root proportion of correct responses. Statistical analyses were conducted in SPSS (IBM SPSS Statistics Version 22 Chicago, IL). Analysis of variance (ANOVA) was run with normalized accuracy as dependent measure and consistency (CP/CS vs. CP/IS vs. IP/CS) as between-subject factor for each task, separately. For the training tasks, repetition block was included as a within-subject factor. For the judgment testing tasks, word type (trained vs. transfer) was included as a within-subject factor. In all analyses, counterbalancing condition was included as a nuisance variable. Data presented in figures and tables are the raw accuracy scores prior to normalization.

Results

Two sets of analyses were conducted. The first focused on the effects of consistency on the learning of the O-P and O-S mappings; the second addressed the cooperative learning hypothesis.

Effects of the statistical structure of the to-be-learned mapping—The primary aim of this experiment was to examine the effect of statistical structure on the learning of both the O-P and O-S mappings under comparable conditions. For the O-P mapping, the results from the phonological tasks in the CP/CS and IP/CS conditions provide a contrast of learning of consistent vs. inconsistent O-P mappings with the structure of the O-S mapping fixed. Similarly, a comparison of performance on the semantic tasks in the CP/CS and CP/IS conditions provides evidence about the role of consistency on O-S learning with the structure of the O-P mapping held constant. Given the hybrid nature of the design, separate analyses of these contrasts were conducted for O-P and O-S learning. The results of these analyses are reported for each training and post-training task in turn.

In the sublexical-focus training tasks, the consistency effect was significant for both phonological and semantic learning. As shown in Figure 2 (upper panel), phonological learning was better in the CP/CS condition than in the IP/CS condition [$F_{(1, 28)} = 66.49$, $\eta_p^2 = 0.704$, $p < .001$, power = 1]; and semantic learning was better in the CP/CS condition than in the CP/IS condition [$F_{(1, 28)} = 55.34$, $\eta_p^2 = 0.664$, $p < .001$, power = 1]. Not surprisingly, in both cases performance improved over blocks [phonological: $F_{(5, 140)} = 9.788$, $\eta_p^2 = 0.259$, $p < .001$, power = 1; semantic: $F_{(5, 140)} = 21.155$, $\eta_p^2 = 0.430$, $p < .001$, power = 1], as did the effect of consistency, as indicated by significant block $\times$ consistency interactions [phonological: $F_{(5, 140)} = 7.937$, $\eta_p^2 = 0.221$, $p < .001$, power = 0.999; semantic: $F_{(5, 140)} = 9.894$, $\eta_p^2 = 0.261$, $p < .001$, power = 1]. In contrast, in the lexical training tasks, the effect of consistency was not significant in either the phonological or semantic tasks, as shown in Figure 2 (lower panel).² In both cases, performance improved over blocks [phonological: $F_{(5, 140)} = 2.470$, $\eta_p^2 = 0.081$, $p < .05$, power = 0.763;

²Data from two participants in the semantic judgment tasks were not recorded due to a technical problem.

semantic: $F_{(5, 140)} = 30.332$, $\eta_p^2 = 0.520$, $p < .001$, power = 1], but in neither case was the block $\times$ consistency interaction significant.

Turning to the post-training tasks, O-P and O-S consistency had similar effects in both the judgment and recall tasks. As shown in Table 1, phonological judgments were more accurate in the CP/CS condition than in the IP/CS condition [$F_{(1, 28)} = 84.77$, $\eta_p^2 = 0.752$, $p < .001$, power = 1] and semantic judgments were more accurate in the CP/CS condition than in the CP/IS condition [$F_{(1, 26)} = 83.17$, $\eta_p^2 = 0.762$, $p < .001$, power = 1]. In both cases, responses were more accurate for trained words than for transfer words [phonological: $F_{(1, 28)} = 4.198$, $\eta_p^2 = 0.130$, $p < .05$, power = 0.507; semantic: $F_{(1, 26)} = 5.067$, $\eta_p^2 = 0.163$, $p < .05$, power = 0.582], but in neither case was the word type $\times$ consistency interaction significant. Finally, performance in the recall tasks also revealed the effects of both O-P and O-S consistency. As Table 2 illustrates, phonological recall was better in the CP/CS condition than in the IP/CS condition [$F_{(1, 28)} = 13.18$, $\eta_p^2 = 0.320$, $p < .01$, power = 0.939]; and semantic recall was better in the CP/CS condition than in the CP/IS condition [$F_{(1, 28)} = 9.194$, $\eta_p^2 = 0.247$, $p < .01$, power = 0.833].

Cooperative learning effects—To examine cooperative learning effects, we examined *context consistency effects*—differences in the learning of the same mapping when paired with different (i.e. consistent vs. inconsistent) levels of the context mapping. Thus, to examine cooperative effects of O-S knowledge on O-P learning, we compared the results of the phonological tasks in the CP/CS and CP/IS conditions. Similarly, to examine cooperative effects of O-P knowledge on O-S learning, we compared the results of the semantic tasks in the CP/CS and IP/CS conditions.

Evidence of cooperative learning was found in the judgment task, although this effect was limited to O-P learning. Specifically, as can be seen in Table 1, phonological judgments were more accurate in the CP/CS condition than in the CP/IS condition [$F_{(1, 28)} = 4.886$, $\eta_p^2 = 0.149$, $p < .05$, power = 0.569]. In contrast, cooperative effects were not significant in the semantic judgment task (Table 1) nor in the training (Figure 2) or recall tasks (Table 2).

Discussion

Experiment 1 yielded two key findings. One is that, as predicted, the statistical structure of a to-be-learned mapping matters. Consistent mappings were better learned than inconsistent mappings and, critically, this was true of both O-P and O-S learning. More specifically, consistent mappings gave rise to more accurate responses in both the phonological and semantic versions of the sublexically-focused training task as well as the post-training judgment and recall tasks. It is worth noting that although correct responses in many of these tasks could be based on either lexical or sublexical knowledge, the transfer conditions in the judgment tasks are particularly diagnostic in that above-chance performance necessarily entails the use of sublexical information. Thus, the results from this condition provide clear evidence that participants acquired knowledge of the sublexical regularities embodied in both the O-P and O-S mappings.

The second key finding is that the O-P mapping was learned better in the context of a consistent O-S mapping than in the context of an inconsistent O-S mapping. This finding

provides evidence supporting the cooperative learning hypothesis. However, it is important to note that cooperative effects in Experiment 1 were limited in scope: they were not observed in any of the semantic tasks, nor in the phonological training or recall tasks. Experiment 2 provides further evidence about this pattern. Therefore, we defer consideration of the implications of these results until the General Discussion.

Experiment 2

The results of Experiment 1 demonstrated that participants were sensitive to the statistical structure of the O-P and O-S mappings, as indicated generally by the pervasive effect of consistency across tasks, and most specifically by the high level of performance (for consistent mappings) on untrained transfer items in the judgment tasks. This influence of statistical structure occurred even though participants were not overtly informed about the regularities embodied in the mappings. That is, participants were never explicitly told that a particular radical was associated with a specific semantic category or a specific second syllable. Nonetheless, it is at least possible that participants were cued into these regularities by the demands of the sublexically-focused training tasks, which required judgments about the very properties associated with sublexical components of the written forms.

There is, of course, practical significance to knowing whether explicit instruction is required for beginning readers to discover the regularities embodied by a writing system. Moreover, although the theoretical framework motivating this study does not deny that explicit instruction can alter the course of reading acquisition (see, e.g., Powell, Plaut, & Funnell, 2006), it does posit that (for typical readers, at least) knowledge of sublexical regularities can be acquired on the sole basis of implicit exposure to such regularities (i.e. whole-word training on words embodying them). Thus, in Experiment 2 we changed the training regime by eliminating the training tasks with a sublexical focus.

Methods

Participants—Forty-eight University of Connecticut undergraduate students with normal hearing and normal or corrected-to-normal vision participated in Experiment 2. Fourteen were bilingual speakers with English as their dominant language, but none of either monolingual or bilingual participants spoke Chinese. Participants were randomly assigned to the three orthography conditions, with sixteen in each condition.

Materials—The same materials were used in Experiment 2 as were used in Experiment 1.

Procedure—The training procedure in Experiment 2 differed from that in Experiment 1 in two respects. First, we eliminated the sublexically focused training tasks and included an additional 12 blocks of lexically focused (6 semantic, 6 phonological). Second, we added two more exposure blocks during the training phase and interleaved them every eight blocks within the 24 lexical training blocks, providing more opportunities to learn the associations between spoken forms and their meanings.

Data Analyses—Methods for data analysis were same as in Experiment 1.

Results

Effects of the statistical structure of the to-be-learned mapping—As expected, accuracy increased over blocks in both the phonological [$F_{(11, 308)} = 6.521$, $\eta_p^2 = 0.189$, $p < .001$, power = 1] and the semantic [$F_{(11, 308)} = 29.892$, $\eta_p^2 = 0.516$, $p < .001$, power = 1] training tasks. More importantly, neither the difference between phonological learning in the CP/CS and IP/CS conditions nor semantic learning in the CP/CS and CP/IS conditions was significant. (See Figure 3.) Note that these results parallel those of the first experiment, where we also failed to observe consistency effects in the lexical-focus training tasks.

In contrast, consistency did affect performance in the judgment and recall tasks (Tables 3 & 4). Phonological judgments were more accurate in the CP/CS condition than in the IP/CS condition [$F_{(1, 28)} = 24.226$, $\eta_p^2 = 0.464$, $p < .001$, power = 0.997] and semantic judgments were more accurate in the CP/CS condition than in the CP/IS condition [$F_{(1, 28)} = 22.508$, $\eta_p^2 = 0.446$, $p < .001$, power = 0.996]. A significant word type effect was observed [$F_{(1, 28)} = 6.571$, $\eta_p^2 = 0.190$, $p < .05$, power = 0.697], but only in the semantic judgment task; in the phonological judgment task the word type effect was not significant, although a significant word type $\times$ position interaction effect was observed [$F_{(1, 28)} = 7.044$, $\eta_p^2 = 0.201$, $p < .05$, power = 0.726]. In neither case was the word type $\times$ consistency interaction significant. Similarly, phonological recall was better in the CP/CS condition than in the IP/CS condition [$F_{(1, 28)} = 4.209$, $\eta_p^2 = 0.13$, $p < .05$, power = 0.508]; and semantic learning was better in the CP/CS condition than in the CP/IS condition [$F_{(1, 28)} = 6.054$, $\eta_p^2 = 0.178$, $p < .05$, power = 0.661].

Cooperative learning effects—The results regarding cooperative learning effects were also similar to those of Experiment 1. As can be seen in Table 3, in the phonological judgment task responses were more accurate in the CP/CS condition than in the CP/IS condition [$F_{(1, 28)} = 5.145$, $\eta_p^2 = 0.155$, $p < .05$, power = 0.591]. In contrast (and again paralleling the results of the first experiment), cooperative effects were not significant in the semantic judgment task (Table 3) nor in the training (Figure 3) or recall tasks (Table 4). (One minor difference between the experiments is that in Experiment 2 there was a marginally significant effect of context consistency in the phonological training task: accuracy was numerically greater in the CP/CS condition than in the CP/IS condition [$F_{(1, 28)} = 3.792$, $\eta_p^2 = 0.119$, $p < .10$, power = 0.468]).

Discussion

The results of Experiment 2 replicate the core findings of the previous experiment. First, consistency had parallel effects on O-P and O-S learning. As in Experiment 1, in both the judgment and recall tasks, participants in the present experiment who had been trained on the consistent version of the relevant mapping responded more accurately than participants trained on an inconsistent mapping. Interestingly, no consistency effects were found in the training task, again paralleling the results of Experiment 1. Finally, cooperative learning effects were found in both experiments—in both cases, these effects were limited to the phonological judgment task.

These results provide clear evidence learners become attuned to the statistical regularities embodied in a writing system even in the absence of explicit attention or other task demands that may draw attention to them.

General Discussion

In this study we employed the artificial orthography paradigm to investigate the mechanism underlying learning to read. Our results support three main conclusions. First, the statistical structure of the O-P and O-S mappings determined how easily each of those mappings were learned, suggesting that the learning of both mappings engages a common statistical learning mechanism. Second, our results suggest that learning to read is a cooperative process, in that learning in any particular component of the reading system is influenced by knowledge stored in the rest of the system. Finally, our results demonstrate that knowledge of sublexical regularities can be acquired as the result of exposure to words embodying those regularities.

With regard to the effects of statistical structure, our results are similar to those of other studies that have investigated the impact of O-P or O-S regularities on artificial orthography learning (e.g. Brooks, 1977; Deng, Booth, et al., 2008; Deng, Chou, et al., 2011; Merkx et al., 2011; Rueckl & Dror, 1994). Our results extend these previous findings by demonstrating parallel effects of statistical structure on O-P and O-S learning. In our experiments the structure of the O-P mapping was manipulated in the same way as that of the O-S mapping and, across tasks, the effect of O-P consistency mirrored that of O-S consistency without fail. In particular, both O-P and O-S consistency resulted in better performance in (i) the judgment tasks of Experiment 1 and Experiment 2 (for both trained and transfer items), (ii) the recall tasks of Experiment 1 and Experiment 2, and (iii) the sublexical-focus training task of Experiment 1. Indeed, the one task that failed to reveal an effect of O-P consistency (lexical-focus training in Experiment) also failed to reveal an effect of O-S consistency.

One caveat that should be noted is that performance was in general better on the semantic tasks than on the phonological tasks. It is possible that this pattern reflects differences in the neurocognitive mechanisms responsible for O-P and O-S learning. For example, based on the distinction between declarative and procedural memory (Squire, 1992; Ullman & Pierpont, 2005), it might be speculated that phonological learning is more reliant on procedural learning and cortical-striatal-thalamic circuits, that semantic learning depends more on declarative learning and circuits in the medial temporal lobe (MTL), and that these mechanisms are differentially effective under the circumstances of these experiments. However, there are a variety of other differences that could be relevant to the overall O-S advantage. For example, in the artificial orthographies we constructed the meanings of the words were familiar concepts but the phonological forms, while composed of familiar segments, are novel and sound like non-words for English speakers. Moreover, in contrast to semantics, phonology is inherently a lower-dimension domain, which could have consequences for learning processes driven by statistical regularities. (For example, although it is difficult to know how to quantify this, it is not unreasonable to suppose that compared to the exemplars of a phonological category (defined by a shared second syllable), the exemplars of a given semantic category are both more distinct from one another and yet

more similar to one another than to the exemplars of other categories.) Thus, although further research is needed to fully reveal the extent that O-P and O-S learning rely on the same or different mechanisms, we take the striking pattern of similar consistency effects in the two cases as evidence of a shared mechanism but we do not find the overall O-S advantage as compelling evidence for qualitative differences in underlying mechanisms.

Turning to the cooperative learning hypothesis, in contrast to these parallel results concerning the effects of statistical structure on O-P and O-S learning, evidence concerning cooperative learning differed for the O-P and O-S learning. In both experiments, phonological judgments were more accurate in the context of a consistent O-S mapping than in the context of an inconsistent O-S mapping, indicating that learning of the O-P mapping differed depending on which O-S mapping a participant had been trained on. However, no effect of O-P context was found in any of the semantic tasks. Again, given the data in hand, we cannot rule out the possibility that this asymmetric pattern stems from differences in the neurocognitive mechanisms responsible for the learning of these mappings. That noted, we believe that a more plausible explanation is that our experiments provided a relatively weak test of the cooperative learning hypothesis. In general, cooperative effects are expected to be strongest when (a) the computation of the to-be-learned mapping is relatively slow, (b) the computation of the context mapping is relatively fast, and (c) there are strong associations between spoken forms and their meanings. While condition (a) was met in our experiments, the others were not: The context mapping was learned concurrently with the to-be-learned mapping, there was little statistical structure in the mapping between phonological forms and their meanings, and there was little opportunity to learn associations between phonological forms and their meanings. In retrospect, it is not surprising that few cooperative effects were observed nor is it surprising, given the overall O-S advantage discussed above, that the effects that were observed involved the influence of O-S on O-P learning. To provide a stronger test of the cooperative learning hypothesis, the methodology of the present experiment could be adapted by, for example, pre-training one of the two (O-S or O-P) mappings and imparting more structure in the mapping between phonology and semantics.

With regard to the conditions that give rise to the acquisition of knowledge of sublexical regularities, several previous artificial orthography studies (e.g., Bitan & Karni, 2003, 2004; Yoncheva et al., 2015) have demonstrated that explicit instruction directing a learner's attention to the statistical properties of a to-be-learned mapping facilitate the learning of that mapping. The results of the our first experiment complement these findings—although participants were not given explicit instruction about the regularities embodied by the orthography, our results suggest that the task demands of the sublexical-focus training task enhanced the effect of statistical regularities in the O-P and O-S mapping, as evidenced for example by the difference in performance on the judgment and recall tasks in the two experiments. This being noted, from our perspective the more important finding concerning this issue is that the participants in the second experiment demonstrated sensitivity to sublexical O-P and O-S regularities as a consequence of exposure to words embodying those regularities in the absence of explicit instruction and while performing a task that focused attention on the phonological and semantic properties of whole words. Thus, our results

suggest that while explicit instruction may help a reader become attuned to regularities in the writing system, such instruction is not strictly necessary.

As noted in the Introduction, the hypotheses addressed by the present study were derived from the triangle model (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989) and the results can be taken as support for that framework. More broadly, theories about the mechanisms underlying reading acquisition are relatively few in number, and those that have been articulated are often rather underspecified. Thus, in our view our results are best seen as providing constraints on the development of such theories rather than as evidence adjudicating between theoretical perspectives. In this spirit, it is interesting to consider how our results relate to several prominent theories. For example, we see our findings concerning statistical learning as support for theories of reading acquisition that posit that the organization of reading reflects the statistical structure of the writing system (Frost, 2012; Seidenberg, 2011) and in particular that the ‘units’ of reading are not formed at a privileged grain size but rather are determined by an orthography’s statistical regularities (Treiman & Kessler, 2006; cf. Ziegler & Goswami, 2005). Similarly, we view our evidence of cooperative learning as broadly consistent with elements of the self-teaching hypothesis (Share, 1995, 1999). In its standard form, this theory holds that the generation of a phonological code via phonological decoding provides an anchor of stability that allows for the creation of lexical representations. Cooperative learning can be understood as the generalization of this idea (albeit most readily with a different set of assumptions about the underlying computational mechanisms) such that any source of stability (i.e., semantic as well as phonological codes) facilitates the acquisition of knowledge about other lexical properties. Finally, we see our results as shedding light on the lexical quality hypothesis (Perfetti, 2007) by demonstrating the critical role of statistical regularities in the creation of high-quality representations. An interesting aspect of our results in this respect is the lack of a consistency effect in the lexical-focus training task in Experiment 2. One possible interpretation of this finding is that regardless of the statistical structure of the mapping, this training task engendered the creation of lexical representations of comparable quality (and hence giving rise to equivalent performance in the training task). The difficulty for this interpretation is that it does not account for the results of the judgment task, where a large consistency effect was evident even in the responses for the trained items. Indeed, performance for inconsistent mappings was near chance in this task. What this pattern suggests is that in the absence of statistical regularities, the learning process generates low-quality representations that are brittle and not readily utilized under different task demands.

Admittedly, the foregoing discussion is largely speculative. Theoretical proposals concerning the computational mechanisms underlying reading acquisition are often in the form of verbal descriptions, and to a significant extent advancing our understanding of the processes involved in learning to read will require a more explicit articulation of the nature of these processes. Computational theories described in sufficient detail to so that it can be implemented in computer simulations (e.g., Harm & Seidenberg, 2004; Perry, Ziegler, & Zorzi, 2007; Pritchard, Coltheart, Marinus, & Castles, 2016) exemplify this approach. However, it should be noted that even theories described at this level of detail are likely to incorporate assumptions that would ultimately require further unpacking. For example, learning in the triangle model is dependent on a ‘teaching signal’ that specifies the correct

phonological and semantic representation for each written word. Open questions remain about the source of such information, the impact of missing or unreliable teaching signals, and so forth (see, e.g., Harm & Seidenberg, 2004).

We conclude by noting that in addition to the need for explicit computational theories, progress understanding the mechanisms of reading acquisition also requires an arsenal of methods appropriate for studying these processes. One of these is the artificial orthography paradigm. As noted in the Introduction, although there are clearly important differences between this paradigm and the circumstances associated with learning to read a natural language, there is good reason to believe that this paradigm provides a window on the mechanisms of reading acquisition. In the present study we introduced a novel way to independently manipulate the structure of the O-P and O-S mappings in an artificial orthography task. Our results suggest that the same process attunes readers to the statistical regularities in both mappings, and they also provide some that learning is cooperative in that learning in any particular component of the reading system is influenced by knowledge stored in the rest of the system.

Acknowledgements

This study was supported by the National Institute of Child Health and Human Development grants P01 HD001994 (PI: Jay Rueckl), P01 HD070837 (PI: Robin Morris), and P20 HD091013 (PI: Donald Compton). The study was also supported by a Graduate School Doctoral Dissertation Fellowship and a Graduate Summer Research Fellowship from the University of Connecticut to Jingjing Zhao, and by the Hundred Talents Program of the Shaanxi Province (Long-term innovation project, SXBR9019) to Jingjing Zhao, the Youth Fund for Humanities and Social Sciences Research of the Ministry of Education (17XJC190010, PI: Jingjing Zhao), and the Fundamental Research Funds for the Central Universities (GK201702011, PI: Jingjing Zhao). The protocol of the study was approved by the institutional review board of the University of Connecticut.

References

- Arciuli J, & Simpson IC (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science*, 36(2), 286–304. [PubMed: 21974775]
- Backman J, Bruck M, Hebert M, & Seidenberg M (1984). Acquisition and use of spelling-sound correspondences in reading. *Journal of Experimental Child Psychology*, 38(1), 114–133. doi: 10.1016/0022-0965(84)90022-5
- Battig WF, & Montague WE (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, 80(3, Pt. 2), 1–46. doi: 10.1037/h0027577
- Bitan T, & Karni A (2003). Alphabetical knowledge from whole words training: effects of explicit instruction and implicit experience on learning script segmentation. *Cognitive Brain Research*, 16(3), 323–337. doi: 10.1016/S0926-6410(02)00301-4 [PubMed: 12706213]
- Bitan T, & Karni A (2004). Procedural and declarative knowledge of word recognition and letter decoding in reading an artificial script. *Cognitive Brain Research*, 19(3), 229–243. doi: 10.1016/j.cogbrainres.2004.01.001 [PubMed: 15062861]
- Bogaerts L, Szmalec A, De Maeyer M, Page MP, & Duyck W (2016). The involvement of long-term serial-order memory in reading development: a longitudinal study. *Journal of Experimental Child Psychology*, 145, 139–156. doi: 10.1016/j.jecp.2015.12.008 [PubMed: 26835842]
- Brooks LR (1977). Visual pattern in fluent word identification. In Reber AS, & Scarborough DL (Eds.), *Toward a Psychology of Reading: The Proceedings of the CUNY Conference* (pp. 143–181). Hillsdale, NJ: Erlbaum.

- Brooks LR, & Miller A (1979). A comparison of explicit and implicit knowledge of an alphabet In Kolers PA, Wrolstad ME, & Bouma H (Eds.), *Processing of Visible Language*, Vol. 1 (pp. 391–401). Plenum, NY.
- Burani C, & Caramazza A (1987). Representation and processing of derived words. *Language and Cognitive Processes*, 2(3–4), 217–227. doi: 10.1080/01690968708406932
- Byrne B (1984). On teaching articulatory phonetics via an orthography. *Memory & Cognition*, 12(2), 181–189. doi: 10.3758/BF03198432 [PubMed: 6203008]
- Connor CM, Morrison FJ, Fishman B, Crowe EC, Al Otaiba S, & Schatschneider C (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science*, 24(8), 1408–1419. doi: 10.1177/0956797612472204 [PubMed: 23785038]
- Deacon SH, & Leung D (2013). Testing the statistical learning of spelling patterns by manipulating semantic and orthographic frequency. *Applied Psycholinguistics*, 34(6), 1093–1108. doi: 10.1017/S0142716412000173
- Deng Y, Booth JR, Chou TL, Ding GS, & Peng DL (2008). Item-specific and generalization effects on brain activation when learning Chinese characters. *Neuropsychologia*, 46(7), 1864–1876. doi: 10.1016/j.neuropsychologia.2007.09.010 [PubMed: 18514678]
- Deng Y, Chou TL, Ding GS, Peng DL, & Booth JR (2011). The involvement of occipital and inferior frontal cortex in the phonological learning of Chinese characters. *Journal Cognitive Neuroscience*, 23(8), 1998–2012. doi:10.1162/jocn.2010.21571
- Feldman LB (1995). *Morphological Aspects of Language Processing*. Hillsdale, NJ: Erlbaum.
- Fowler CA, Napps SE, & Feldman L (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory & Cognition*, 13(3), 241–255. doi: 10.3758/BF03197687 [PubMed: 4046825]
- Frost R (2012). A universal approach to modeling visual word recognition and reading: Not only possible, but also inevitable. *Behavioral and Brain Sciences*, 35(5), 310–329. doi: 10.1017/S0140525X12000635 [PubMed: 23251930]
- Frost R, Siegelman N, Narkiss A, & Afek L (2013). What predicts successful literacy acquisition in a second language?. *Psychological Science*, 24(7), 1243–1252. [PubMed: 23698615]
- Glushko RJ (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5(4), 674–691. doi: 10.1037/0096-1523.5.4.674
- Goswami U, Gombert JE, & de Barrera LF (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics*, 19(1), 19–52. doi: 10.1017/S0142716400010560
- Harm MW, McCandliss BD, & Seidenberg MS (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, 7(2), 155–182. doi: 10.1207/S1532799XSSR0702_3
- Harm MW, & Seidenberg MS (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, 106(3), 491–528. doi: 10.1037/0033-295x.106.3.491 [PubMed: 10467896]
- Harm MW, & Seidenberg MS (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review*, 111(3), 662–720. doi: 10.1037/0033-295X.111.3.662 [PubMed: 15250780]
- Havas V, Laine M, & Fornells AR (2017). Brain signatures of early lexical and morphological learning of a new language. *Neuropsychologia*, 101, 47–56. doi: 10.1016/j.neuropsychologia.2017.04.005 [PubMed: 28411058]
- Havas V, Waris O, Vaquero L, Rodríguez-Fornells A, & Laine M (2015). Morphological learning in a novel language: a cross-language comparison. *Quarterly Journal of Experimental Psychology*, 68(7), 1426–1441. doi: 10.1080/17470218.2014.983531
- Ho CSH, Ng TT, & Ng WK (2003). A “radical” approach to reading development in Chinese: The role of semantic radicals and phonetic radicals. *Journal of Literacy Research*, 35(3), 849–878. doi: 10.1207/s15548430jlr3503_3

- Jared D (2002). Spelling-sound consistency and consistency effects in word naming. *Journal of Memory and Language*, 46(4), 723–750. doi: 10.1006/Jmla.2001.2827
- Lee CY, Tsai JL, Kuo WJ, Yeh TC, Wu YT, Ho LT, Hung DL, Tzeng OJ-L, & Hsieh JC (2004). Neuronal correlates of consistency and frequency effects on Chinese character naming: an event-related fMRI study. *Neuroimage*, 23(4), 1235–1245. doi: 10.1016/j.neuroimage.2004.07.064 [PubMed: 15589089]
- Linkersdörfer J, Jurcoane A, Lindberg S, Kaiser J, Hasselhorn M, Fiebach CJ, & Lonnemann J (2015). The association between gray matter volume and reading proficiency: a longitudinal study of beginning readers. *Journal of Cognitive Neuroscience*, 27(2), 308–318. doi: 10.1162/jocn_a_00710 [PubMed: 25203270]
- Marslen-Wilson W, Tyler LK, Waksler R, & Older L (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101(1), 3–33. doi: 10.1037/0033-295X.101.1.3
- Merkx M, Rastle K, & Davis MH (2011). The acquisition of morphological knowledge investigated through artificial language learning. *The Quarterly Journal of Experimental Psychology*, 64(6), 1200–1220. doi: 10.1080/17470218.2010.538211 [PubMed: 21271463]
- Moore MW, Brendel PC, & Fiez JA (2014). Reading faces: Investigating the use of a novel face-based orthography in acquired alexia. *Brain and Language*, 129, 7–13. doi: 10.1016/j.bandl.2013.11.005 [PubMed: 24463310]
- Pan J, Song S, Su M, McBride C, Liu H, Zhang Y, Li H, & Shu H (2016). On the relationship between phonological awareness, morphological awareness and Chinese literacy skills: evidence from an 8-year longitudinal study. *Developmental Science*, 19(6), 982–991. doi: 10.1111/desc.12356 [PubMed: 26537834]
- Perfetti C (2007). Reading ability: lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383.
- Perry C, Ziegler JC, & Zorzi M (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114(2), 273–315. doi: 10.1037/0033-295X.114.2.273 [PubMed: 17500628]
- Plaut DC, & Gonnerman LM (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, 15(4–5), 445–485. doi: 10.1080/01690960050119661
- Plaut DC, McClelland JL, Seidenberg MS, & Patterson K (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103(1), 56–115. doi: 10.1037/0033-295x.103.1.56 [PubMed: 8650300]
- Powell D, Plaut D, & Funnell E (2006). Does the PMSP connectionist model of single word reading learn to read in the same way as a child? *Journal of Research in Reading*, 29 (2), 229–250. doi: 10.1111/j.1467-9817.2006.00300.x
- Pritchard SC, Coltheart M, Marinus E, & Castles A (2016). Modelling the implicit learning of phonological decoding from training on whole-word spellings and pronunciations. *Scientific Studies of Reading*, 20(1), 49–63. doi: 10.1080/10888438.2015.1085384
- Pugh KR, Frost SJ, Sandak R, Landi N, Rueckl JG, Constable RT, Seidenberg MS, Fulbright RK, Katz L, & Mencl WE (2008). Effects of stimulus difficulty and repetition on printed word identification: An fMRI comparison of nonimpaired and reading-disabled adolescent cohorts. *Journal of Cognitive Neuroscience*, 20(7), 1146–1160. doi: 10.1162/jocn.2008.20079 [PubMed: 18284344]
- Ricketts J, Davies R, Masterson J, Stuart M, & Duff FJ (2016). Evidence for semantic involvement in regular and exception word reading in emergent readers of English. *Journal of experimental child psychology*, 150, 330–345. doi: 10.1016/j.jecp.2016.05.013 [PubMed: 27416563]
- Rueckl JG, & Dror IE (1994). The effect of orthographic-semantic consistency on the acquisition of new words In Umiltà C & Moscovitch M (Eds.) *Attention and Performance*, XV (pp. 571–588). Hillsdale, NJ: Erlbaum.
- Rueckl JG, & Raveh M (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language*, 68(1–2), 110–117. doi: 10.1006/Brln.1999.2106 [PubMed: 10433747]
- Sandak R, Mencl WE, Frost SJ, Rueckl JG, Katz L, Moore DL, Mason SA, Fulbright RK, Constable RT, & Pugh KR (2004). The neurobiology of adaptive learning in reading: A contrast of different

training conditions. *Cognitive, Affective, & Behavioral Neuroscience*, 4(1), 67–88. doi: 10.3758/CABN.4.1.67

- Seidenberg MS (2011). Reading in different writing systems: One architecture, multiple solutions In McCardle P, Miller B, Lee J & Tzeng O (Eds.), *Dyslexia Across Languages: Orthography and the Brain-Gene-Behavior Link* (pp. 151–174). Baltimore, MD: Paul H. Brookes Publishing.
- Seidenberg MS, & McClelland JL (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523–568. doi: 10.1037/0033-295X.96.4.523 [PubMed: 2798649]
- Seymour PH, Aro M, & Erskine JM (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174. doi: 10.1348/000712603321661859 [PubMed: 12803812]
- Share DL (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55(2), 151–218. doi: 10.1016/0010-0277(94)00645-2 [PubMed: 7789090]
- Share DL (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72(2), 95–129. doi: 10.1006/jecp.1998.2481 [PubMed: 9927525]
- Shu H, Anderson RC, & Wu N (2000). Phonetic awareness: Knowledge of orthography–phonology relationships in the character acquisition of Chinese children. *Journal of Educational Psychology*, 92(1), 56–62. doi: 10.1037/0022-0663.92.1.56
- Shu H, Chen X, Anderson RC, Wu N, & Xuan Y (2003). Properties of school Chinese: implications for learning to read. *Child Development*, 74(1), 27–47. doi: 10.1111/1467-8624.00519 [PubMed: 12625434]
- Snodgrass JG, & Vanderwart M (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174–215. doi: 10.1037/0278-7393.6.2.174 [PubMed: 7373248]
- Squire LR (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, 4(3), 232–243. doi: 10.1162/jocn.1992.4.3.232 [PubMed: 23964880]
- Strain E, Patterson K, & Seidenberg MS (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Human Learning and Memory*, 21(5), 1140–1154. doi: 10.1037/0278-7393.34.2.381
- Taft M, & Forster KI (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 638–647. doi: 10.1016/S0022-5371(75)80051-X
- Taylor JS, Davis MH, & Rastle K (2017). Comparing and validating methods of reading instruction using behavioural and neural findings in an artificial orthography. *Journal of Experimental Psychology: General*, 146(6), 826–858. doi: 10.1037/xge0000301 [PubMed: 28425742]
- Taylor JS, Plunkett K, & Nation K (2011). The influence of consistency, frequency, and semantics on learning to read: an artificial orthography paradigm. *Journal of Experimental Psychology: Human Learning and Memory*, 37(1), 60–76. doi: 10.1037/a0020126
- Treiman R, & Kessler B (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology*, 98(3), 642–652. doi: 10.1037/0022-0663.98.3.642
- Trudeau JJ (2006). Semantic contributions to word naming with artificial lexicons. *Dissertations Collection for University of Connecticut*. Paper AAI3236153.
- Ullman MT, & Pierpont EI (2005). Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex*, 41(3), 399–433. doi: 10.1016/S0010-9452(08)70276-4 [PubMed: 15871604]
- Wang HC, Nickels L, Nation K, & Castles A (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading*, 17(5), 369–384. doi: 10.1080/10888438.2012.749879
- Xue G, Chen C, Jin Z, & Dong Q (2006). Cerebral asymmetry in the fusiform areas predicted the efficiency of learning a new writing system. *Journal of Cognitive Neuroscience*, 18(6), 923–931. doi: 10.1162/jocn.2006.18.6.923 [PubMed: 16839300]

- Yang J, Shu H, McCandliss BD, & Zevin JD (2009). Simulating language-specific and language-general effects in a statistical learning model of Chinese reading. *Journal of Memory and Language*, 61(2), 238–257. doi:10.1016/j.jml.2009.05.001 [PubMed: 20161189]
- Yang J, Shu H, McCandliss BD, & Zevin JD (2013). Orthographic influences on division of labor in learning to read Chinese and English: Insights from computational modeling. *Bilingualism: Language and Cognition*, 16(2), 354–366. doi: 10.1017/S1366728912000296
- Yeatman JD, Dougherty RF, Ben-Shachar M, & Wandell BA (2012). Development of white matter and reading skills. *Proceedings of the National Academy of Sciences*, 109(44), E3045–E3053. doi: 10.1073/pnas.1206792109
- Yoncheva YN, Wise J, & McCandliss B (2015). Hemispheric specialization for visual words is shaped by attention to sublexical units during initial learning. *Brain and Language*, 145, 23–33. doi: 10.1016/j.bandl.2015.04.001 [PubMed: 25935827]
- Ziegler JC, & Goswami U (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131(1), 3–29. doi: 10.1037/0033-2909.131.1.3 [PubMed: 15631549]

(a) Consistent Mapping

Inconsistent Mapping		士	中	下	刀	山
	士	dog	bed	pear	shoe	nose
	中	cat	desk	banana	coat	leg
	下	horse	lamp	peach	shirt	arm
	刀	cow	chair	grapes	sock	eye
	山	lion	table	apple	pants	foot

(b) Consistent Mapping

Inconsistent Mapping		小	大	儿	广	父
	小	/faIbi/	/waIleI/	/hifu/	/keIgΔ/	/raIdo/
	大	/gUbi/	/geleI/	/keIfu/	/naUgΔ/	/bædo/
	儿	/næbi/	/poleI/	/lafu/	/faIgΔ/	/feIdo/
	广	/sΔbi/	/tuleI/	/waIfu/	/higΔ/	/gUdo/
	父	/pobi/	/sΔleI/	/bæfu/	/nægΔ/	/naUdo/

Figure 1.
 (a) An illustration of how the structure of the orthographic-semantic (O-S) mapping was manipulated. Each semantic radical was embedded in 5 words. Each column lists five exemplars of a semantic category. Thus, pairing the words containing a particular radical with the meanings listed in its respective column results in a consistent O-S mapping. In contrast, pairing forms and meaning by row results in an inconsistent mapping. (b) Orthographic-phonological (O-P) mappings were constructed in the same way. Phonological categories (defined by the identity of the second syllable) are organized by columns. Pairing

the words containing a particular radical with the pronunciations listed in its respective column results in a consistent O-P mapping. In contrast, pairing forms and meaning by row results in a inconsistent mapping.

Figure 2. Consistency effects in the sublexical and lexical training tasks in Experiment 1. Error bars represent standard errors (SES).

Figure 3.
Consistency effects in the lexical training tasks in Experiment 2. Error bars represent SES.

Table 1.

Mean accuracy (standard deviation) of the judgment tasks of Experiment 1.

	CP/CS		CP/IS		IP/CS	
	Trained	Transfer	Trained	Transfer	Trained	Transfer
Phonological	0.93 (0.14)	0.89 (0.16)	0.77 (0.25)	0.79 (0.23)	0.52 (0.08)	0.50 (0.08)
Semantic	0.93 (0.12)	0.90 (0.14)	0.57 (0.09)	0.45 (0.07)	0.89 (0.14)	0.86 (0.20)

Note : CP/CS = Consistent Phonology/Consistent Semantics. CP/IS = Consistent Phonology/Inconsistent Semantics. IP/CS = Inconsistent Phonology/Consistent Semantics.

Table 2.

Mean accuracy (standard deviation) of the recall tasks of Experiment 1.

	CP/CS	CP/IS	IP/CS
Phonological	0.11 (0.08)	0.08 (0.08)	0.02 (0.03)
Semantic	0.44 (0.28)	0.21 (0.11)	0.40 (0.27)

Note : CP/CS = Consistent Phonology/Consistent Semantics. CP/IS = Consistent Phonology/Inconsistent Semantics. IP/CS = Inconsistent Phonology/Consistent Semantics.

Table 3.

Mean accuracy (standard deviation) of the judgment tasks of Experiment 2.

	CP/CS		CP/IS		IP/CS	
	Trained	Transfer	Trained	Transfer	Trained	Transfer
Phonological	0.67 (0.18)	0.68 (0.18)	0.56 (0.17)	0.56 (0.13)	0.47 (0.12)	0.50 (0.11)
Semantic	0.76 (0.17)	0.66 (0.19)	0.51 (0.09)	0.48 (0.05)	0.79 (0.15)	0.77 (0.16)

Note : CP/CS = Consistent Phonology/Consistent Semantics. CP/IS = Consistent Phonology/Inconsistent Semantics. IP/CS = Inconsistent Phonology/Consistent Semantics.

Table 4.

Mean accuracy (standard deviation) of the recall tasks of Experiment 2.

	CP/CS	CP/IS	IP/CS
Phonological	0.08 (0.08)	0.07 (0.07)	0.03 (0.04)
Semantic	0.35 (0.22)	0.20 (0.12)	0.28 (0.25)

Note : CP/CS = Consistent Phonology/Consistent Semantics. CP/IS = Consistent Phonology/Inconsistent Semantics. IP/CS = Inconsistent Phonology/Consistent Semantics.