Different processes for reading words learned before and after onset of literacy

Padraic Monaghan (p.monaghan@lancaster.ac.uk)
Ya-Ning Chang (y.chang5@lancaster.ac.uk)
Department of Psychology, Lancaster University, Lancaster, LA1 4YF, UK
Stephen Welbourne (stephen.welbourne@googlemail.com)
Neuroscience and Aphasia Research Unit (NARU), University of Manchester, Manchester, M13 9PL, UK

Abstract

Learning to read has a substantial effect on the representations of spoken and meaning forms of words. In this paper we assess literacy effects beyond representational changes, focusing on adaptations to the architecture of the reading system that maps between these representations. We present a connectionist model of reading that predicted distinct processing of pre- and post-literacy acquired words. For reading for meaning, words learned prior to literacy were processed more indirectly via phonological representations, whereas for post-literacy acquired words, processing was more direct along the orthography to semantics pathway. This more computationally intensive route was prioritised because indirect phonology to semantics mappings were unavailable. Such an effect was less apparent for naming, because learning direct orthography to phonology mappings is less computationally intensive. These results were confirmed in an analysis of naming and lexical decision behavioural data. The effect of literacy onset remains an observable artefact in adult reading.

Keywords: literacy; age of acquisition; language development; reading fluency; reading comprehension; computational modelling.

Effects of literacy on reading

There are multiple influences on readers' speed and accuracy of reading, and these have been extensively documented in the literature over the last 50 years of reading research. For instance, higher-frequency words tend to be accessed more quickly and accurately than lower-frequency words, and early-acquired words tend to be responded to faster and more accurately than later-acquired words, referred to as an "age of acquisition" (AoA) effect (Brysbaert, & Ghyselinck, 2006; Cortese & Khanna, 2007; Juhasz, 2005; Monaghan & Ellis, 2002).

Theories of the origin of the AoA effect on reading are two-fold. One view is that early acquired words result in prioritised lexical semantic representations, because they enter first of all into the lexical semantic associative network, and subsequently learned words are then connected to previously acquired words (Brysbaert & Ghyselinck, 2006). Analyses of semantic associations by Steyvers and Tenenbaum (2005) confirmed that early acquired words do have more words associated with them than later acquired words, and they demonstrated that small-scale illustrative versions of this growing semantic associative network could prioritise early acquired words in semantic processing.

An alternative perspective is that AoA effects are instead found in the mappings between representations, rather than the representations themselves (Monaghan & Ellis, 2010). Early acquired words are learned when the neural network supporting the mappings among print, sound and meaning is plastic and able to acquire mappings effectively. Mappings for later acquired words are required to fit around the previously learned mappings, when the neural network has lower plasticity, resulting in prioritisation for early over later acquired words. Such AoA effects are predicted to be greater for arbitrary mappings, such as between meaning and sound, rather than for (quasi-)regular mappings such as between print and sound, because learning arbitrary mappings is more computationally intensive and therefore affected more by reduced plasticity (Lambon Ralph & Ehsan, 2006). However, AoA effects ought still to be observed even for regular mappings because of the smaller, but still present, effect of reducing plasticity in learning the mappings.

These predictions have been supported by meta-analyses of behavioural studies (Brysbaert & Ghyselinck, 2006) which have investigated AoA effects for naming and for lexical decision. It is generally assumed that for naming, semantic representations of words are minimally involved in producing the phonological form of a word from its orthographic form (Harm & Seidenberg, 1999). However, lexical decision appears to implicate semantic representations to a greater degree (Chang et al., 2016; Plaut, 1997), in that semantic properties of words, such as imageability or concreteness, account for more variance in lexical decision or picture naming responses and little for written word naming (Balota et al., 2004; Catling & Johnston, 2009). Brysbaert and Ghyselinck (2006) showed that AoA effects were much greater for tasks involving semantics, including lexical decision, than for tasks involving production of phonology (see also Cortese & Khanna, 2007). However, the fact that AoA does still account for some variance in naming indicates the effects of plasticity in the quasi-regular print to sound mapping for English (see Lambon Ralph and Ellis, 2000, and Monaghan and Ellis, 2010, for computational illustrations of this).

Conversely, the size of the AoA effect can be used to indicate the extent to which the pathways to and from lexical semantics in the reading system are involved in reading. If the AoA effect is large, then semantics is likely to be involved, if the effect is small then semantics is less likely to be involved. Chang et al. (2016) implemented a triangle computational model of single word reading, and varied the point at which words were presented to the model, to simulate different AoA of words. For simulations of naming, AoA had a significant effect, but for simulations of lexical decision, AoA accounted for a substantially larger proportion of variance.

One absence from these theoretical and implemented models of reading, however, is the role not only of AoA but also of different modes by which words are acquired. Literacy is known to have profound effects on language processing, resulting in changes to phonological awareness (Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012; Morais, Cary, Alegria & Bertelson, 1979), changes to phonological processing of words (Smith, Monaghan, & Huettig, 2014), as well as semantic fluency (Kosmides, Tsapkini, Folia, Vlahou, & Kiosseoglou, 2004), and even visual processing (Szwed, Ventura, Querido, Cohen, & Dehaene, 2012).

However, less studied are the potential effects of literacy on the architecture of the reading system in terms of pathways employed between different representations of words. Prior to literacy, the learner acquires mappings between sound and meaning representations of words, through listening and comprehending words, and speaking words for others' comprehension. However, once the child begins to learn to read for these already known words, mappings will be generated from print to the stored sound and meaning representations. But for new words, the print form will be mapped onto newly acquired sound and meaning representations, where the mappings between sound and meaning are not available in advance.

In terms of the operation of the reading system, this difference between pre-literacy and post-literacy acquired words is likely to be profound. In the triangle model of reading (Seidenberg & McClelland, 1989) there are two routes by which a printed word can be pronounced. This can occur directly, through learned mappings between print and sound, or indirectly from print via semantics to sound (see Figure 1). Similarly, for reading comprehension, the mapping from print can be directly to meaning, or indirectly, from print to sound to meaning. For pre-literacy acquired words, the indirect route is more likely to be available, because the sound to meaning routes are already acquired, whereas for post-literacy words, the indirect route requires two mappings to be acquired.

Furthermore, the properties of the mappings from print to sound and meaning will also contribute to the extent to which direct and indirect mappings are utilised. Regular mappings, such as between print and sound in English, are easier to acquire than arbitrary mappings, such as between print and meaning. Thus, the direct route is more likely to be prioritised for print to sound mappings than the indirect route, and the indirect route is more likely to be prioritised for print to meaning mappings than the direct route, because the indirect route is more easily acquired, at least for words acquired pre-literacy, where the sound to meaning mapping is already in place in the language processing system.

Based on this theory, we predict that there is likely to be a distinction between pre-literacy and post-literacy processing of words' print to meaning mappings, as in lexical decision. Pre-literacy, the indirect route is more likely to have a greater influence on processing. Post-literacy, the direct route is likely to have a greater influence. Whereas for print to sound mappings, as in word naming, we predict no difference between pre- and post-literacy processing, because both will be mapped via fast-acquired direct print to sound mappings, which will have an equal influence on reading.

In this paper, we first provide a computational test of the extent to which the triangle model of reading predicts different processing routes for words pre- and post-literacy. We then test whether the predictions of the model are observed in behavioural data on word naming and lexical decision response times. For both the simulation and the behavioural data, we use the size of the AoA effect as an index of the extent to which direct mappings from orthography to semantics are implicated in the reading system. For naming, a larger AoA effect indicates greater use of indirect mappings via semantics for reading tasks, for lexical decision a larger AoA effect indicates greater use of direct mappings from orthography, where arbitrary mappings between orthography and semantics are implicated. A smaller AoA effect for lexical decision indicates that the indirect quasi-regular mapping from orthography to phonology is being prioritised. It is the case that mappings between phonology and semantics are also arbitrary, but these mappings would exert a smaller AoA effect than that observed for the newly acquired mappings because they are intensively trained, and acquired earlier in acquisition, thus reducing distinctions between words due to greater plasticity of resources for early-learned mappings (see e.g., Ellis & Lambon Ralph, 2000; Monaghan, Chang, Welbourne, & Brysbaert, 2017).

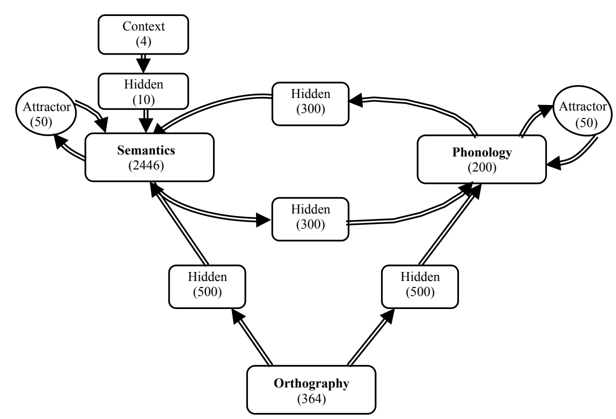

Figure 1. The architecture of the triangle model of reading.

A computational model of literacy effects on reading processes

Method

Network Architecture

The model is based on the connectionist triangle model of reading (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989), as shown in Figure 1. The critical property of the model is that it incorporates three representations of single words – print, sound, and meaning. Each of these representations is interconnected by sets of hidden units that permit the mappings between representations to be acquired as a consequence of exposure. The sound and meaning layers were also each connected to a set of attractor units that enable the model to develop high-fidelity phonological and semantic representations of words.

Also included was a context layer to enable disambiguation of the meaning of homonyms (e.g., /beIs/ as the instrument *bass* and as the location *base,* see Chang et al., 2016, for more details).

Representations

The representations of words were derived from Harm and Seidenberg's (2004) version of the triangle model. Printed words were represented across 14 letter slots, with each letter slot comprising 26 units relating to one letter of the alphabet. If a letter was present in a slot, then the unit corresponding to the letter had activity 1, otherwise the units were inactive. Spoken words were represented in terms of segmental phonological features, across 8 phoneme slots of 25 binary phonological feature units, with distinct phonemes represented in terms of overlapping subsets of the units representing the features. Finally, lexical meaning representations were constructed from semantic features in WordNet (Miller, 1990). Each word activated a subset of the 2446 semantic features in the semantic layer of the model, with activity 1 if the semantic feature was associated with the word.

The model was eventually trained to read 6229 monosyllabic words, which were presented during reading training according to log-compressed frequency, where frequency was taken from the Wall Street Journal corpus (Marcus, Santorini, & Marcinkiewicz, 1993), to be consistent with Harm and Seidenberg's (2004) implementation of the triangle model.

Training Procedures

The training process had two phases: a pre-literacy and a post-literacy phase. In the pre-literacy training, the model learned to map between phonology and semantics on a subset of words from the entire training set, that children are more likely to have learned before beginning reading. In the post-literacy phase, the model was trained to learn to map all words from orthographic forms onto phonology and semantics.

In pre-literacy training, the model was trained on oral language tasks, including a speaking task (mapping from semantic to phonological representations), a hearing task (mapping from phonological to semantic representations), as well as tasks that assisted in developing stable attractors at phonology and semantics (mapping from phonological to phonological representations, and from semantic to semantic representations). For the speaking task, the semantic input pattern for a selected word was clamped for eight time steps, then in the last two time steps, the model was required to reproduce the phonological form for the word. The difference between the model's actual production and the target phonological production was backpropagated through the network and connections were adjusted to reduce error. Similarly, for the hearing task, the phonological input and the context were clamped for 8 time steps, and the model was required to produce the target semantic form at the output. For the stable attractor tasks, the input was presented then activation cycled for 6 time steps, before the model was required to reproduce the originally inputted phonological or semantic representation. For pre-literacy training, the four tasks were interleaved, with 40% of trials each for the speaking and hearing tasks, and 10% each for the phonological and semantic attractor trials. There were 600,000 trials altogether.

For pre-literacy training, the model was exposed to 2,973 monosyllabic words, which were selected to be the most common words occurring in reading materials before age 18, and therefore those words most likely that children come across prior to literacy onset. Words were presented randomly, but selected according to their frequency. The model was trained with a learning rate of 0.05 using back-propagation through time, and cross-entropy error was computed. No adjustments to weights were made if the model was within 0.1 of the target for each output unit.

In the post-literacy training, the model was given printed word forms, and required to learn to map onto phonological and semantic representations. Words were presented to the model incrementally, according to the reading-age at which words occurred. Similar to Monaghan and Ellis (2010), reading developed cumulatively, over 14 reading stages reflecting reading materials experienced from age 5 to 18, determined from the educator's word frequency guide (Zeno et al., 1995), see Chang et al. (2016) for more details.

For each word, the model cycled for 12 time steps of activation after which the model had to generate the phonological and semantic representations of the word. These reading trials were interleaved with hearing and speaking trials, and phonological and semantic attractor trials, to ensure that the pre-literacy mappings were maintained during reading training. There were 1.74 million post-literacy training trials altogether.

Critically, by the end of training, the model had been exposed to all words, but some of these had been acquired prior to literacy onset, and others were acquired from print. We refer to these words as pre- and post-literacy words.

Testing Procedures

To measure pre-literacy oral language skills, the model was tested on its productions for the speaking and hearing tasks. For semantics, if the model was closer to the target word than any other word, then it was judged to be accurate. For phonology, if the model was closer to the target phoneme at each phoneme slot then it was judged to be correct.

For the analysis of reading performance, we interpreted orthographic to phonological representations to be analogous to behavioural naming responses (Chang, Furber, & Welbourne, 2012), and orthographic to semantic mappings to relate to lexical decision responses (see, e.g., polarity measure in Plaut, 1997, and Chang et al., 2016).

Results

At the end of pre-literacy training, of the words to which the model had been exposed prior to onset of literacy, the model was able to speak 90.7%, and comprehend 91.7% correctly. After reading training, the model was accurate for 99.4% of phonology and 93.3% of semantics for the reading task.

To assess whether literacy changed patterns of processing in the model, multiple regression analyses were conducted for the model's simulations of word naming and lexical decision tasks. The mean square error of the model's productions was taken as the dependent variable, and a set of psycholinguistic variables were included as predictors, to relate to previous regression analyses of behavioural data (e.g., Balota et al., 2004; Cortese & Khanna, 2007). These variables were cumulative frequency (CF), orthographic neighbourhood size (OrthN) (Coltheart, 1977), word length (Len), consistency (Cons) (which was the proportion of words with the same pronunciation of the orthographic rime, e.g., "gave/save" versus "have"), and AoA, which was the reading stage during training for the model. Error scores were log transformed and all the predictor variables were centred.

To examine the effect of literacy onset on the model's performance, hierarchical regression analyses were conducted. At step 1, all psycholinguistic variables were entered, then at step 2 whether the word appeared pre- or post-literacy was entered as a variable interacting with AoA. If processing changes from pre- to post-literacy, then the effect of AoA at the point of literacy onset should change, as an index of the involvement of semantics – reflected in a significant interaction. It was not possible to include literacy onset as a separate variable because it is highly correlated with the interaction term. The results for naming and lexical decision are shown in Table 1.

Table 1. Results from the regression analysis for naming and for lexical decision in the computational model.

		Naming	Lexical Decision
		β	β
Step 1	CF	-0.179***	-0.107***
	OrthN	-0.256**	0.012
	Cons	-0.247***	-0.016
	Len	-0.071***	-0.127***
	AoA	0.198***	0.452***
Step 2	AoA x Literacy onset	0.219*** ΔR^2 = 0.37%	0.501*** ΔR^2 = 1.96%

***p<.001; **p<.01; β is a standardized beta value.

Literacy onset was a significant predictor of changes in the model's performance – at the point of literacy onset, the regression gradient for the AoA effect changed, such that words acquired pre-literacy demonstrated a smaller change in response times associated with increasing AoA compared to words acquired post-literacy. This effect was substantially larger for lexical decision than for naming responses, suggesting that processing for pre-literacy acquired words used the indirect route from orthography to semantics via phonology, whereas the post-literacy acquired words used the direct orthography to semantics route.

We next tested whether a similar change in processing was associated with literacy onset in naming and lexical decision behaviour.

Testing the literacy effect in word processing

Method

The data were a subset of responses from the English Lexicon Project (Balota et al., 2007), comprising naming and lexical decision response times from a set of 816 young adult participants from a range of universities. We acquired data for 2,536 monosyllabic words, for which all the psycholinguistic variables could be generated.

Word-form frequency, orthographic neighbourhood size, and word length were taken from the CELEX database (Baayen, Pipenbrock, & Gulikers, 2005). These three measures were taken from the same dataset to ensure consistency across these measures. AoA was taken from Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012). Consistency of words was determined in the same way as for the computational simulation.

Results

We first aimed to replicate the results of Balota et al. (2004, 2007) in determining the role of frequency, word length, neighbourhood size, consistency, and AoA in a linear regression on naming response times and lexical decision response times.

Then, we measured whether there was an effect of onset of literacy in the behavioural data through adding an interaction between AoA and literacy onset. Age of literacy onset could not be included *a priori* as with the simulation, however, we assumed that if there is an effect of onset of literacy, then this should occur somewhere close to the age of 5. Onset of literacy was thus determined iteratively between the age of 3, 4, 5, 6, and 7 years in order to assess whether there is a discontinuity in response times predicted by AoA that changes around the age children begin formal literacy. We took as an indicator of discontinuity a significant interaction between AoA and literacy onset, though see Baayen, Feldman, and Schreuder (2006) for an alternative means of measuring discontinuities (note they were unable to test AoA because of small sample size).

For naming and lexical decision response times, the results of the multiple regression are shown in Table 2. For naming, adding the interaction between onset of literacy and

AoA for any of the ages 3 to 7 did not significantly improve the model fit (Bonferroni corrected).

For lexical decision response times, there were significant effects of literacy onset found at ages 5, 6, and 7, with the largest effect for age 6. Figure 2 shows the effect of this discontinuity in predicting response times for lexical decision when the onset of literacy is implemented at age 6. The same Figure illustrates no statistically significant discontinuity effect for naming response times.

Table 2. Results from the regression analysis for naming and for lexical decision in the behavioural data.

		Naming	Lexical Decision
		β	β
Step 1	Log-frequency	-0.156***	-0.305***
	OrthN	-0.255***	-0.001
	Cons	-0.115***	-0.032*
	Len	0.165***	-0.062***
	AoA	0.174***	0.440***
Step 2	AoA x Literacy onset age 3	-3.220	1.572
	AoA x Literacy onset age 4	-0.653	0.369
	AoA x Literacy onset age 5	0.150	0.387*
	AoA x Literacy onset age 6	0.151	0.348***
	AoA x Literacy onset age 7	0.154	0.310***

***p<.001; **p<.01; *p<.05; β is standardized beta value

General Discussion

Onset of literacy has a profound effect on cognition, but generally these effects have been assessed on the representations involved in reading, rather than the pathways involved in mapping between these representations (Hulme et al., 2012; Morais et al., 1979; Smith et al., 2014). In this paper, we show that onset of literacy likely has a long-standing impact on the architecture of the reading system. For words that are in the learner's vocabulary prior to onset of literacy, reading can proceed via two routes – directly, by newly learned mappings from orthography to semantics, or orthography to phonology, or can instead exploit indirect pathways that incorporate learned mappings between phonology and semantics that the learner already has cemented in their language system.

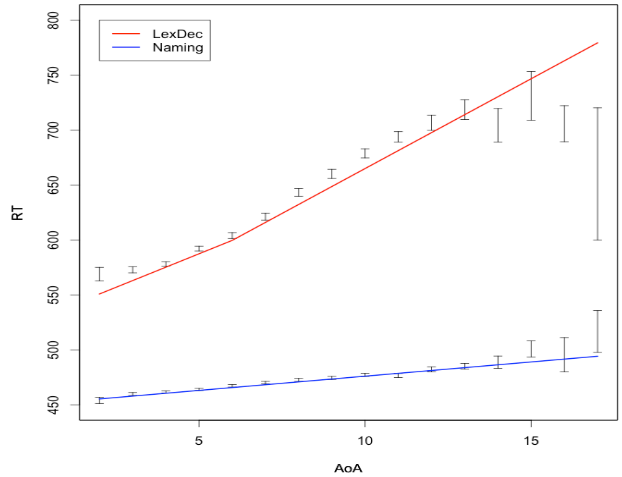

Figure 2. Interaction between AoA and onset of literacy at age 6 in lexical decision but not in naming responses.

For naming tasks, the use of this prior semantics to phonology knowledge has a minimal effect, because the quasi-regularity of orthography to phonology mappings is relatively easy to acquire. The greater difficulty of learning an arbitrary mapping from orthography to semantics, then using this semantic representation to activate the previously acquired phonological representation for the known word, means that this indirect processing is unlikely to be involved differentially for words learned pre- versus post-literacy.

For lexical decision, or other tasks involving activation of semantic representations, the role of literacy onset appears to be quite different. The computational model predicted that when prior knowledge about phonological and semantic associations is available, as it is for pre-literacy acquired words, then an indirect route is likely to be involved in mapping from orthographic to semantic representations. For words learned post-literacy, this prior knowledge is not available, and so the reading system has to proceed via generating either a new mapping from orthography to semantics, or a new mapping from phonology to semantics. Thus, a distinct pattern of response is likely to be observed for lexical decision of pre- and post-literacy words.

The behavioural results provide support for the computational predictions of different pathways used in reading pre- versus post-literacy. Even though literacy onset was several years before the participants in the lexical decision study were tested, the vestiges of literacy onset appear to be still observable in reading behaviour. We acknowledge that literacy onset is not a sudden change, as some new words will still be acquired aurally even after reading training has commenced, and proficient reading is not immediate, but requires extensive, sometimes strenuous, training (e.g., Seidenberg, 2017). Nevertheless, we have shown that literacy onset changes the use that the reader makes of the language system, and this differential use of the system survives to be observed in behavioural responses even after decades of reading practice.

Acknowledgments

This research was supported by ESRC grant RES-000-22-4049.

References

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, *55*, 290-313.

Baayen, R.H., Pipenbrock, R. & Gulikers, L. (1995). *The CELEX Lexical Database* (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283-316.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445-459.

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition, 13*(7-8), 992-1011.

Catling, J. C., & Johnston, R. A. (2009). The varying effects of age of acquisition. *Q J Exp Psychol, 62*(1), 50-62.

Chang, Y. N., Furber, S., & Welbourne, S. (2012). "Serial" effects in parallel models of reading. *Cognitive Psychology, 64*(4), 267-291.

Chang, Y. N., Monaghan, P., & Welbourne, S. (2016). Effects of experience in a developmental model of reading. *Proceedings of the 38th Cognitive Science Society Conference.*

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535-555). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: an analysis of 2,342 words. *Quarterly Journal of Experimental Psychology, 60*(8), 1072-1082.

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*(5), 1103-1123.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica, 115*(1), 43-67.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*(3), 662-720.

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read combining intervention studies with mediation analyses. *Psychological Science, 23(6),* 572-577.

Juhasz, B. J. (2005). Age-of-Acquisition effects in word and picture identification. *Psychological Bulletin, 131*(5), 684-712.

Kosmidis, M. H., Tsapkini, K., Folia, V., Vlahou, C. H., & Kiosseoglou, G.. (2004). Semantic and phonological processing in illiteracy. *Journal of the International Neuropsychological Society, 10(6),* 818-827.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978-990.

Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition, 13,* 292-248.

Marcus, M. P., Marcinkiewicz, M. A., & Santorini, B. (1993). Building a large annotated corpus of English: the penn treebank. *Computational Linguistics, 19*(2), 313-330.

Miller, G. A. (1990). WordNet: An on-line lexical database. *International Journal of Lexicography, 3,* 235-312.

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling--sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 183-206.

Monaghan, P., Chang, Y.N., Welbourne, S., & Brysbaert, M. (2017). Exploring the relations between word frequency, language exposure, and bilingualism in a computational model of reading. *Journal of Memory and Language, 93*, 1-21.

Monaghan, P., & Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language, 63*(4), 506-525.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7(4),* 323-331.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes, 12*(5-6), 765-805.

Steyvers, M., & Tenenbaum, J. B. (2005). The Large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41-78.

Szwed, M., Ventura, P., Querido, L., Cohen, L., & Dehaene, S. (2012). Reading acquisition enhances an early visual process of contour integration. *Developmental Science, 15(1),* 139-149.

Zeno, S. (Ed.). (1995). *The educator' s word frequency guide.* Brewster, NJ: Touchstone.