Literacy and early language development: Insights from computational modelling

Special issue: *Computational modelling of language development with a focus on integration across linguistic levels and (traditionally) non-linguistic levels*

Wordcount: 4832 (6699 including title page+references)

Padraic Monaghan

Lancaster University, UK

Correspondence to:

Padraic Monaghan

Department of Psychology

Lancaster University

Lancaster LA1 4YF

UK

Email: p.monaghan@lancaster.ac.uk

Abstract

Computational models of reading have tended to focus on the cognitive requirements of mapping among written, spoken, and meaning representations of individual words in adult readers. Consequently, the alignment of these computational models with behavioural studies of reading development has been limited. Computational models of reading have provided us with insights into the architecture of the reading system, and these have recently been extended to investigate literacy development, and the early language skills that influence children's reading. These models show us: how learning to read builds on early language skills, why various reading interventions might be more or less effective for different children, and how reading develops across different languages and writing systems. Though there is growing alignment between descriptive models of reading behaviour and computational models, there remains a gap, and I lay out the groundwork for how translation may become increasingly effective through future modelling work.

Keywords: literacy development; computational modelling; oral language; phonological development; vocabulary; comprehension.

Language and Early Literacy: Insights from Computational Modelling

Understanding literacy and translating between models

Literacy is a foundational skill in children's education, and early literacy development has a profound effect on life outcomes (Harold, Acquah, Sellers, & Chowdry, 2016; von Hippel et al., 2017). Understanding how children learn to read, and how to best support their literacy development are thus crucially important issues in children's early development. There has been substantial progress made in describing literacy development in children, and uncovering sets of tasks relating to children's early language and educational development that are predictive of children's reading skills (e.g., Castles, Rastle, & Nation, 2018).

Reading is frequently differentiated into reading fluency and reading comprehension (Foorman et al., 2015; Ouellette, 2006; Ouellette & Beers, 2010). Reading fluency measures the reader's ability to produce spoken forms of words from written forms. Reading comprehension relates to the reader's ability to determine the meaning of text, either in terms of knowing the meaning of individual words or in terms of understanding events described in sentences, or narratives described in paragraphs (Language and Reading Research Consortium, 2015). Whereas these skills are related, they do show distinct trajectories during literacy development, with fluency and comprehension closely related in early stages of reading, but tending to bifurcate later in reading development, as syntactic and discourse structures tend to become more complex (Language and Reading Research Consortium, 2015; Ouellette, 2006).

The Simple View of Reading (SVR) (Gough & Tunmer, 1986) provided a milestone in developing a framework for how early language skills affect reading development. The SVR proposed that reading rests upon two key abilities; first, the ability to decode letters and sets of letters and map them onto speech sounds; and second, oral language comprehension skills. In a large-scale meta-analysis of children's reading comprehension as related to preschool language abilities, Hjetland et al. (2020) estimated the extent to which decoding and reading comprehension related to various early language abilities. They found that children's reading fluency was predicted directly by decoding skills, involving children's letter knowledge and abilities to manipulate and recognize phonemes and rhymes, and indirectly by oral language skills involving vocabulary and grammatical skills, which contributed by enhancing children's decoding skills (see Figure 1). Reading comprehension was predicted directly by both oral language skills and by reading fluency, consistent with the SVR framework (see also Catts et al., 2006). The meta-analysis also revealed that reading comprehension was influenced more by oral language skills and less by reading fluency skills in more advanced reading stages than earlier stages, where reading fluency had a stronger influence on comprehension.

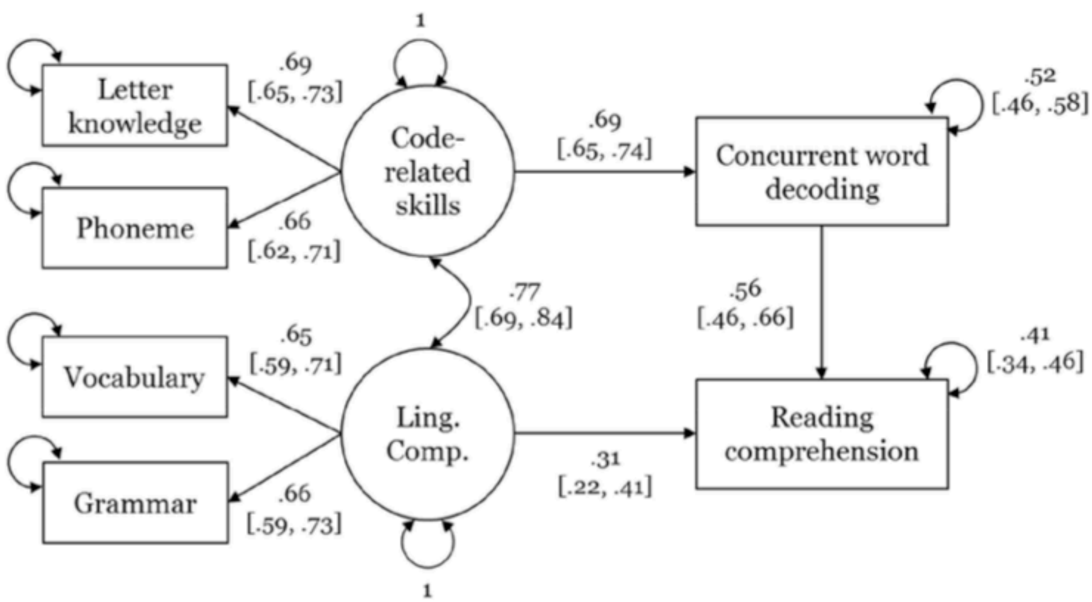

Figure 1. Meta-analysis structural equation model (Fig. 7 in Hjetland et al., 2020) showing relations among oral language skills and reading fluency (concurrent word decoding) and reading comprehension.

These large-scale behavioural studies provide substantial insight into the kinds of tasks that can predict children's reading development, yet there remains a disconnect between behavioural descriptions of reading and theoretical models that examine the cognitive mechanisms involved in children's learning to map between written and spoken and meaning forms of words. In pursuit of this, computational models of reading have attempted to articulate precisely the processing requirements for converting written into spoken and meaning forms of words, and consequently attempt to clarify the cognitive mechanisms required for learning to read.

Computational models of (mature) reading

Computational modelling provides a stringent test of theoretical models of behavioural phenomena. By constructing a model, and simulating behaviour, the adequacy of assumptions about the cognitive mechanisms involved in a cognitive task can be revealed (Sawi & Rueckl, 2019). Models can then be assessed for the extent to which they approximate behaviour, and hence assessed for whether the mechanisms they implement can underly behaviour. Computational models of reading have tended to cluster around two traditions, each of which has been very productive in determining the task constraints and cognitive mechanisms involved in adult reading skills (see Seidenberg et al., 2022).

The first tradition derived from the connectionist modelling approach, and gave rise to the triangle model of reading (Figure 2). The triangle model approach investigated the computational requirements from mapping written to spoken and meaning forms that were purely consequent on the nature of those mappings themselves. Models are exposed to written forms of words and trained to learn to map onto target spoken and meaning forms of words, by adjusting connection strengths between sets of units representing letters and sets of units representing sounds and meanings. The triangle model was thus a minimally defined architecture – the architecture of the reading system was emergent from the computational requirements of forming mappings among representations.

In early manifestations of the triangle model, just written and spoken forms of words were implemented in order to simulate reading fluency (Seidenberg & McClelland, 1989; Harm & Seidenberg, 1999; Zevin & Seidenberg, 2002). However, meaning representations have also been included in fuller implementations of the triangle model (e.g., Chang et al., 2019; Harm & Seidenberg, 2004; Plaut et al., 1996) which can simulate reading comprehension (of single words). Including meaning representations also opens up the possibility of including the role of oral language skills in reading, by investigating the nature of mappings between spoken and meaning forms of words.

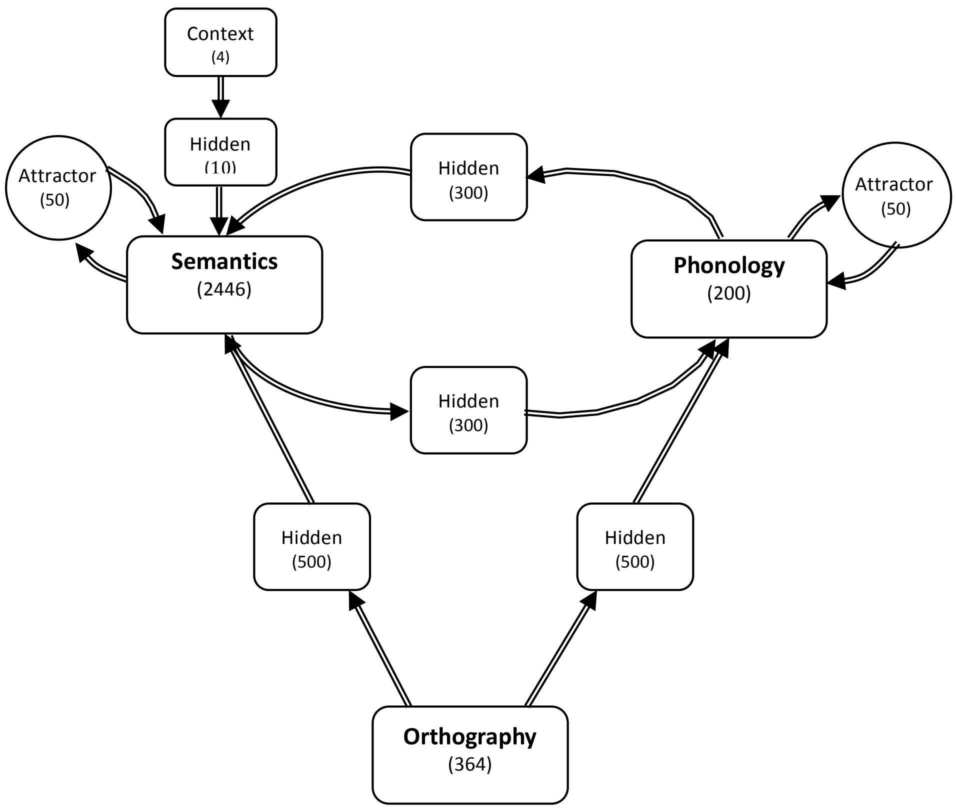

Figure 2. The triangle model (from Chang et al., 2019).

The second influential tradition in computational models of reading is more consistent with a symbolic modelling tradition (Foorman, 1994; Schneider & Graham, 1992), where the architecture that converts written to spoken forms of words is explicitly defined. Key among these models is the dual route cascaded (DRC) model of reading (Coltheart et al., 2001), where written words are pronounced via two routes: one containing lexical representations where whole word written forms are mapped onto whole word spoken forms, and the other, operating simultaneously, via a set of grapheme-to-phoneme correspondence rules where letters and sets of letters are mapped onto phonemes. Unlike the triangle model tradition, mappings are not learned, but are pre-specified and hard-wired into the model. These hard-wired connections are then weighted according to observations about behaviour. For instance certain whole word forms are programmed to be activated more quickly than others, to reflect behaviour showing that particular words being accessed more quickly than others. An adaptation of the DRC, the Connectionist Dual Process model (CDP+, Perry et al., 2006), has included a learning component, where the lexical route is pre-specified as in the DRC model, but the grapheme-to-phoneme correspondence rules can be learned in a connectionist system as a consequence of exposure to written and spoken forms of words.

Both the triangle model and the dual route modelling traditions have been extremely successful in providing detailed descriptions of adult reading behaviours, and clarifying the cognitive processes underlying them (Adelman & Brown, 2008), as well as providing explanations for reading impairments such as phonological dyslexia (Harm & Seidenberg, 1999). Computational models of reading have also had success in reading in different languages (Seidenberg, 2011), consequent on different language properties (Ziegler & Goswami, 2005) and different writing systems (Frost, 2012). For instance, Pagliuca and Monaghan (2010) demonstrated that a connectionist model of reading in Italian learned to map orthography to phonology very quickly as it was able to exploit the systematic letter to sound mappings, more quickly than similar models learning to read English. This reflected the fact that children learn reading fluency in Italian more quickly than children learning English (Seymour et al., 2003). Similarly, the dual route architecture tradition of modelling has been shown to apply to European languages other than English, for instance by encoding a separate set of grapheme to phoneme correspondence rules for German (Ziegler et al., 2010).

Developmental studies of children in cultures that are not WEIRD are sparse (Nielsen et al., 2017), and this applies equally to computational models of reading (though see, e.g., Chang et al., 2016; Ueno & Lambon-Ralph, 2013, for rare exceptions). However, in order to fully understand reading development, it is critical to examine the range of literacy systems globally – there are, for instance, approaching 2 billion readers of non-alphabetic writing systems (Smith et al., 2021). Yang et al. (2013) explored the cognitive consequences of reading in Chinese, where the Chinese logographic writing system has a large number of distinct characters, which correspond to morphemes (Zhou, 1978) but not transparently to sounds (Tong et al., 2009). Yang et al. (2013) found that the model of Chinese was able to learn reading fluency and word reading comprehension accurately, but that it acquired the mappings in a different way than a comparable English model, learning more easily the written to meaning mappings than the written to spoken mappings, whereas the opposite pattern was observed in English. In a controlled comparison among several writing systems using the triangle model framework, Smith et al. (2021) showed how the writing system can have a profound impact on the architecture of the reading system, with greater use of written to meaning forms for writing systems that are progressively more opaque (in terms of how systematic are relations between written and spoken forms). The pathways in the adult reading system may vary substantially across different literacy cultures.

These implemented models of reading have thus proven successful in demonstrating how the computational requirements of mapping among representations of words – written, spoken, and meaning – affect, and are affected by, the cognitive components involved in reading and reading disorders. As a default these computational models of reading aim to simulate adult reading behaviour (Coltheart et al., 2001; Harm & Seidenberg, 1999; Perry et al., 2007; Seidenberg & McClelland, 1989) – the outcome, rather than the process of learning – and so they are not intended to capture reading development. Hence, translating from a model of mappings among *representations* as in the connectionist triangle model shown in Figure 2 to a model of pathways among *tasks* in a behavioural model such as in Figure 1 remains an ongoing problem. However, computational models of reading that have attended to language development enable this gap to be narrowed, and these models are described in the next section.

Computational models of reading development

Computational models that have aimed to capture aspects of reading development have tended to take one of two different approaches. The first approach investigates whether there are residual effects of the process of learning to read as exhibited in adult reading behaviour. The second approach recognizes – and attempts to simulate – the effect of children's early language experience prior to learning to read, and then following the model as it incrementally learns to read to examine the impact of this prior language experience on reading development.

Reading development crystallised in adult reading

The age of acquisition (AoA) of a word has an effect on adult reading behaviour: the earlier a word is acquired the more quickly and accurately it is read (Brown & Watson, 1987; Brysbaert et al., 2000). This effect is independent of frequency, such that words with similar frequency (either cumulative frequency or current frequency) but that differ in age of acquisition are responded to differently (Morrison et al., 2002). Ellis and Lambon Ralph (2000) constructed a computational model that learned to map between abstract input and output patterns (mimicking written and spoken forms of words, such that there was a degree of systematicity between the input and output representations, reflecting the close but not perfect correspondence between letters and speech sounds). The model was a neural network and was trained with backpropagation, such that the connections between units were adjusted to reduce the model's error when producing spoken output from written input. They found that the mappings for patterns to which the model was exposed earlier in its training were learned more accurately than patterns that the model experienced later in its training. This was due to a higher degree of plasticity for the computational system early in its development – when the model's experience is limited it can adjust strengths of connections more effectively to learn earlier mappings, than later when subsequent learning depends on reorganizing the connections to adapt to new mappings.

Monaghan and Ellis (2010) tested early plasticity as a potential explanation for AoA effects observed in adult reading in a reading model of English. They trained the written to spoken part of the triangle model of reading, presenting written words incrementally to the model according to an order that approximated a child learning to read. After learning to read all the words in the language, the AoA effect was apparent in reading fluency: words experienced earlier in training were read more accurately than those experienced later, even when frequency of the word had been taken into account (Zevin & Seidenberg, 2004). Chang et al. (2019) expanded the Monaghan and Ellis (2010) model to the full triangle model of reading. They trained the model to map from written to spoken *and* meaning representations with words presented incrementally according to age-appropriate reading material from age 5 years and upwards. After the model had learned to read all words in the language, the model demonstrated AoA effects for both reading fluency and reading comprehension tasks, but a substantially larger effect of AoA for reading comprehension – where the arbitrary mappings between written to meaning representations were implicated which reflected human behaviour (Brysbaert et al., 2000). The individual's personal history of learning to read affects their processing as a mature reader, and demonstrate how this learning trajectory exerts stronger or weaker effects depending on the reading task (Taylor et al., 2005).

Early oral language development

The second approach to computational modelling of literacy is to take into account children's oral language experience and its effect on learning to read. This approach draws the computational approach a step closer to behavioural models of reading, such as that illustrated in Figure 1.

Harm and Seidenberg (1999)'s model of reading was initially exposed to spoken representations of English words prior to receiving any orthographic representations. This was to simulate readers having substantial experience of hearing language before learning to read. The model was able to develop phonological "attractors", where knowledge of the structure of spoken words in English was manifested before the model began to learn to read. However, the model did not contain any meaning representations of words, and so prior experience of words' meanings was not incorporated into the model. Furthermore, gradual development of the model was also not included in Harm and Seidenberg's (1999) model, as it experienced all words orally that it would later learn to read, hence, there was no incrementality in the model's design.

Chang and Monaghan (2019) included oral language experience as a key principle in determining how preschool language experience influenced later literacy learning in their computational model of reading. As described above, according to the Simple View of Reading (Gough & Tunmer, 1986), children's early oral language comprehension ability, coupled with decoding ability, is a key contributor to successful reading. Children's experience of oral language is enormously variable, meaning that some children have substantial training on mappings between words' sounds and meanings, whereas other children have much less opportunity to learn these representations (Anderson et al., 2021). In addition to quantity of oral language experience, quality of the experience matters too for children's early language development. Rowe (2012) showed that the properties of child-directed speech by adults influenced children's language development, particularly when more complex constructions, involving longer utterances and broader vocabulary, were used.

Chang and Monaghan (2019) exposed the triangle model to varying quantities of preschool oral language experience – where the model learned to map between spoken and meaning representations of English words. They also varied the quality of oral language experience, in terms of the range of vocabulary that the preschool model was exposed to. After this preliteracy oral language training, they then gave the model written representations of words and required the model to learn to pronounce the words (reading fluency task) and identify the meaning of the word (word reading comprehension task). They found that both quantity and quality of oral language experience exerted an effect: the more words heard, and the wider the range of the vocabulary from which those words were sampled, the more quickly and accurately the model learned to read, especially for the word reading comprehension task. This was because the model could use the pretrained spoken to meaning, and meaning to spoken, mappings to assist in increasing the fidelity of the spoken and meaning representations of words (Perfetti & Stafura, 2014), and to assist in dividing the labour among the pathways in the model: a model that had a good ability to produce the meaning of a spoken word can then map the written form onto its meaning via the easier to acquire written to spoken pathway in the model. Similarly, but to a lesser degree, the meaning to spoken pathway could also be incorporated into producing the spoken form of written words with support from the harder to learn, but still contributory, written to meaning pathway.

Curiously, however, the model also showed that lots of experience of a limited vocabulary was detrimental to later learning to read. This was because of the principles of plasticity in the model. If it learned to limit its understanding of a smaller vocabulary, then it was harder for the model to reconfigure to read words that were not part of its preliteracy oral language experience. This lack of plasticity was only the case when the model had had a lot of experience of the small vocabulary. A little experience of a small vocabulary rendered a system that still retained plasticity to expand to a broader vocabulary when the opportunity for that enriched reading environment arose.

Chang et al. (2019) extended the simulations from Chang and Monaghan's (2019) model to examine how this preschool oral language experience affected learning to read when reading was simulated with incremental growth of the vocabulary, according to children's gradual expansion of reading material from age 5 upwards. The model showed that division of labour between spoken and meaning representations was critical for solving the cognitive task of learning to map written words onto their spoken and meaning forms. They found greater division of labour between spoken and meaning representations earlier in reading training than later (see also Yang et al., 2013), but also critical was whether words were learned before or after the onset of literacy. Chang et al. (2019) found that words the reader did not know orally prior to learning to read were acquired more slowly, and were read with greater involvement of the written to meaning pathway in the model (so meaning was more involved even in reading fluency tasks). Words which the model knew orally prior to learning to read were read more quickly and tended to have greater involvement of the written to spoken pathway, even for word reading comprehension tasks. Chang et al. (2019) tested this prediction of the model in a reanalysis of adult reading behaviour data, which demonstrated that words are not all read in the same way, and that the reading architecture itself is not a monolithic system, but one that adaptively and flexibly applies to different requirements of different words experienced at different life stages.

These triangle models of oral language experience affecting reading development (Chang & Monaghan, 2019; Chang et al., 2019) had further implications for the consequences of reading training for children with different pre-literacy experience. The models predict that children with larger vocabularies will read more of those words more efficiently, decoding from written to spoken forms, and accessing meaning from the spoken form. Children with smaller vocabularies – who acquire more of their vocabulary from reading – will acquire words more effortfully, requiring the arbitrary mapping between written and meaning representations to be acquired in order to support effective comprehension of these words. Siegelmann et al. (2020) showed that children with better reading fluency were more influenced by regularities between written and spoken forms, indicating that they tended to rely more on the written to spoken over the written to meaning pathway, consistent with Chang et al.'s (2019) computational model of reading.

Differences in preschool oral language skills may also have an important role in determining how effective different literacy training schemes might be. If children have good oral language skills, then they will be able to learn to read words (in an alphabetic writing system) for pronunciation and for meaning via the easier to acquire, more systematic written to spoken mapping, with meaning activated partially via the spoken to meaning mappings that were trained pre-literacy. However, children who have poorer oral language skills will be required to acquire the direct written to meaning mappings for word reading comprehension, which are arbitrary and harder to learn. Hence, focusing on the pronunciation of written words will be most beneficial when the child already has good oral language skills, otherwise a focus on pronunciation will end in a cul-de-sac in terms of generating meaning: the meaning will not be effectively activated via the spoken form of the word.

Chang et al. (2020) applied the triangle model of reading to investigate how alternative methods for literacy training might influence reading development differently according to children's oral language skills. The triangle model of reading in English was trained to learn to read under two training regimes, one that simulated a meaning-focused training, and the other that reflected a sound-focused training. These were distinguished in terms of which of the pathways in the model (written to meaning, or written to spoken) were being trained most often. The model with good preschool oral language skills responded well to both forms of training, whereas the model with poor preschool oral language skills performed more poorly on word comprehension after sound-focused compared to meaning-focused training. The model results are consistent with calls to focus during early literacy on simultaneously improving oral language skills (Castles et al., 2018; Ricketts et al., 2007), and requests not to neglect the importance of direct written to meaning mappings in understanding reading development (Taylor et al., 2015).

Future directions and challenges

So, how close is the alignment between computational models (e.g., Figure 2) and behavioural studies (e.g., Figure 1) of reading development? There are several ways in which the divide is being narrowed by these developmental computational models of reading. One benefit of the developmental computational models is that they have incorporated oral language skills as well as reading training into their performance, testing explicitly early language skills and their relation to reading development. Thus, individual differences in the model's oral language skills enable the computational models to make hypotheses about *why* oral language skills predict reading fluency (decoding) and comprehension, and why links from oral language to comprehension are stronger than those to fluency (see, e.g., Figure 1). The models also provide explanations for how the role of oral language skills may change over time, and why certain reading schemes may be more or less successful for readers with different oral language profiles.

However, substantial gaps still remain. There are limitations to extant computational models of reading in terms of how closely they resemble human reading behaviour. With very few exceptions (Ans et al., 1998; Pagliuca & Monaghan, 2010; Perry et al., 2010; Rastle et al., 2000), models have tended to simulate reading only of monosyllables. In English, monosyllables represent 70% of word tokens that children read, but polysyllabic words present crucial (and interesting) challenges both for children learning to segment words into syllables during reading (Duncan & Seymour, 2003; Kearns, 2020; Mousikou et al., 2017, for a study with adults), and for computational models that have attempted to simulate those behaviours, usually requiring substantial additional machinery compared to monosyllables (e.g., Ans et al., 1998; Perry et al., 2010; Plaut, 1999; Rastle & Coltheart, 2000).

Another lack of resemblance is that models of word reading tend to be just that – models of individual, isolated word reading. Reading comprehension, in contrast, tends to be tested with questions about narrative and discourse, rather than merely identifying the meaning of a single word (which often cannot be isolated in its meaning from the rest of the text) (Ouellette, 2006). Computational models of reading development at the word level therefore need to become more closely aligned with models of text reading (Reichle, 2021), such that insights at the word level can permeate models that simulate readers' responses to longer texts.

Greater specificity in the spoken and meaning representations used in models of reading – and inclusion of more facts about visual processing of stimuli – can also help to clarify some of the other preschool pre-literacy language skills that relate to reading skill development. For instance, phonological awareness tasks have been related to fidelity of spoken representations in the triangle model (Harm & Seidenberg, 1999; Smith et al., 2021), and phonics training has been shown to increase the precision of individual phoneme representations further (Harm et al., 2003). However, computational models represent sequences of phonemes either categorically (one unit for each phoneme) or in terms of sets of phoneme features, which is an abstraction away from the auditory properties of words. Implementing auditory features may better highlight how phonological processing skills exert an influence on decoding skills, and consequent reading development. Similarly, meaning is represented in these models either categorically, or in terms of semantic features (derived from encyclopedic definitions of words or vectors derived from contextual co-occurrences in text (Mandera et al., 2017), and links among meanings are seldom included (apart from abstract contextual units (Chang et al., 2019), or via weighted links based on free association norms (Li et al., 2004; Steyvers & Tenenbaum, 2005)). There are also relevant models that more closely describe and simulate how vocabulary develops, which is not due only to frequency in the environment, but also how the new concept links to the knowledge network of previously acquired words (Jiménez & Hills, 2022). These alternative representations of meaning may help further unpack the role of grammatical skills and vocabulary knowledge that are generally elided into a composite measure of oral language in behavioural models of reading (e.g., Duff et al., 2015; Foorman et al., 2015; Hjetland et al., 2018).

Other opportunities for future computational models of reading to enable closer alignment with behavioural studies can expand on the groundwork that has been laid in terms of cross-linguistic and cross-orthography analyses of reading. Implementing different writing systems, and determining how variety in phonological systems across the world's languages affect learning to read will give us new insights in constraining the cognitive processes that we have inferred as involved in reading (from studies of English and the handful of other languages that have been investigated). These models will also enable us to determine how transfer effects from one language to another (and across writing systems) may affect reading development in children that move across cultures (Melby-Lervåg & Lervåg, 2014). Such models would also require determining how language skills (both oral and orthographic) transfer from one language to another, and these may have surprising and unexpected effects in terms of where advantages of transfer can occur (see, e.g., Monaghan et al., 2017 for transfer from English to Dutch in a model of reading, and Paulesu et al., 2021, for insights into how bilingualism may alter reading behaviour compared to monolingualism).

Conclusion

There is a now long-standing tradition of computational models of reading that have provided numerous insights into the processes involved in adult reading. Recent extensions of these computational models to investigate reading development have opened up possibilities for closer alignment of propositions about cognitive mechanisms involved in reading with behavioural observations of early language skills affecting children's reading development. There are points where divergence remains between computational and behavioural studies, but the consideration of a range of oral language skills and their relation to reading development in the computational models results in widening opportunities for co-creation of new insights into reading acquisition (see, e.g., Ziegler et al., 2020). Among these opportunities are: examining the causes of individual differences in learning to read; understanding the key pressure points in learning to read (such as transitions from monosyllabic to polysyllabic reading); predicting the impacts of different interventions on children's learning; and discovering how learning may vary across different languages and writing systems. The gap between behavioural and computational studies can then be narrowed through future behavioural work that specifies more closely the cognitive mechanisms involved in children's language task performance (e.g., Siegelmann et al., 2022), together with increasingly detailed computational modelling that gets closer to children's language experience, and the constraints on their perception and production of language.

Acknowledgments

This work was supported by the International Centre for Language and Communicative Development (LuCiD) at Lancaster University, funded by the Economic and Social Research Council (UK) [ES/S007113/1].

References

Adelman, J. S., & Brown, G. D. (2008). Methods of testing and diagnosing model error: Dual and single route cascaded models of reading aloud. *Journal of Memory and Language*, *59*(4), 524-544.

Adlof, S. M., Catts, H. W., & Little, T. D. (2006). Should the Simple View of Reading include a fluency component? *Reading and Writing, 19*(9), 933-958.

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development*, *92*(2), 484-501.

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, *105*(4), 678.

Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition, 15,* 208–216.

Brysbaert, M., van Wijnendaele, I., & Deyne, S. D. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica, 104*(2), 215-226.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5-51.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the Simple View of Reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278–293.

Chang, Y-N., & Monaghan, P. (2019). Quantity and diversity of pre-literacy language exposure both affect literacy development: Evidence from a computational model of reading. *Scientific Studies of Reading, 23*(3), 235-253.

Chang, Y-N., Monaghan, P., & Welbourne, S. (2019). A computational model of reading across development: Effects of literacy onset on language processing. *Journal of Memory and Language, 108*, 104025.

Chang, Y.-N., Taylor, J. S., Rastle, K., & Monaghan, P. (2020). The relationships between oral language and reading instruction: Evidence from a computational model of reading. *Cognitive Psychology, 123*, 101336.

Chang, Y.-N., Welbourne, S., & Lee, C.-Y. (2016). Exploring orthographic neighborhood size effects in a computational model of Chinese character naming. *Cognitive Psychology, 91*, 1-23.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204-256.

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, 56*(8), 848-856.

Duncan, L. G., & Seymour, P. H. (2003). How do children read multisyllabic words? Some preliminary observations. *Journal of Research in Reading*, *26*(2), 101-120.

Foorman, B. R. (1994). The relevance of a connectionist model of reading for "The great debate". *Educational Psychology Review*, *6*(1), 25-47.

Foorman, B. R., Herrera, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). The structure of oral language and reading and their relation to comprehension in kindergarten through grade 2. *Reading and Writing, 28*(5), 655-681.

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.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education,7*(1), 6-10.

Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for dis- abled readers. *Scientific Studies of Reading*, *7*, 155-182.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106*(3), 491-528.

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.

Harold, G., Acquah, D., Sellers, R., & Chowdry, H. (2016). *What works to enhance inter-parental relationships and improve outcomes for children? department for work and pensions*. *Early intervention foundation report*. UK Government: Department for Work and Pensions. http://www.eif.org.uk/publication/what-works-to-enhance- inter-parental-relationships-and-improve-outcomes-for-children-3. Retrieved 13/4/21.

Hjetland, H. N., Brinchmann, E. I., Scherer, R., Hulme, C., & Melby-Lervåg, M. (2020). Preschool pathways to reading comprehension: A systematic meta-analytic review. *Educational Research Review*, *30*, 100323.

Jiménez, E., & Hills, T. (2022). Semantic maturation during the comprehension-expression gap in typical and late talkers. *Child Development, in press.*

Kearns, D. M. (2020). Does English have useful syllable division patterns? *Reading Research Quarterly, 55*(S1), S145–S160.

Language and Reading Research Consortium (2015). The dimensionality of language ability in young children. *Child Development, 86(6)*, 1948-1965.

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin*, *140*(2), 409-433.

Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks, 17*(8), 1345-1362.

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language, 92*, 57-78.

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.

Morrison, C. M., Hirsh, K. W., Chappell, T., & Ellis, A. W. (2002). Age and age of acquisition: An evaluation of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology*, *14*(4), 435-459.

Mousikou, P., Sadat, J., Lucas, R., & Rastle, K. (2017). Moving beyond the monosyllable in models of skilled reading: Mega-study of disyllabic nonword reading. *Journal of Memory and Language, 93,* 169–192.

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31-38.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554-566.

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, *23*, 189-208.

Pagliuca, G., & Monaghan, P. (2010). Discovering large grain sizes in a transparent orthography: Insights from a connectionist model of Italian word naming. *European Journal of Cognitive Psychology*, *22*(5), 813-835.

Paulesu, E., Bonandrini, R., Zapparoli, L., Rupani, C., Mapelli, C., Tassini, F., ... & Zorzi, M. (2021). Effects of orthographic consistency on bilingual reading: Human and computer simulation data. *Brain sciences*, *11*(7), 878.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22-37.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review*, *114*(2), 273-315.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive psychology*, *61*(2), 106-151.

Plaut, D. C. (1999). A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. *Cognitive Science*, *23*(4), 543-568.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115.

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language, 42*(3), 342–364.

Reichle, E.D. (2021). *Computational models of reading: A handbook*. Oxford: Oxford University Press.

Sawi, O. M., & Rueckl, J. (2019). Reading and the neurocognitive bases of statistical learning. *Scientific Studies of Reading, 23*(1), 8–23.

Schneider, W., & Graham, D. J. (1992). Introduction to connectionist modeling in education. *Educational Psychologist*, *27*(4), 513-530.

Seidenberg, M. S. (2011). Reading in different writing systems: One architecture, multiple solutions. In McCardle, P., Ren, J. & Tzeng, O. (Eds.), *Dyslexia Across Languages: Orthography and the gene-brain behavior link*. Baltimore, MD: Paul Brooke.

Seidenberg, M. S., Farry-Thorn, M., & Zevin, J. D. (2022). Models of word reading: What have we learned? In Snowling, M.J., Hulme, C., & Nation, K. (Eds.), *The science of reading: A handbook (2nd edition)*. Chichester, UK: Wiley, pp.36-59.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523-568.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143-174.

Siegelman, N., Rueckl, J. G., Steacy, L. M., Frost, S. J., van den Bunt, M., Zevin, J. D., Seidenberg, M. S., Pugh, K. R., Compton, D. L., & Morris, R. D. (2020). Individual differences in learning the regularities between orthography, phonology and semantics predict early reading skills. *Journal of Memory and Language, 114*, 10415.

Siegelman, N., Rueckl, J. G., van den Bunt, M., Frijters, J. C., Zevin, J. D., Lovett, M. W., ... & Morris, R. D. (2022). How you read affects what you gain: Individual differences in the functional organization of the reading system predict intervention gains in children with reading disabilities. *Journal of Educational Psychology*, *114*(4), 855-869.

Smith, A. C., Monaghan, P., & Huettig, F. (2021). The effect of orthographic systems on the developing reading system: Typological and computational analyses. *Psychological Review, 128*, 125-159.

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.

Taylor, J. S. H., Duff, F. J., Woollams, A. M., Monaghan, P., & Ricketts, J. (2015). How Word Meaning Influences Word Reading. *Current Directions in Psychological Science, 24*(4), 322-328.

Tong, X., McBride-Chang, C., Shu, H., & Wong, A. M. (2009). Morphological awareness, orthographic knowledge, and spelling errors: Keys to understanding early Chinese literacy acquisition. *Scientific Studies of Reading*, 13(5), 426-452.

Ueno, T., & Lambon Ralph, M. (2013). The roles of the “ventral” semantic and “dorsal” pathways in conduite d'approche: A neuroanatomically-constrained computational modeling investigation. *Frontiers in Human Neuroscience, 7*, 422.

von Hippel, P., Workman, J., & Downey, D. B. (2017). Inequality in reading and math skills comes mainly from early childhood: A replication, and partial correction, of ‘Are schools the great equalizer?’ *SSRN working paper,* 3036094.

Yang, J., Shu, H., McCandliss, B. D. & Zevin, J. D. (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.

Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition*, *32*(1), 31-38.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3-29.

Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology*, *12*(3), 413-430.

Ziegler, J. C., Perry, C., & Zorzi, M. (2020). Learning to read and dyslexia: From theory to intervention through personalized computational models. *Current Directions in Psychological Science*, *29*(3), 293-300.

Zhou, Y. G. (1978). To what extent are the “phonetics” of present-day Chinese characters still phonetic? *Zhongguo Yuwen, 146*, 172-177.