Published in final edited form as:
Biling (Camb Engl). 2013 April ; 16(Spec Iss 2): 354–366. doi:10.1017/S1366728912000296.

Orthographic influences on division of labor in learning to read Chinese and English: Insights from computational modeling

Jianfeng Yang[1], **Hua Shu**[2], **Bruce D. McCandliss**[3], and **Jason D. Zevin**[4,*]

[1]Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China

[2]State Key laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China

[3]Department of Psychology, Vanderbilt University Peabody College of Education and Human Development, Nashville, TN 37240, USA

[4]Sackler Institute for Developmental Psychobiology, Weill Cornell Medical College, NY 10021, USA

Abstract

Learning to read any language requires learning to map among print, sound and meaning. Writing systems differ in a number of factors that influence both the ease and rate with which reading skill can be acquired, as well as the eventual division of labor between phonological and semantic processes. Further, developmental reading disability manifests differently across writing systems, and may be related to different deficits in constitutive processes. Here we simulate some aspects of reading acquisition in Chinese and English using the same model architecture for both writing systems. The contribution of semantic and phonological processing to literacy acquisition in the two languages is simulated, including specific effects of phonological and semantic deficits. Further, we demonstrate that similar patterns of performance are observed when the same model is trained on both Chinese and English as an "early bilingual." The results are consistent with the view that reading skill is acquired by the application of statistical learning rules to mappings among print, sound and meaning, and that differences in the typical and disordered acquisition of reading skill between writing systems are driven by differences in the statistical patterns of the writing systems themselves, rather than differences in cognitive architecture of the learner.

Keywords

Computational modeling; Reading; Dyslexia

Most research on the development of reading has been done in English. This raises questions about whether insights and models of reading skill and its development will generalize to other languages and writing systems. In fact, reading acquisition and use appear to be quantitatively and qualitatively different across writing systems, in ways that have prompted some theorists to propose that different cognitive architectures would be required to understand reading in different writing systems (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Coltheart, Curtis, Atkins, & Haller, 1993; Perfetti, Liu, & Tan, 2005; Frost, 2005). An alternative view is that these differences may be better understood in terms of statistical properties of the writing system (Ziegler & Goswami, 2005) and the impact these

---

*Address for Correspondence: Jason D Zevin, Box 140, Sackler Institute for Developmental Psychobiology, Weill Cornell Medical College, New York, NY 10021, USA, Tel: (718)619-9757, jdz2001@med.cornell.edu.

may have on the "division of labor" between semantic and phonological processing in reading (Seidenberg, 1995). The current models apply the same basic architecture and learning rules to two very different writing systems – English and Chinese – in order to test these two possibilities. This is also a first-order question in the modeling of biliteracy for these two writing systems, because if the languages cannot be accommodated in a single functional architecture, it will have important consequences for the modeling of biliteracy in these two languages.

The difference between Chinese and English can be understood in the statistical properties of spelling-to-sound and spelling-to-meaning mappings. Although English is something of an "outlier" writing system in mapping from print to sound (Malone, 1925; Venezky, 1999), it has an alphabet of letters that correspond roughly to individual speech sounds (Venezky, 1970). In contrast, Chinese is an extremely "deep" orthography (Frost, Katz, & Bentin, 1987) in that the pronunciation of a character cannot be computed sound-by-sound from its constituent parts (DeFrancis, 1989), although probabilistic cues to pronunciation do exist (Zhu, 1988; Li & Kang, 1993) and are used by both children learning to read and adults (Lee, Tsai, Su, Tzeng, & Hung, 2005; Shu, Meng, Chen, Luan, & Cao, 2005). Chinese is sometimes characterized as the logographic system, in contrast to alphabetic systems, because of the morphemic (Leong, 1973) or even morphosyllabic (e.g., DeFrancis, 1989; Mattingly, 1992) mappings characters afford. Characters, as basic writing units, map onto morphemes – not phonemes – in the spoken language. Furthermore, Chinese characters typically contain a "semantic radical" that provides some probabilistic information that aids in the translation from orthography to semantics. Alphabetic writing systems rarely contain semantic information that is not somehow encoded phonologically. Even where there is ambiguity about spelling to sound for morphological forms (final *-s* and *-ed* in English), it cannot be said that these convey no phonological information at all, in the way that semantic components of Chinese characters do (see Frost, In press; Mirkovic, MacDonald, & Seidenberg, 2005, for discussion).

There is clear evidence among alphabetic orthographies that shallower systems are easier to learn than deeper ones, as reflected by both word and non-word performance in beginning readers (e.g., Ellis & Hooper, 2001; Goswami, Gombert, & Barrera, 1998; Seymour, Aro, & Erskine, 2003). Differences between alphabetic orthographies and Chinese are starker still: The average English-reading child can recognize 3000–5000 words after first grade (White, Graves, & Slater, 1990) whereas Chinese children can typically read fewer than and 800 characters with the same amount of schooling (Xing, Shu, & Li, 2004). Thus, the overall consistency of mappings from units in the writing system to their phonological counterparts has clear effects on the rate at which reading skill can be acquired.

Another consequence of orthographic depth for the acquisition of reading is that related language skills (such as semantic and phonological processing) contribute differentially to reading success across writing systems. Shallow orthographies are characterized by weak effects of semantic variables in skilled reading (e.g., Bates, Burani, D’amico, & Barca, 2001), and a limited contribution of semantic processing skills to the development of reading (McBride-Chang, Cho, et al., 2005; Saiegh-Haddad & Geva, 2008). In relatively "deep" orthographies, such as English, semantic knowledge plays some role in reading aloud, particularly in the reading of words whose spellings are highly atypical (Strain, Patterson, & Seidenberg, 1995, 2002) and there is some evidence for a role of semantic processing abilities in beginning reading skill (Nation & Snowling, 1999; Carlisle, 2000, 2003). In part because print-to-sound cues are even less reliable in Chinese, the role of semantic processing in reading aloud is greater, and the contribution of semantics to the development of Chinese reading is particularly important (Shu, McBride-Chang, Wu, & Liu, 2006; Shu, Peng, & McBride-Chang, 2008).

The differential contribution of semantic and phonological processing across writing systems may also explain differences in the manifestation of reading disability across languages. In English, there is evidence for subtypes of developmental dyslexia: "phonological dyslexics" who have specific difficulty with decoding and "surface dyslexics" who have specific difficulty with atypically spelled words, but relatively spared performance on regular words and non-words (e.g., Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996). These subtypes are often explained as resulting from distinct pre-existing deficits: in semantic processing for the developmental delay/surface dyslexics and phonological processing for the phonological dyslexics. The reading performance of children with developmental surface dyslexia is very similar to that of younger normal readers with respect to the relative difficulty of pseudowords, regular words and irregular-inconsistent words (Manis et al., 1996). Their specific difficulty reading words with unusual spelling to sound correspondences, may thus be associated wither with semantic deficits or with a general delay in the development of reading skill (Plaut, McClelland, Seidenberg, & Patterson, 1996; Nation & Snowling, 1998). In contrast, developmental phonological dyslexia is associated with deficits in phonological processing (Stanovich & Siegel, 1994). In English, phonological dyslexics present with a reading impairment that is most pronounced for non-words, but, in milder cases, can leave exception word reading more or less intact (Castles & Coltheart, 1993).

In Chinese, semantic and phonological processing deficits impact reading in different ways. Poor semantic processing is associated with difficulties reading all types of words, even those with more typical spelling to sound correspondences, although reading of atypically spelled words does suffer relatively more (Shu et al., 2005). Children with phonological deficits are also impaired relative to age-matched controls on reading of all words, but the impairment is greater for words with typical spelling-to-sound correspondences, with the result that phonological dyslexics do not show the typical advantage for regular-consistent over irregular-inconsistent words (Shu et al., 2005). In sum, there are gross differences between writing systems in the relative contribution of phonological and semantic processing abilities to the development of reading skill that appear to be driven largely by the consistency of print-to-sound mappings across languages. In previous studies, we have demonstrated that the same basic architecture and learning rules appropriate to English could model the acquisition and use of reading skill in Chinese, and simulate both effects that are directly analogous to English and effects that are specific to Chinese (Yang, McCandliss, Shu, & Zevin, 2009).

Here, we simulate typical and disordered reading acquisition in English and Chinese by applying the same functional architecture (modified to represent the inputs and outputs for each language) and learning rules for both writing systems. The models implement the theory that reading is acquired via a process of statistical learning of mappings among spelling, sound and meaning and test the hypothesis that differences in the patterns of typical and disordered reading development across writing systems may be explained in terms of differences in the statistical properties of the writing systems rather than by differences in functional architecture. In a second simulation, we model the simultaneous acquisition of Chinese and English, to examine whether the same learning trajectories and sequelae of pre-literate deficits would be observed across languages learned by the same individual. Simulation 2 addresses a first order question in the modeling of Chinese/English biliteracy: Can English and Chinese are learned in the same set of mappings among orthography, phonology and semantics? Or do they require fundamentally different processing assumptions? Further, we can examine how the two languages interact when learned by the same system. It is possible that learning these two very different writing systems at the same time will lead to differences in how reading skill is acquired and used in both writing systems, but it is also possible that when both languages are learned at the same time, the outcome is equivalent to monolingual learning of each. Either outcome would have important consequences for understanding bilingualism and biliteracy in reading development.

Simulation 1: Modeling differential division of labor between Chinese and English in monolingual models

Here we examine development of typical and disordered reading in two parallel models implementing the same functional architecture for English and Chinese. Both models have feed-forward connections from an orthographic input layer to a phonological attractor network (Harm & Seidenberg, 1999), supplemented with a semantic input layer that functions mainly to provide a secondary source of input about word identity that is particularly useful for words with ambiguous spelling-to-sound mappings (Plaut et al., 1996). Following Plaut (1997), we used random bit patterns to capture this contribution of word-specific knowledge to generating a correct pronunciation. While this has the disadvantage of not providing a realistic representation of the similarity of the meanings of words within a language, it has the advantage of permitting us to use the same semantic patterns for both languages, thus allowing a direct investigation of the role of properties of the print-to-sound system on the division of labor.

Methods

Architecture—The same basic architecture (Figure 1) was used for two models: One for Chinese and one for English. Each model had an orthographic input layer designed to represent the spellings of words in each writing system, fully connected to a hidden layer of 100 hidden units, which was in turn fully connected to a phonological output layer designed to represent the pronunciations of words in that language. The phonological output layer was fully connected both directly to itself and to 50 cleanup units, permitting the formation of attractor states, following Harm and Seidenberg (1999). The English representations of orthography and phonology were adapted from the scheme of Harm and Seidenberg (2004): 101 units were used to represent 10 slots of letters in the orthographic layer and 200 units were used for 8 slots to represent phonemes in phonological layer. The Chinese orthographic representation consisted of 270 units based on a linguistic description of Chinese orthography including radicals, number of strokes and radical position, adapted from Xing et al. (2004) by excluding slots that explicitly coded the location of the phonetic component (see Yang et al., 2009, for details). Ninety-two units were used to code each Chinese syllable, which includes five slots: one onset slot, three rime slots, and a fifth slot for tone. As in Zhao & Li’s (2009) PatPho system, each phoneme slot was encoded with the same basic featural representation, but with a slightly different configuration for two languages (e.g., Chinese has palatal and retroflex in addition to bilabial, alveolar and velar, used in the English models).

A second input layer was included to simulate the contribution of semantics in print-to-sound translation. Semantic patterns were 3000 random bit patterns clustered into 120 categories over 200 semantic features. Categories were created by generating a set of 120 prototypes, in which each feature had a probability of 0.1 of being active. Each prototype was then used to generate 25 exemplars by randomly selecting 10% of all features and resetting their probability of activation to 0.05, under the constraint that each exemplar differ from all other exemplars by at least three features. A subset of 2881 patterns was assigned randomly to the words in the English training corpus. A subset of 2689 patterns from the English training patterns were selected and randomly assigned to Chinese characters. In both versions of the model, the semantic input layer was connected to the output layer via 100 hidden units.

Training—Training was carried out in the same way for the English and Chinese versions of the model. We first pre-trained the phonological attractor net to an error threshold of 0.01, and the final weights (240K in Chinese and 60K in the English model) of phonological attractor net were embedded in the reading model. To avoid "catastrophic interference", interleaved training (Hetherington & Seidenberg, 1989) on phonological and reading was adopted. Training mixed 10% "listening" trials, on which only the phonological attractor was trained, with 90% "reading" trials, on which the whole model was trained. A learning rate of 0.005 and momentum of 0.9 were used. Online learning was used with the continuous recurrent back-propagation algorithm (Pearlmutter, 1995). Each word was selected according to the training probability transformed via square root compression.

The Chinese training corpus of 2689 characters consisted of 2390 characters from a set of naming norms (Liu, Shu, & Li, 2007) and 299 additional items from phonetic families represented in the testing materials. Frequency estimates taken from the Modern Chinese Frequency Dictionary (Language and Teaching Institute of Beijing Language College, 1986). The English training corpus consisted of 2,881 monosyllabic words assigned frequencies taken from the Marcus, Santorini, and Marcinkiewicz (1993) norms, which are based on 43 million tokens from the Wall Street Journal.

In both languages, two subtypes of developmental dyslexia were simulated by applying decay to either the hidden units from orthography to phonology (to simulate phonological dyslexia, hereafter PD) or the hidden units from semantics to phonology (to simulate surface dyslexia, hereafter SD). Decay on each weight ω was reduced in magnitude according to the formula $\Delta\omega = -\omega \times \sigma$ where σ was the decay constant. In order to simulate a wide range of deficit severity, 20 different decay values were used, varying from 0.25×10^{-5} to 5×10^{-5} in steps of 0.25×10^{-5}. Unimpaired models were also run 20 times. Each run of the model used a different random seed for the initial randomization of weights and selection order of stimuli.

Testing—Naming accuracy was computed to test the model’s performance. It was determined by applying a winner-take-all scoring system: for each slot on the output layer, we determined which phoneme was closest to the pattern on the output at the final time tick and reported this as the model’s pronunciation.

Test items were drawn from studies of consistency, regularity and frequency effects in the two languages: the 120 Chinese test items were from Yang et al. (2009), and the 144 English test items were those used by Plaut et al. (1996) from Taraban and McClelland (1987). In both languages, the items were sets of regular-consistent, regular-inconsistent and exception words matched for frequency, phonetic family size and other Chinese script properties, such as structure type, the number of strokes and radicals.

The definition of regularity in English and Chinese is slightly different. In English, regular words are those that can be pronounced correctly by rule (although there is some discrepancy between rule sets, due to disagreements about whether rules for units larger than single graphemes are considered, see e.g., Andrews & Scarratt, 1998; Zevin & Seidenberg, 2006). In the current study, "regular" words are those with pronunciations consistent with the rule set of the Dual-Route Cascade model of word reading (Coltheart et al., 2001) which has a large number of multi-grapheme rules, but nonetheless counts many highly inconsistent items as "irregular." In Chinese, a character is considered regular if its pronunciation matches the pronunciation of its phonetic component when this occurs as a single character (see Peng & Yang, 1997; Yang et al., 2009). In both languages, exception words or characters are just those that are not considered regular. Consistency is defined essentially the same way in both languages – completely consistent words share the pronunciation of some critical sub-lexical component with all of the words that contain that component – although the sub-lexical structures of the two languages are of course different. In English, regular inconsistent words were items such as DOLL and BROTH, that have exception words as neighbors (e.g., POLL and BOTH). In Chinese, consistency (like regularity) is defined at the level of the phonetic component. Characters that are regular, but contain a phonetic component that is pronounced different ways in different (exception) characters are regular and inconsistent. Simulations of surface and phonological dyslexia in Chinese children used the items from the original study (Shu et al., 2005).

Results

Overall performance across languages, for typical and disordered reading models—Figure 2 shows the models' accuracy over time for all items in the training set. For the typically developing model, on average the English model reached 90% overall accuracy after 292K trials (*SD* = 10.1K) and the Chinese model reached 90% accuracy after 665K trials (*SD* = 15.7K). A 3 (Deficit: Typical, PD, SD) × 2 (Language: Chinese and English) ANOVA with maximum naming accuracy as the dependent variable revealed significant main effects of Deficit, $F(2,114) = 187.93$, $MSE = 0.20$, $p < 0.01$, and of Language, $F(1, 114) = 8.90$, $MSE = 0.01$, $p < 0.01$, as well as an interaction between the two, $F(2, 114) = 119.08$, $MSE = 0.13$, $p < 0.01$. The interaction arises because there was a greater effect of PD in the English model (81.0% accuracy) than in the Chinese model (89.4%), and the reverse pattern for SD, with a very modest effect in English (98.7% accuracy) but a large effect in Chinese (84.7%). Accuracy in the typical model was nearly perfect (> 99%) for both languages.

Reading deficits in English—To further investigate the patterns of reading disability resulting from particular patterns of deficit, we conducted a 2 (Regularity: regular-consistent, irregular-inconsistent) x 3 (Deficit: Typical, PD, SD) ANOVA analysis on maximum naming accuracy. The main effect of regularity was significant, $F(1, 57) = 54.91$, $MSE = 0.02$, $p < 0.01$, as was the main effect of deficit, $F(2, 57) = 26.74$, $MSE = 0.05$, $p < 0.01$, and the interaction of the two, $F(2, 57) = 14.41$, $MSE = 0.01$, $p < 0.01$. As seen in Figure 3, regular-consistent words were read more accurately than irregular-inconsistent words, and the Typical model's overall performance (100%) was significantly better than the PD model (93.0%), $p < 0.01$, and marginally better than the SD model (97.7%), $p = 0.056$. The interaction between deficit and stimulus condition arose because performance on all items was impaired in the PD model, whereas the SD model was impaired only in irregular word reading. In the Typical model, all words were named accurately. Semantic impairment had no impact on the regular-consistent items (100% accuracy), but resulted in reduced accuracy for the irregular-inconsistent items (95.4%). In contrast, the PD model was impaired for both regular-consistent and irregular-inconsistent items (94.8% and 91.3% accuracy, respectively).

Poor nonword reading is a particular hallmark of phonological dyslexia in English, but the status of "non-character" reading in Chinese (i.e., whether it reflects normal reading processes or meta-linguistic guessing) is a topic of debate (Weekes, Yin, Su, & Chen, 2006; Shu et al., 2005). Because of the higher degree of arbitrariness in spelling-to-sound mappings, it is hard to create non-characters in Chinese. We therefore tested nonword reading in a separate set of statistical tests for English. Nonword reading was strongly influenced by deficit, $F(2, 57) = 44.09$, $MSE = 0.06$, $p < 0.01$. Post-hoc tests showed no effect of SD on nonword reading (86.2% accuracy, compared to 86.1% accuracy for the Typical model), $p = 0.86$, and a large effect of PD (77.0%), $p < 0.01$.

Reading deficits in Chinese—In parallel with the analysis of the English model, we conducted a 2 (Regularity: regular-consistent, irregular-inconsistent) x 3 (Deficit: Typical, PD, SD) ANOVA with maximum accuracy as the dependent variable in Chinese. The main effect of regularity was significant, $F(1, 57) = 124.23$, $MSE = 0.11$, $p < 0.01$, as was the main effect of deficit, $F(2, 57) = 111.93$, $MSE = 0.17$, $p < 0.01$, and the interaction of the two, $F(2, 57) = 96.12$, $MSE = 0.09$, $p < 0.01$. As seen in Figure 4, regular-consistent words were read more accurately than irregular-inconsistent words, and the Typical model's overall performance was significantly better than both the PD model, $p < 0.01$ and than the SD model $p < 0.01$. The interaction between deficit and stimulus condition arose because performance on the word classes was differentially impacted by semantic and phonological impairments. In the Typical model, all words were named accurately (100%). Semantic impairment influenced the naming accuracy more for irregular-inconsistent (79.75%) than regular-consistent (96.5%) words, $t(19) = 12.67$, $p < 0.01$. In contrast, the PD model performed equally poorly on both regular-consistent (90.5%) and irregular-inconsistent (89.0%) words, $t(19) = 1.55$, $p = 0.14$.

Simulating three cases of Chinese dyslexia—Shu et al. (2005) reported three cases of developmental dyslexia in Chinese, along with data on the children's phonological and semantic processing abilities. One case, Child L (age 9:0, male), was classified as surface dyslexic because of his relatively specific impairment on exception words. The two remaining cases (J, 10:8, and Q, 12:2, both male) were phonological dyslexics. One important feature of this study is that semantic and phonological processing skills were also tested independently. Child L's reading impairment was accompanied by frank impairments in morphological awareness, a meta-linguistic task used to assess semantic processing in Chinese readers (McBride-Chang, Shu, Zhou, Wat, & Wagner, 2003), but his performance on phonological awareness tasks was within normal range, whereas J and Q showed the converse pattern.

Simulation of case studies was undertaken by identifying a point in training at which the appropriate model (SD for Child L, PD for J and Q) achieved the same overall accuracy (on all test items) as the case being simulated. Data from all three are shown in Figure 5. The SD model attained overall performance of 44% after average 364K ($SD = 58$K) trials. At this point in training, the SD model's ability to read words was strongly influenced by stimulus regularity (61.25% accuracy for regular items vs. 50.63% for irregular, $t(19) = 7.64$, $p < 0.01$.

A diagnostic feature of phonological dyslexia in Chinese is the lack of a regularity effect, seen in both of the cases under consideration here. Child J's overall accuracy was 49%, which the PD model reached after 433K ($SD = 88$K) trials. At this point in training, the model exhibited no effect of regularity with 59.1% and 58.1% accuracy for regular and irregular characters, respectively, $t(19) = 1.00$, $p = 0.33$. Child Q's overall accuracy was 73%, which the PD model reached after 910K ($SD = 192$K) trials. At this point in training, the model did not show evidence for a regularity effect — 79.8% and 78.6% accuracy for regular and irregular items, $t(19) = 1.58$, $p = 0.13$.

Discussion

When the same functional architecture is trained to read English and Chinese, distinct patterns of typical and atypical development are observed across languages. Gross differences in the rate of learning of the two writing systems are clearly captured by the models, as are differences in the patterns of deficits observed in reading disability. Specifically, the same constitutive deficits (in phonological and semantic processing) have distinct effects that are language-specific, suggesting that these patterns are driven by statistical properties of the writing systems themselves, and not by differences in the basic architecture of reading across languages.

In both English and Chinese, phonological deficits have relatively broad effects, and are a key factor in predicting reading disability (McBride-Chang & Zhong, 2006; Snowling, 2000; Vellutino & Fletcher, 2005). This is captured in the overall pattern of effects in the two models. Further, language-specific features of phonological dyslexia are also observed. In English, children with phonological difficulties have particular difficulty with nonword pronunciation (Temple & Marshall, 1983; Castles & Coltheart, 1993; Manis et al., 1996). This was also found in the English model. In contrast, nonword reading is a difficult task for even skilled Chinese readers, and is rarely tested in development, but there is a specific pattern that is a hallmark of phonological dyslexia – the reduced size of the regularity effect observed by Shu et al. (2005). This effect was also captured in the model.

Semantic deficits had strikingly different effects across writing systems. In English, semantic support is mainly necessary for irregular-inconsistent items, and deficits in semantic processing have relatively specific effects on these items (Castles & Coltheart, 1993, 1996; Manis et al., 1996). In Chinese, in contrast, effects of semantic deficits are quite general, impacting both regular-consistent and irregular-inconsistent items nearly equally. Again, this is consistent with case observations from Shu et al. (2005), and is also generally consistent with the relatively strong correlation of morphological awareness with reading ability (Ku & Anderson, 2003; McBride-Chang, Cho, et al., 2005).

Simulation 2: Modeling Chinese-English bilingualism

Although they shared many features, the models in Simulation 1 differed in important ways, because their phonological and orthographic representations were language-specific. Here we explore whether the same model, when trained to read both English and Chinese will show similar patterns of results to parallel models described in Simulation 1. We did this by training a single model with a single phonological output attractor, a single semantic system, and two orthographic input layers, one for each language.

Methods

Architecture—The architecture of the bilingual model (Figure 6) was modified from Simulation 1 to two orthographic layers as its input: one for Chinese and one for English. The orthographic representations themselves were the same as in Simulation 1, as were the semantic representation. In order to represent Chinese and English syllables in the same phonological layer, the number of units used to represent phonology was increased to accommodate the larger number of contrasts. Eight phoneme slots were used to encode the CCCVVVCC structure of syllables both for Chinese and English (where C is a consonant and V is a vowel). Like Zhao and Li’s (2009) PatPho system, a set of 37 phonological features were used to describe each phoneme: 20 features were used for consonants including 2 for voiced or not, 10 for place and 8 for manner of articulation (see Table 1); 17 features were used for vowels including 7 for height, 5 for backness, 3 for roundedness and 2 for long or short. All features were binary, taking values of 0 or 1. The 37 features per phoneme over eight phoneme slots yielded a total of 296 features. Including extra 4 features added to represent 5 tones in Chinese syllables, the number of features on the phonological layer was increased to 300 features.

Training and testing—Training was carried out in the same way as in the monolingual models. The learning rate, momentum, and learning algorithm of model, as well as the training materials were the same to simulation 1. We trained the reading model directly without pre-training the phonological attractor net. Simultaneous bilingualism was modeled by interleaved training on both languages, with an equal likelihood of training on either language on each trial. Twenty runs of normal model were trained at a different random seed for the initial randomization of weights and selection order of stimuli. As in Simulation 1, two types of decay were applied to either the hidden units from orthography to phonology (PD) or the hidden units from semantics to phonology (SD). Decay on each weight ω was reduced in magnitude according to the formula $\Delta\omega = -\omega \times \sigma$ where σ was the decay constant. Different decay values for each of twenty runs (both for PD and SD model) were used, varying from 0.25×10^{-5} to 5×10^{-5} in steps of 0.25×10^{-5}.

Testing also followed the same procedures as in Simulation 1.

Results

The model learned English more rapidly than Chinese, achieving 90% accuracy in 600K trials for Chinese, but only after 300K trials in English, $t(19) = 112.45$, $p < 0.01$. Further, semantic and phonological deficits had different effects on performance across languages, as revealed by a 3 (Deficit: Typical, Phonological, Semantic) x 2 (Language: English, Chinese) ANOVA with maximum accuracy as the dependent variable, and iterative runs of the model as a random variable. Main effects were observed for both deficit, $F(2, 38) = 111.57$, $p < 0.01$, and language, $F(1, 19) = 5.73$, $p < 0.05$, with an interaction between these two factors, $F(2, 38) = 96.29$, $p < 0.01$. The interaction arose because the effect of phonological deficits was greater for English (59.0% accuracy) than Chinese (71.7%) overall, whereas the effect of semantic deficits was greater in Chinese (90.1%) than English (98.8%).

Figure 7 shows the differential effect of impairments on stimulus types. In English, the effect of semantic impairments was specific to irregular inconsistent items (accuracy for these items was 97.3%, compared to 99.96% for regular consistent items). In contrast, semantic deficits impacted both types of items in Chinese (88.3% accuracy for I-I, 95.3% accuracy for R-C). Phonological deficits impacted both types of items in both languages. In English, the phonologically impaired model nonetheless read regular-consistent items more accurately than irregular items (76.2% vs. 65.7%, $t(19) = 13.48$, $p < 0.01$). In contrast, the model's performance in Chinese was impacted more equally for regular-consistent (73.5%) and irregular-inconsistent (70.3%), although the regularity effect was significant at this point in training, $t(19) = 3.22$, $p < 0.01$. Finally, we tested nonword reading for English only, and found substantial deficits for the phonologically impaired model (22.1% accuracy, vs. 59.7% in the typical model, $t(19) = 9.07$, $p < 0.01$). A modest (57.3%) but significant effect of semantic impairment was also found, $t(19) = 3.18$, $p < 0.01$.

Discussion

The pattern of results observed in a single model trained on both English and Chinese generally replicated the results from two models with the same architecture trained on each language separately. The results further support the view that differences between languages in typical and atypical patterns of learning can be explained without assuming differences in functional architecture between the languages. In particular, simulated deficits in the same set of hidden units had distinct effects on reading outcomes in English and Chinese.

General Discussion

The simulations presented here capture differences between two very different writing systems in the rate of acquisition as well as the differential sequelae of pre-existing deficits in phonological and semantic processing over the course of reading development. They also represent the first computational simulations of specific cases of developmental dyslexia in Chinese. That the same architecture and learning rules give rise to different outcomes across writing systems suggests that differences in typical and atypical reading development can be understood as resulting from statistical properties of the mapping from spelling to sound inherent to specific writing systems (Ziegler & Goswami, 2005), rather than structural differences in the reading system itself. This is in contrast to previous models of reading in Chinese, in particular (e.g., Perfetti et al., 2005), in which it is assumed that there is no sub-lexical spelling to sound conversion in Chinese, and implement a model in which phonological and semantic activation from print are mediated by lexical representations only (cf. models of alphabetic reading, which are equipped with an "assembled phonology" routine that permits sub lexical spelling to sound conversion). While it is possible in principle that such basic architectural differences could emerge as a result of statistical differences between the writing systems, it is difficult to characterize the results of the current simulations in those terms. Further, in Simulation 2, the same implemented model showed distinct patterns of typical and disordered reading ability across languages. This is the first simulation to address the acquisition of reading skill in a shared architecture for two languages, and is an important first step toward modeling the acquisition of biliteracy, because it addresses the first-order question of whether reading in the two languages can even be learned within the same basic processing and learning architecture.

When print-to-sound mappings are generally systematic, as in English, the contribution of semantic processing to reading aloud is limited largely to the pronunciation of the words that benefit least from regularities in the print-to-sound mapping (Strain et al., 1995, 2002). The mappings from print-to-sound in Chinese are much less reliable overall, with the result that the confluence of semantic and phonological processing is important to processing even words with relatively consistent mappings. These differences in the division of labor between phonological and semantic contributions to reading play out as differences in the manifestation of reading disability in the two writing systems. In English, semantic deficits have a relatively modest effect on reading in general, causing a highly selective impairment on exception words, whereas phonological impairments are more general and severe, and are marked by particular difficulty with generalization to nonwords. This disparity in the sequelae of phonological and semantic impairments has been exhaustively discussed in the current framework (Harm & Seidenberg, 2004). The division of labor for Chinese is more equitable, and therefore both semantic and phonological impairments have widespread effects (McBride-Chang, Cho, et al., 2005; Shu et al., 2006; Toyoda & Scrimgeour, 2009).

We have simulated differences between writing systems in the contribution of semantics to reading aloud, but a full account of the development of literacy across languages will require consideration of other factors. For example, mappings from print to meaning are actually much more systematic for Chinese characters than for monomorphemic words in English. This is likely to have consequences for reading development, and may be related to the finding that morphological awareness contributes more strongly to reading skill in Chinese than in other languages (McBride-Chang, Cho, et al., 2005). The current model uses a simplified random-bit semantic representation designed to capture only the fact that semantics can serve as an additional source of support for arriving at a pronunciation in the naming task. This was sufficient to model differential contributions of semantics and phonology to reading aloud, but mappings from spelling to meaning, and their impact on development across languages have begun to be studied behaviorally (Carlisle, 2000, 2003; Ku & Anderson, 2003; McBride-Chang et al., 2003; McBride-Chang, Wagner, Muse, Chow, & Shu, 2005) and would require a more elaborate model (e.g., with a similar architecture to Harm & Seidenberg, 2004) to simulate. The greater orthographic complexity of Chinese may also contribute to differences between languages. Ho, Chan, Chung, Lee, and Tsang (2007) describes orthographic processing deficits that may be directly related to reading disability, and these will be important to incorporate in a full account of reading acquisition across writing systems. Further, one major difference in patterns of brain activity during reading across writing systems – robust activation in the middle frontal gyrus for Chinese, which is not observed for alphabet languages – has been attributed to the increased demands on spatial processing involved in identifying Chinese characters (Tan, Laird, Li, & Fox, 2005; Bolger, Perfetti, & Schneider, 2005). One challenge for future modeling will be to incorporate an account of how orthographic knowledge emerges from more general aspects of visual processing (Polk & Farah, 1997, 1998).

In sum, the current simulations reflect an important step in cross-language modeling of the development of typical and disordered reading, and the modeling of biliteracy. The model instantiates two principles that have emerged from the study of reading acquisition, that the grain size of spelling-to-sound mappings determines the functional units that emerge in the spelling-to-sound system (Ziegler & Goswami, 2005), and that the "division of labor" between phonological and semantic contributions to the development of reading depends on the reliability of mappings among print, sound and meaning (Harm & Seidenberg, 2004; Seidenberg, 1993). As demonstrated in prior work (Yang et al., 2009; Zevin & Seidenberg, 2006), the model’s simulation of consistency effects depends on its ability to identify the appropriate grain sizes at which spelling-to-sound mappings exhibit regularities. Rather than assume a priori that a particular grain size should be privileged in the acquisition and use of these regularities, the model arrives at an appropriate level of description for each language as a result of the statistical structure of the input (Frost, in press). Similarly, the division of labor between phonological and semantic processes in reading plays out very differently in the two languages, even within the context of the same implemented model in Simulation 2. Specifically, the impact of semantic impairments is much greater and more general in Chinese than in English. Thus, the models presented here elucidate how these two principles can interact to produce different developmental trajectories across writing systems within a universal functional architecture.

Acknowledgments

This research was supported by Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning (CNKOPZD1005, J.Y.), NSF of China 31171077 (J.Y.), 30870758 (H.S.) and the Fundamental Research Funds for the Central Universities (H.S.), NIH R01 HD06579, R01-HD067364 (J.D.Z.) and R01 DC007694 (B.D.M.), NSF REC 0337765 (B.D.M.).

References

Andrews S, Scarratt DR. Rule and analogy mechanisms in reading nonwords: Hough dou peapel rede gnew wirds? Journal of Experimental Psychology: Human Perception and Performance. 1998; 24(4):1052–1086.

Bates E, Burani C, D’amico S, Barca L. Word reading and picture naming in Italian. Memory and Cognition. 2001; 29(7):986–999. [PubMed: 11820758]

Bolger DJ, Perfetti CA, Schneider W. Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. Human Brain Mapping. 2005; 25(1):92–104. [PubMed: 15846818]

Carlisle JF. Awareness of the structure and meaning of morphologically complex words: Impact on reading. Reading and Writing. 2000; 12(3–4):169–190.

Carlisle JF. Morphology matters in learning to read: A commentary. Reading Psychology. 2003; 24:291–322.

Castles A, Coltheart M. Varieties of developmental dyslexia. Cognition. 1993; 47(2):149–180. [PubMed: 8324999]

Castles A, Coltheart M. Cognitive correlates of developmental surface dyslexia: A single case study. Cognitive Neuropsychology. 1996; 13(1):25–50.

Coltheart M, Curtis B, Atkins P, Haller M. Models of reading aloud: Dual-route and parallel-distributed-processing approaches. Psychological Review. 1993; 100(4):589–608.

Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. DRC: A dual route cascaded model of visual word recognition and reading aloud. Psychological Review. 2001; 108(1):204–256. [PubMed: 11212628]

DeFrancis, J. Visible speech: The diverse oneness of writing systems. Honolulu, HI: University of Hawaii Press; 1989.

Ellis NC, Hooper AM. It is easier to learn to read in Welsh than in English: Effects of orthographic transparency demonstrated using frequency-matched cross-linguistic reading tests. Learning, Language and Development. 2001; 22(4):571–599.

Frost, R. Orthographic systems and skilled word recognition processes in reading. In: Snowling, MJ.; Hulme, C., editors. The science of reading: A handbook. Oxford: Blackwell; 2005. p. 272-295.

Frost R. Towards a universal model of reding. Behavioral and Brain Sciences. in press

Frost R, Katz L, Bentin S. Strategies for visual word recognition and orthographical depth: a multilingual comparison. Journal of Experimental Psychology: Human Perception and Performance. 1987; 13(1):104–115. [PubMed: 2951484]

Goswami U, Gombert JE, Barrera LFde. Children’s orthographic representations and linguistic transparency: nonsense word reading in English, French, and Spanish. Applied Psycholinguistics. 1998; 19(1):19–52.

Harm MW, Seidenberg MS. Phonology, reading acquisition, and dyslexia: Insights from connectionist models. Psychological Review. 1999; 106(3):491–528. [PubMed: 10467896]

Harm MW, Seidenberg MS. Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. Psychological Review. 2004; 111(3):662–720. [PubMed: 15250780]

Hetherington, P.; Seidenberg, MS. the 11th annual conference of the cognitive science society. USA:Ann Arbor, MI: Hillsdale, NY: Erlbaum; 1989. Is there "catastrophic interference" in connectionist networks?; p. 26-33.

Ho CS, Chan DW, Chung KKH, Lee S-H, Tsang S-M. In search of subtypes of Chinese developmental dyslexia. Journal of Experimental Child Psychology. 2007; 97(1):61–83. [PubMed: 17320097]

Ku Y-M, Anderson RC. Development of morphological awareness in Chinese and English. Reading and Writing. 2003; 16(5):399–422.

Language and Teaching Institute of Beijing Language College. Modern Chinese frequency dictionary. Beijing: Beijing Language College Publisher (In Chinese); 1986.

Lee C-Y, Tsai J-L, Su EC-I, Tzeng OJL, Hung DL. Consistency, regularity, and frequency effects in naming Chinese characters. Language and Linguistics. 2005; 6(1):75–107.

Leong, C. Reading in Chinese with reference to reading practices in Hong Long. In: Downing, JA., editor. Comparative reading: Cross national studies of behavior and processes in reading and writing. New York: Macmillan; 1973. p. 383-402.

Li, Y.; Kang, J. Information analysis of usage of characters in modern Chinese. Shanghai: Shanghai Education Publisher (In Chinese); 1993. Analysis of phonetics of the ideophonetic characters in modern Chinese; p. 84-98.

Liu Y, Shu H, Li P. Word naming and psycholinguistic norms: Chinese. Behavior Research Methods. 2007; 39(2):192–198. [PubMed: 17695344]

Malone K. Benjamin Franklin on spelling reform. American Speech. 1925; 1(2):96–100.

Manis FR, Seidenberg MS, Doi LM, McBride-Chang C, Petersen A. On the bases of two subtypes of development dyslexia. Cognition. 1996; 58(2):157–195. [PubMed: 8820386]

Marcus M, Santorini B, Marcinkiewicz MA. Building a large annotated corpus of English: The penn treebank. Computational Linguistics. 1993; 19:313–330.

Mattingly, IG. Linguistic awareness and orthographic form. In: Frost, R.; Katz, L., editors. Orthography, phonology, morphology, and meaning. Amsterdam: Elsevier Science; 1992. p. 11-26.

McBride-Chang C, Cho J-R, Liu H, Wagner RK, Shu H, Zhou A, et al. Changing models across cultures: Associations of phonological awareness and morphological structure awareness with vocabulary and word recognition in second graders from Beijing, Hong Kong, Korea, and the united states. Journal of Experimental Child Psychology. 2005; 92(2):140–160. [PubMed: 15904930]

McBride-Chang C, Shu H, Zhou A, Wat CP, Wagner RK. Morphological awareness uniquely predicts young children's Chinese character recognition. Journal of Educational Psychology. 2003; 95(4): 743–751.

McBride-Chang C, Wagner RK, Muse A, Chow BW-Y, Shu H. The role of morphological awareness in children's vocabulary acquisition in English. Applied Psycholinguistics. 2005; 26(3):415–435.

McBride-Chang, C.; Zhong, Y. Emergent literacy skills in Chinese. In: Li, P.; Tan, LH.; Bates, E.; Tzeng, OJL., editors. The handbook of east Asian psycholinguistics. Vol. Vol. 1. Cambridge, UK: Cambridge University Press; 2006. p. 81-89.Chinese

Mirkovic J, MacDonald MC, Seidenberg MS. Where does gender come from? evidence from a complex inflectional system. Language and Cognitive Processes. 2005; 20(1/2):139–167.

Nation K, Snowling M. Semantic processing and the development of word recognition skills: Evidence from children with reading comprehension difficulities. Journal of Memory and Language. 1998; 39:85–101.

Nation K, Snowling M. Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. Cognition. 1999; 70:B1–B13. [PubMed: 10193058]

Pearlmutter BA. Gradient calculations for dynamic recurrent neural networks: A survey. IEEE Transactions on Neural Networks. 1995; 6(5):1212–1228. [PubMed: 18263409]

Peng D, Yang H. The phonological processing of Chinese phonograms. Asia and Pacific Journal of Speech, Language and Hearing. 1997; 2(2):177–195.

Perfetti CA, Liu Y, Tan LH. The lexical constituency model: Some implications of research on Chinese for general theories of reading. Psychological Review. 2005; 112(1):43–59. [PubMed: 15631587]

Plaut DC. Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. Language and Cognitive Processes. 1997; 12:767–808.

Plaut DC, McClelland JL, Seidenberg MS, Patterson K. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. Psychological Review. 1996; 103(1): 56–115. [PubMed: 8650300]

Polk TA, Farah MJ. A simple common contexts explanation for the development of abstract letter identities. Neural Computation. 1997; 9(6):1277–1289. [PubMed: 9248063]

Polk TA, Farah MJ. The neural development and organization of letter recognition: Evidence from functional neuroimaging, computational modeling, and behavioral studies. Proceedings of the National Academy of Sciences. 1998; 95(3):847–852.

Saiegh-Haddad E, Geva E. Morphological awareness, phonological awareness, and reading in English-Arabic bilingual children. Reading and Writing. 2008; 21(5):481–504.

Seidenberg M. A connectionist modeling approach of the word recognition and dyslexia. Psychological Science. 1993; 4(5):299–304.

Seidenberg, M. Visual word recognition. In: Miller, J.; Eimas, P., editors. Handbook of perception and cognition. Vol. Vol. 11. San Diego: Academic Press; 1995. p. 138-163.

Seymour PHK, Aro M, Erskine JM. Foundation literacy acquisition in European orthographies. British Journal of Psychology. 2003; 94:143–174. [PubMed: 12803812]

Shu H, McBride-Chang C, Wu S, Liu H. Understanding Chinese developmental dyslexia: Morphological awareness as a core cognitive construct. Journal of Educational Psychology. 2006; 98(1):122–133.

Shu H, Meng X, Chen X, Luan H, Cao F. The subtypes of developmental dyslexia in Chinese: evidence from three cases. Dyslexia. 2005; 11(4):311–329. [PubMed: 16355750]

Shu H, Peng H, McBride-Chang C. Phonological awareness in young Chinese children. Developmental Science. 2008; 11(1):171–181. [PubMed: 18171377]

Snowling, MJ. Dyslexia. 2nd. Oxford, UK: Blackwell; 2000.

Stanovich KE, Siegel LS. Phenotypic performance profile of reading-disabled children: A gression-based test of the phonological core variable-difference model. Journal of Educational Psychology. 1994; 86:24–53.

Strain E, Patterson K, Seidenberg MS. Semantic effects in single-word naming. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1995; 21(5):1140–1154.

Strain E, Patterson K, Seidenberg MS. Theories of word naming interact with spelling-sound consistency. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2002; 28(1): 207–214.

Tan LH, Laird AR, Li K, Fox PT. Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. Human Brain Mapping. 2005; 25(1):83–91. [PubMed: 15846817]

Taraban R, McClelland JL. Conspiracy effects in word pronunciation. Journal of Memory and Language. 1987; 26(6):608–631.

Temple CM, Marshall JC. A case study of developmental phonological dyslexia. British Journal of Psychology. 1983; 74(4):517–533. [PubMed: 6640236]

Toyoda E, Scrimgeour A. Common and script-specific awareness in relation to word recognition in English and Chinese. Language Awareness. 2009; 18(1):61–73.

Vellutino, FR.; Fletcher, JM. Developmental Dyslexia. In: Margaret, CH.; Snowling, J., editors. The Science of Reading: A Handbook. Oxford, UK: Blackwell Publishing; 2005. p. 362-378.

Venezky, RL. The structure of English orthography. The Hague: Mouton; 1970.

Venezky, RL. American way of spelling: The structure and origins of American English orthography. New York: The Guildford Press; 1999.

Weekes BS, Yin W, Su F, Chen MJ. The cognitive neuropsychology of reading and writing in Chinese. Language and Linguistics. 2006; 7(3):595–617.

White TG, Graves MF, Slater WH. Growth of reading vocabulary in diverse elementary schools: Decoding and word meaning. Journal of Educational Psychology. 1990; 82(2):281–290.

Xing H, Shu H, Li P. The acquisition of Chinese characters: Corpus analyses and connectionist simulations. Journal of Cognitive Science. 2004; 5(1):1–49.

Yang J, McCandliss BD, Shu H, Zevin JD. Simulating language-specific and language-general effects in a statistical learning model of Chinese reading. Journal of Memory and Language. 2009; 61(2): 238–257. [PubMed: 20161189]

Zevin JD, Seidenberg MS. Simulating consistency effects and individual differences in nonword naming: A comparison of current models. Journal of Memory and Language. 2006; 54(2):145–160.

Zhao X, Li P. An online database of phonological representations for Mandarin Chinese. Behavior Research Methods. 2009; 41(2):575–583. [PubMed: 19363200]

Zhu, X. Analysis of cueing function of phonetic components in modern Chinese. In: Yuan, X., editor. Proceedings of the symposium on the Chinese language and characters. Bejing, China: Guang Ming Daily Press; 1988. p. 85-99.

Ziegler JC, Goswami U. Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. Psychological Bulletin. 2005; 131(1):3–29. [PubMed: 15631549]

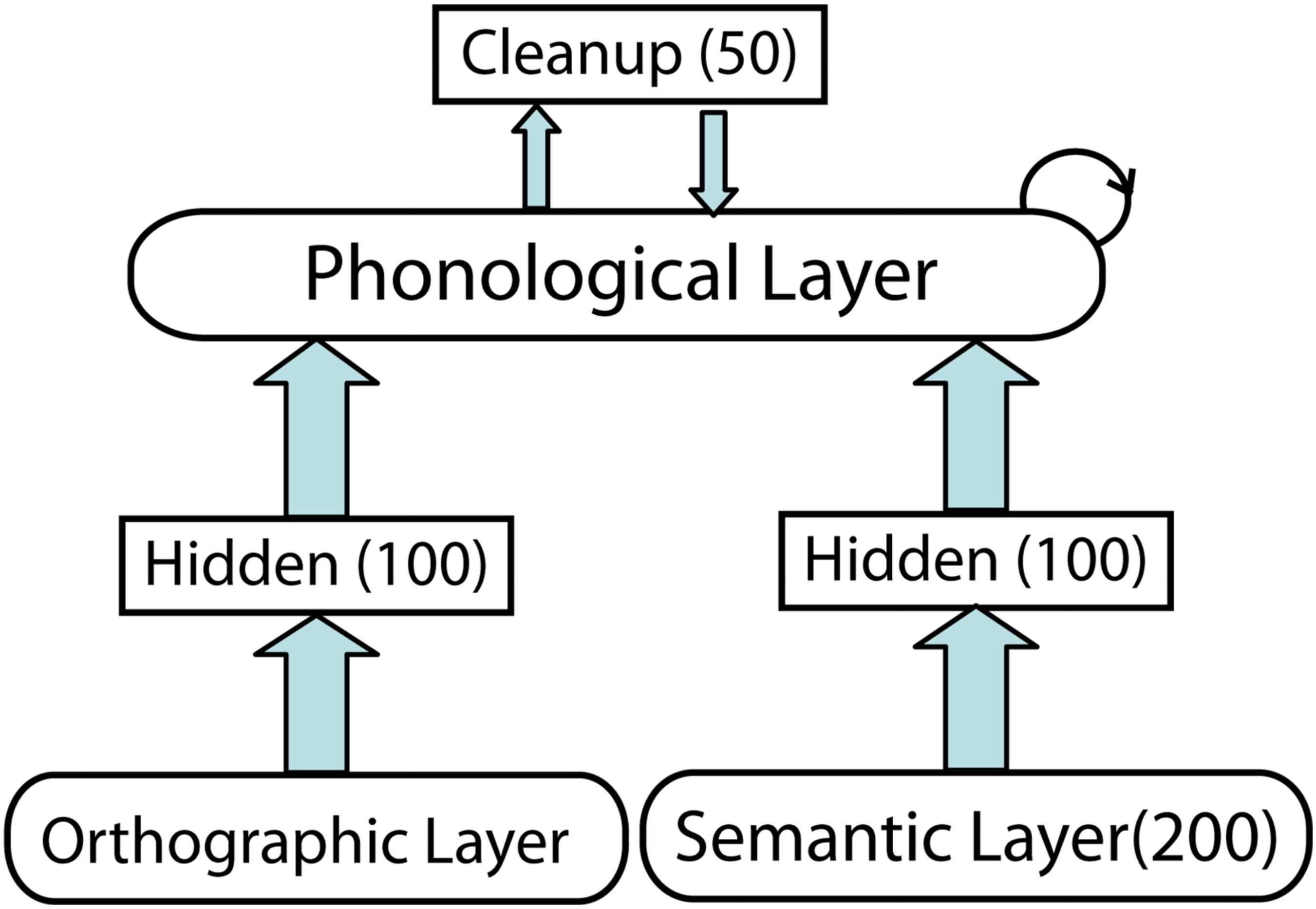


Figure 1.

Architecture of the monolingual Model.

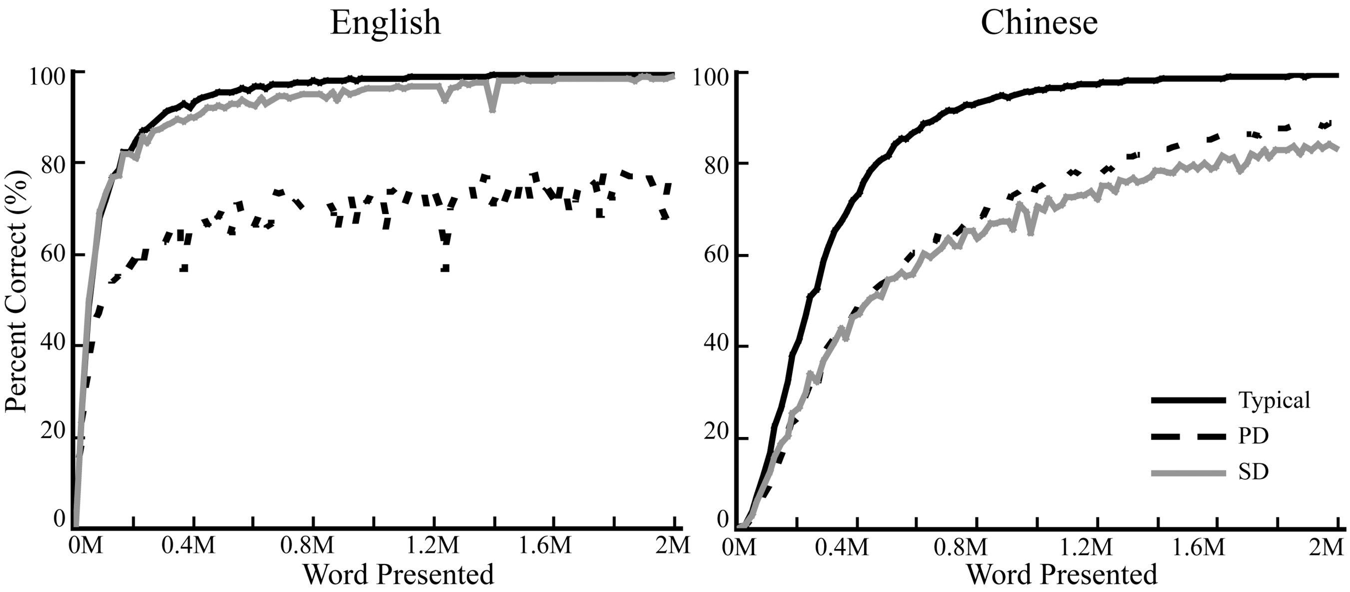


Figure 2.

Learning trajectories of English and Chinese models show differential effects phonological (dashed) and semantic (grey) impairment across languages.

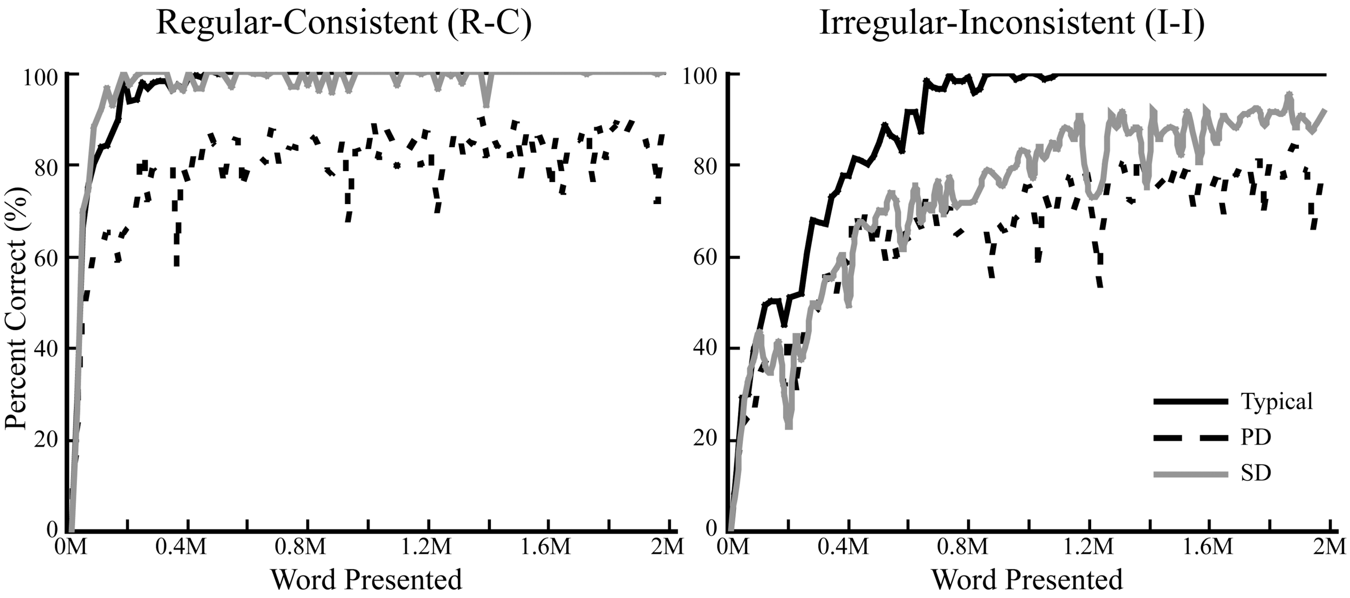


Figure 3.

Learning trajectories for different stimulus types in the English model. Regular consistent items are impacted only by phonological impairment (dashed line) whereas both phonological and semantic (grey line) impacted irregular-inconsistent items.

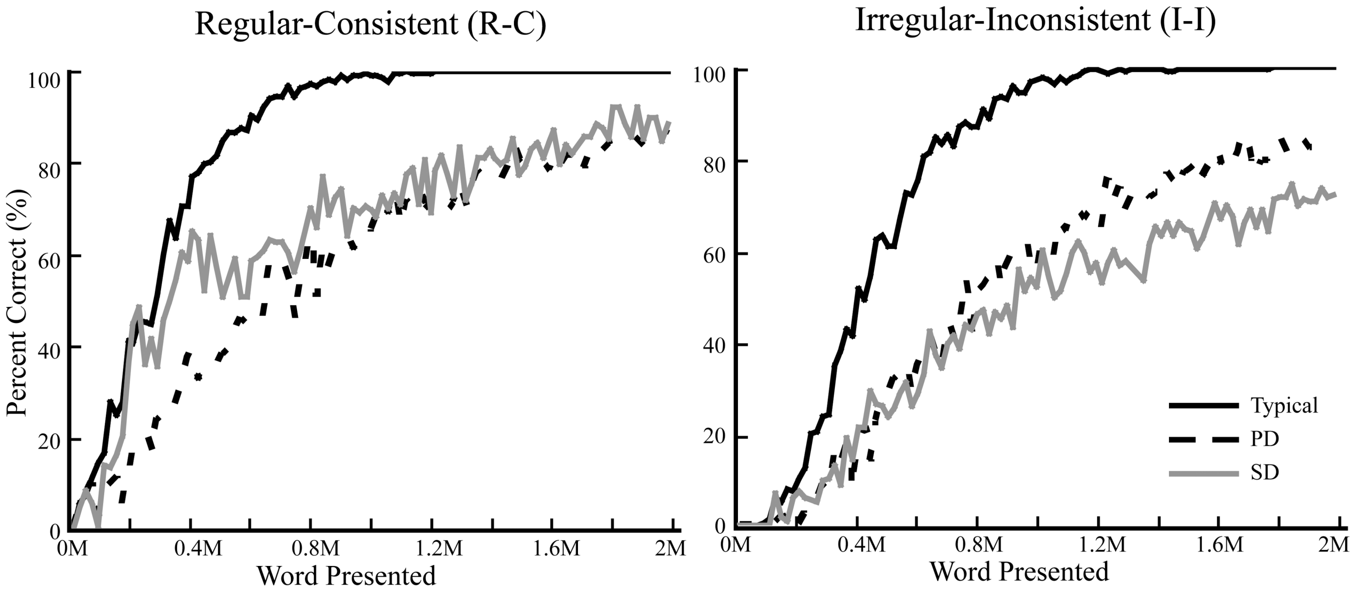


Figure 4.

Learning trajectories for different stimulus types in the Chinese model. Both regular-consistent and irregular-inconsistent items are impacted by both phonological (dashed) and semantic (gray) impairments.

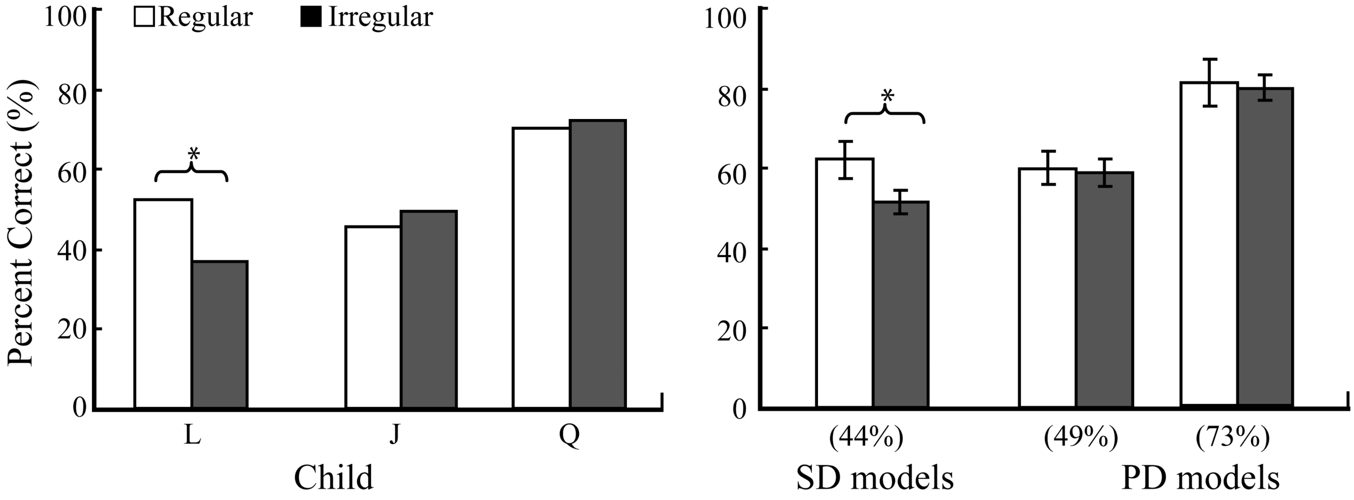


Figure 5.

Simulations of three case studies. Child L has semantic deficits, and shows a strong regularity effect, whereas children J and Q have phonological deficits and show no regularity effect.

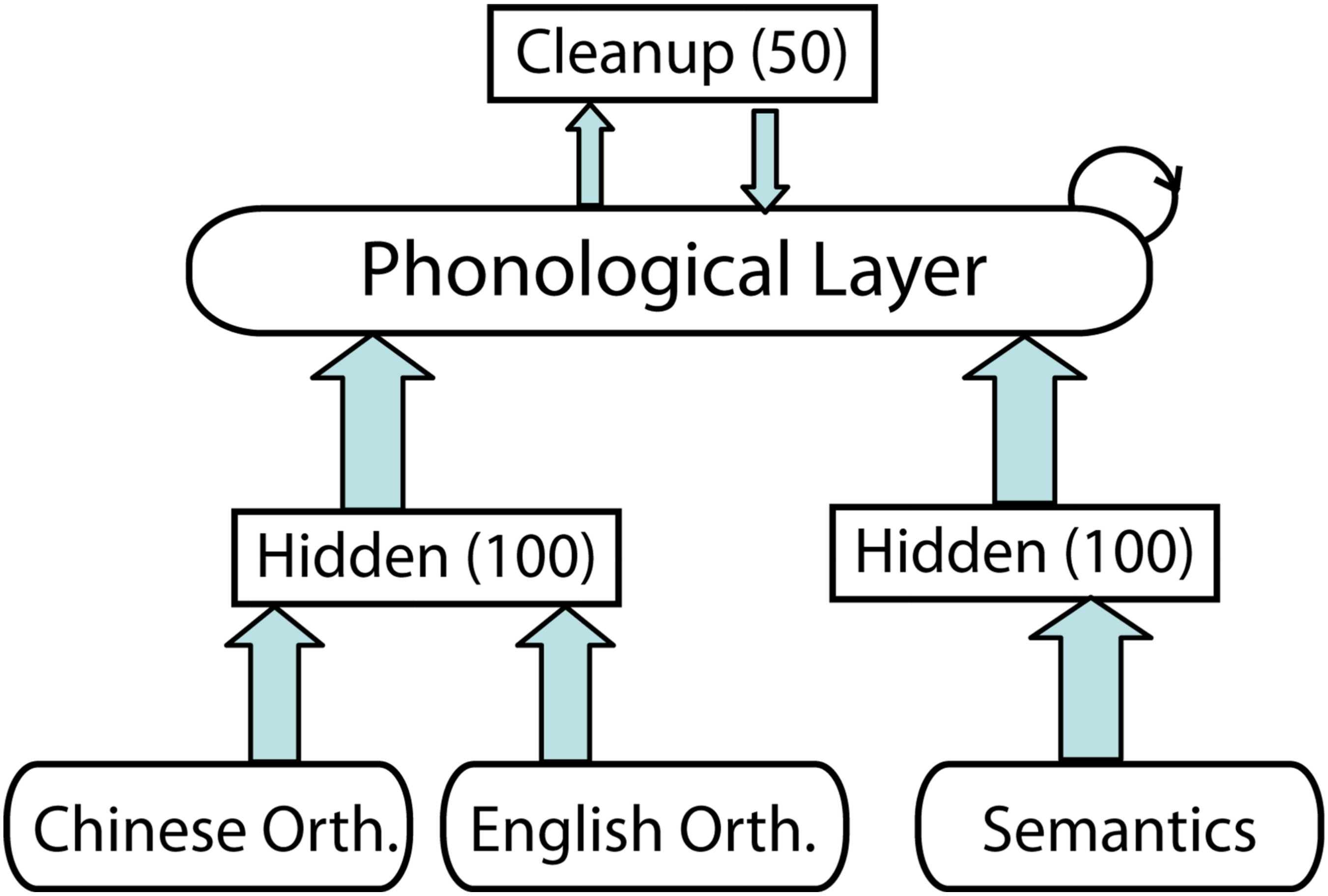


Figure 6.
Architecture of the bilingual model.

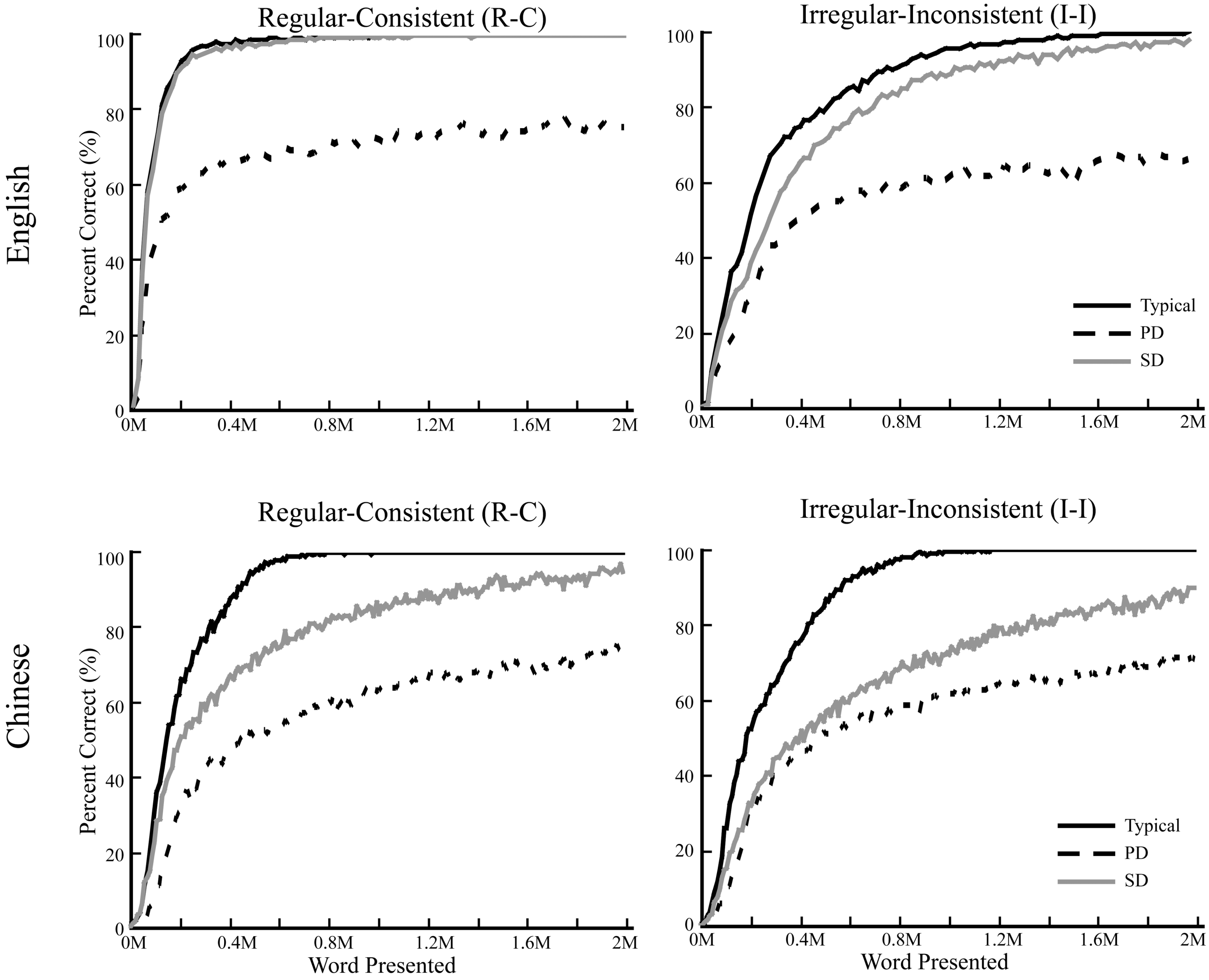
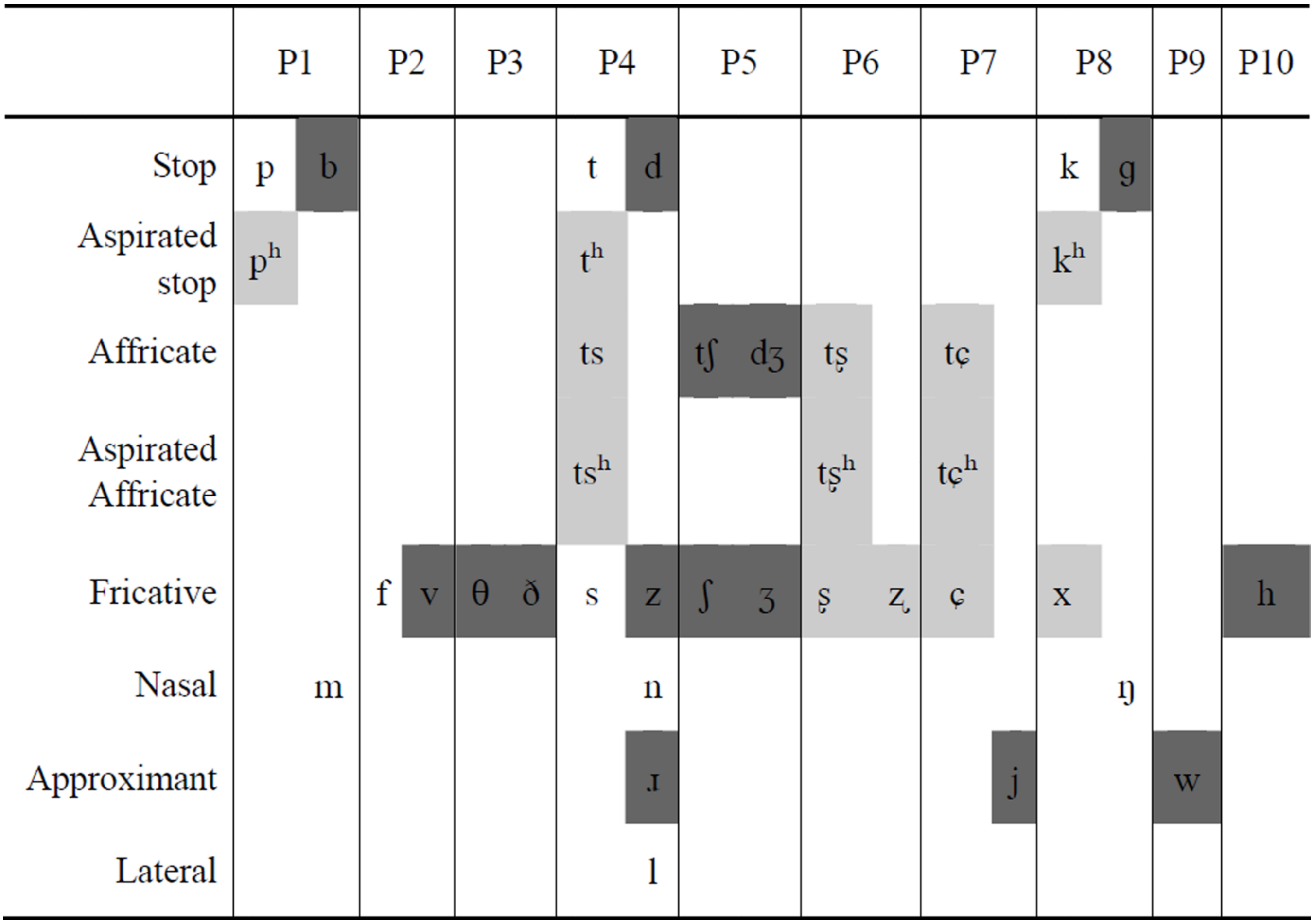


Figure 7.

Deficits on bilingual model showed differential effect of impairments on stimulus types.

Table 1

Phonetic features for consonants in English and Chinese



	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Stop	p b			t d				k g		
Aspirated stop	pʰ			tʰ				kʰ		
Affricate				ts	tʃ dʒ	tʂ	tɕ			
Aspirated Affricate				tsʰ		tʂʰ	tɕʰ			
Fricative		f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ɕ	x		h
Nasal	m			n				ŋ		
Approximant				ɹ			j		w	
Lateral				l						

Notes: P1, Bilabial; P2, Labio-dental; P3, Dental; P4, Alveolar; P5, Post-Alveolar; P6, Retroflex; P7, Palatal; P8, Velar; P9, Labio-velar; P10, Glottal. Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded phonemes in light grey only for Chinese, the dark phonemes only for English, unshaded phonemes present in both languages.