Published online September 2014, *Scientific Studies of Reading*

Phonological Representations and Early Literacy in Chinese

Joanna C. Kidd

Kathy Kar-Man Shum

Connie Suk-Han Ho

Terry Kit-fong Au

University of Hong Kong

We are grateful to the children, parents, and schools for their support, to the University of Hong Kong for a postdoc fellowship (Kidd) and research funding (Au), and to a team of dedicated research assistants for data collection and coding. We thank K.K. Luke for access to his Cantonese spoken words corpus, Johnson Li for help with data analysis, and Karen Ravn for helpful comments on the manuscript. Corresponding author: Terry Kit-fong Au, Department of Psychology, the University of Hong Kong, Pokfulam Road, Hong Kong; terryau@hku.hk

Abstract

Phonological processing skills predict early reading development, but what underlies developing phonological processing skills? Phonological representations of 140 native Cantonese-speaking Chinese children (age 4 to 10) were assessed with speech gating, mispronunciation detection, and nonword repetition tasks; their nonverbal IQ, reading, and phonological processing were assessed with standard tests. Results indicated that even without explicit script-sound correspondence at the phonemic level in Chinese orthography, young Chinese speakers developed representations segmented at this level, and such representations were more fine-grained for older children. Further, the quality of kindergarteners' phonological representations (specified by sensitivity to mispronunciation in lexical judgment) significantly predicted their emergent reading abilities, and this relation was fully mediated by phonological processing skills, with rapid naming showing the strongest mediation effect. Such mediation was no longer found with the primary-school sample, suggesting plausible developmental changes in the relations between phonological representations, phonological processing, and reading during early reading development.

Phonological Representations and Early Literacy in Chinese

Phonological processing refers to a cluster of interrelated skills that are integral to detecting, identifying, and discriminating speech sounds. This cluster comprises: *phonological awareness* – the ability to access and reflect upon units of a spoken language, as in a syllable deletion or rhyme identification task; *phonological retrieval* (often known as *phonological recoding in lexical access*) – the ability to efficiently retrieve a lexical referent to a written code from storage, as in a rapid automatized naming task; and *phonological short-term memory* – the ability to maintain phonological information in short-term memory for online processing or storage, as in a nonword repetition or sentence repetition task (Wagner & Torgesen, 1987). Phonological processing skills constitute a reliable predictor of reading ability and developmental dyslexia (Snowling, 2000). But the basis for individual differences in phonological processing skills is not well understood.

The quality of *phonological representations* – the sound-based codes required to recognize, distinguish, and produce spoken words (Stackhouse & Wells, 1997) – is a candidate predictor (Claessen, Heath, Fletcher, Hogben, & Leitão, 2009; Elbro & Jensen, 2005). By *phonological representations* we refer to the codes upon which phonological processing skills act. The quality of these representations (i.e., their level of segmentation and/or specification) can arguably be assessed by speech gating, lexical judgment, and nonword repetition tasks directly, rather than by proxy through phonological processing tasks that reveal only how well the codes are used. This study explores the nature of phonological representations, the ways their structure and quality might change with age, and their potential for explaining individual differences in phonological processing and reading in Chinese.

Phonological Awareness and Reading

Phonological awareness is a key phonological processing skill. Pre-readers typically show awareness of syllables and onset-rimes across languages (Ziegler & Goswami, 2005) merely from exposure to language (Cheung, Chen, Lai, Wong, & Hills, 2001; Morais, Bertelson, Cary, & Alegria, 1986). By contrast, phonemic awareness and literacy seem to develop reciprocally. Illiterate adults (Loureiroa et al., 2004; Morais et al., 1986) and pre-literate children (Wagner et al., 1997) have great difficulty manipulating phonemes (e.g., deleting the ‘l’ sound from ‘fleet’ to say ‘feet’). This makes sense in alphabetic languages: analyzing words into fine-grained speech sounds (e.g., phonemes) enables mapping to graphemes (letters), which in turn facilitates phoneme-level understanding (Horlyck, Reid, & Burnham, 2012). Reading, then, can develop in a ‘self-teaching’ manner (Share, 1995). But what about non-alphabetic, logographic languages such as Chinese?

Phonological awareness does seem important in learning to read Chinese (Chow, McBride-Chang, & Burgess, 2005; Ho & Bryant, 1997b; Huang & Hanley, 1997; McBride-Chang & Ho, 2005; Perfetti, Cao, & Booth, 2013). Each syllable in Chinese corresponds to a morpheme and is represented via a logograph (a Chinese character). Around 80% of Chinese characters include a phonetic component (Ho, Ng, & Ng, 2003), offering reliable information to the pronunciation of the character – either the whole syllable or the rime – about 40% of the time (Chung & Leung, 2008; Shu, Chen, Anderson, Wu, & Xuan, 2003). Phonological awareness appears to be important for learning such script-sound regularities in Chinese (Ho & Bryant, 1997a,b; Hu & Catts, 1998; McBride-Chang & Kail, 2002), yet it does not correlate significantly with word reading in Chinese when rapid naming, orthographic skills, and morphological awareness are taken into consideration (Yeung et al., 2011).

While phoneme-onset awareness is important for reading English, syllable awareness suffices for reading Chinese (McBride-Chang, Bialystok, Chong, & Li, 2004; Perfetti, Cao, & Booth, 2013). Some Chinese readers are nonetheless sensitive to phoneme-onset and -coda, but they typically have learned an alphabetic script (e.g., pinyin) or a phonetic-symbol system (e.g., zhuyin-fuhao in Taiwan) for Mandarin Chinese (Cheung et al., 2001). Mandarin speakers’ developing phonological awareness, then, may actually depend on such alphabetic or phonetic scripts. At least that is what Ziegler and Goswami’s (2005) ‘grain size’ hypothesis would predict: orthographies or phonetic symbol systems coded at the phoneme level foster phonemic awareness. The question thus remains: does phoneme-level sensitivity occur in Chinese readers who have not learnt an alphabetic script? Cantonese Chinese is an important test case for studying phonological development relatively uncomplicated by phonemes explicitly encoded in its writing system; no phonetic script is taught in schools in Cantonese-speaking communities such as Hong Kong, so any phoneme-level sensitivity cannot be attributed to orthographic clues.

Development of Phonological Representations

To what extent does developing phonological awareness reflect changes in underlying phonological representations? According to the 'lexical restructuring' view, such representations start out crude and holistic, with little inter- and intra-word differentiation. But as vocabulary grows, and more words are stored in the mental lexicon, a more refined and systematic phonological organization is needed. Whole-word representations become increasingly segmented, rendering the phonemes more accessible (Fowler, 1991; Metsala & Walley, 1998). Even words with extensive acoustic overlap become readily distinguished as high-quality speech-sound codes develop. Such restructuring likely proceeds from the onset-syllable level (Jusczyk, 1997) to the phoneme level. The more segmented a person's phonological representations, the less input the person should need to identify a word.

Evidence for lexical restructuring has come from studies of partial word recognition. Fernald, Swingley, and Pinto (2003) examined infant eye-gaze to familiar pictures upon hearing corresponding whole or partial words. Even 18-month-olds could recognize partial words, and one- to two-year-olds with larger vocabularies (about 100 words) did so more accurately than those with smaller vocabularies (about 60 words). The speech-gating paradigm that quantifies the point within a spoken word when recognition is first possible (i.e., the isolation point, or IP; Grosjean, 1980) also revealed that phonological structures become increasingly segmented. In speech gating, each word is heard from its onset in accumulating segments (i.e., gates) until the whole word has been heard. Significantly, isolation points for adolescents and adults are typically much earlier – suggesting more segmented phonological representations – than those for five- to seven-year-olds (Elliott, Hammer & Evan, 1987; Walley, Michela, & Wood, 1995). Poorly segmented phonological representations may also mean making more, or detecting fewer, subtle mispronunciations (Elbro, 1998). If the underlying representation is of lesser quality, a variety of spoken word inputs could be a potential match. Likewise, a specific input might fail to match a lexical representation if it is inadequately specified. Indeed, mispronunciation detection improves with age (Bernthal, Greenlee, Eblen, & Marking, 1987; Walley & Metsala, 1990).

The Phonological Representations Hypothesis of Reading

Phonological processing skills should to some extent reflect the underlying quality of phonological representations; how well the representations can be processed depends on their integrity (Snowling, 2000). Well-structured representations specifying phonemic information should benefit all aspects of phonological processing (e.g., phonological awareness, retrieval, and short-term memory) and in turn facilitate reading; poor phonological representations can lead to lesser phonological processing skills, poor reading, and developmental dyslexia (Elbro, 1998). This hypothesis therefore traces a causal link from phonological representations to reading via phonological processing skills.

Children and adults with dyslexia do seem to have poorer phonological representations, as evidenced from a range of tasks. Compared to non-dyslexic counterparts, they make non-semantic naming errors more frequently for even familiar multi-syllabic nouns (Dietrich & Brady, 2001; Swan & Goswami, 1997a,b), and their pronunciation reveals less phonemic information in repetition and naming tasks when target words are presented with sparse phonological information (e.g., less likely to produce the complete word ‘crocodile’ when hearing only [ko:di]; Elbro & Jensen, 2005; Elbro, Nielson, & Petersen, 1994).

Kindergarteners who are poorer at detecting subtle mispronunciations – suggesting poor phonological representations – also tend to have less phonological awareness, and are at higher risk for becoming poor readers in second grade (Claessen et al., 2009). Detection of subtle mispronunciations therefore seems a good tool for examining how phonological representations relate to phonological awareness. Speech-gating has also revealed sub-par phonological representations in dyslexic children, who need more speech input than typically-developing children to identify words. Furthermore, isolation points of words (especially from low-density neighborhoods) correlate with measures of reading, nonword reading, and phonological awareness (Boada & Pennington, 2006; Bruno et al. 2007; Metsala, 1997b; but see Griffiths & Snowling, 2001). This difficulty extends to children with specific language impairment, who often have severe reading problems (Dollaghan, 1998; Wong, Kidd, Ho, & Au, 2010).

Note, however, while Wesseling and Reitsma (2001) in one study found kindergarten speech gating significantly predicting nonword repetition, vocabulary, phonological awareness, and decoding in first grade, their attempts at replication brought conflicting results. It therefore remains unclear whether speech gating is related to reading and phonological processing skills in non-clinical samples.

The Phonological Representations Hypothesis and the Chinese Language

If phoneme-level representational structure arises from, and is useful for, reading alphabetic scripts (e.g., Ziegler & Goswami, 2005), this structure might not be useful, efficient, or economical for a syllabic, logographic language like Chinese. Recall that each Chinese character (the basic orthographic unit of the language) maps onto a syllable and a morpheme.

Hu (2004) examined the development of phonological representations, phonological processing, and reading in Chinese. While generally consistent with lexical restructuring, the findings did not support Ziegler and Goswami's (2005) cross-linguistic predictions: in a bisyllabic nonword repetition task, Mandarin-speaking children's errors included both incorrect ordering of syllables and transposing phonemes within syllables. Such errors increased with age (from kindergarten to third grade). Hu (2004) suggested that phoneme-level segmentation enabled the transposition errors, and that the developmental increase in these errors suggested ongoing lexical restructuring. One complication: many Mandarin-speaking children begin learning a phonetic-symbol system (e.g., pinyin, or zhuyin-fuhao) from kindergarten on. Such exposure might account for the phonemic transposition errors observed by Hu (2004). Therefore, the present study focused on children in Hong Kong who speak Cantonese-Chinese and are not taught phonetic symbols of Cantonese either at school or at home.

Study Aims

This study focused on three questions. First, over what time course do phonological representations develop in Cantonese-Chinese – uncomplicated by any phonemic-scripts? We used a cross-sectional design to look broadly across development for the most striking age-related differences. Second, how do phonological processing and reading skills relate to phonological representations in Cantonese? Third, do relations among phonological representations, phonological processing, and reading depend on grade levels during early reading development?

Method

Participants

In Hong Kong, three years of kindergarten (K1, K2, K3, with K1 starting around age three) precede the primary-school years (P1, P2, etc.). Our sample (total N = 140) was recruited from two kindergartens and three primary schools: 28 K2 children (aged 4;4 to 5;5, M= 4;10); 25 K3 children (aged 5;4 to 6;4, M= 5;8); 22 P1 children (aged 6;4 to 7;5, M= 6;10); 21 P2 children (aged 7;4 to 8;6, M= 7;10); 22 P3 children (aged 8;3 to 9;2, M= 8;9); 22 P4 children (aged 8;10 to10;5, M= 9;9). The final sample represented children from a variety of socioeconomic backgrounds (based on parental report of maternal and paternal education levels). All children spoke Cantonese as a first language, were educated in Cantonese, and had been living in Hong Kong for at least the past three years. Participants had all passed an initial screen for visual or hearing impairments based on parental report (Purdy, Farrington, Moran, Chard, & Hodgson, 2002), as well as hyperactivity and other emotional or behavioral problems, assessed via a parental questionnaire.

Materials and Testing Procedures

Psychometric Tasks

Nonverbal IQ. Primary-school children took Raven's Standard Progressive Matrices; kindergarteners took Raven's Colored Progressive Matrices. Scoring was referenced to local norms obtained by the Hong Kong Education Department (test-retest reliability = .88; Raven, 1986).

Reading. (1) One-Minute Word Reading: Kindergarteners read aloud as many words as possible in one minute from a list of 80 simple two-character Chinese words (Ho, Leung & Cheung, 2011). Primary-school children read from a list of 90 two-character words adopted from the Hong Kong Test of Specific Learning Difficulties in Reading and Writing (HKT-SpLD; Ho, Chan, Tsang, & Lee, 2000). (2) Chinese Word Reading: In this un-timed task, children read from a list in ascending order of difficulty. The list included 35 single Chinese characters for kindergarteners (Ho et al., 2011) and 150 two-character words for primary-school children (the HKT-SpLD Chinese Word Reading subtest; Ho et al., 2000).

Phonological Retrieval. (1) Rapid Automatized Naming—Objects (RAN Objects): Children saw pictures of five familiar objects with two-syllable Chinese labels (sun, butterfly, airplane, fan, apple), each appeared four times randomly in a 5x4 matrix (Ho et al., 2011). (2) Rapid Automatized Naming—Digits (RAN Digits): This HKT-SpLD subtest (Ho et al., 2000) consisted of a 5x8 matrix, with 5 digits (2, 4, 6, 7, and 9) arranged in random order in each row. For both RANs, the task was to name the items in order as quickly as possible. The child's score was the average completion time for two trials.

Phonological Awareness. (1) Syllable Deletion (for kindergarteners): For each of the 12 three-syllable test-words, the experimenter pronounced the word and asked the child first to repeat it, then to omit a specified syllable and say the remaining two syllables (Ho et al., 2011). (2) Rhyme Awareness (for primary-school children): For each of the 18 items in this HKT-SpLD subtest (Ho et al., 2000), children saw three pictures and heard a one-syllable label for each via a CD. Two of the labels rhymed; children pointed to the two corresponding pictures. Due to the nature of the Chinese language and writing system, syllable and onset-rime awareness, instead of phonemic awareness, is important for learning to read Chinese. Hence, phonological awareness was measured by syllable awareness and rhyme awareness in this study.

Phonological Short-Term Memory. (1) Nonword Repetition (non-gap): Children listened to, and repeated, nine nonwords. Each nonword contained 2 to 6 syllables that were meaningful in isolation but meaningless in combination. Scores were given in real time and later checked for reliability against audio-recordings (Ho et al., 2011). The maximum score was 63 (1 point for each correctly repeated syllable; 1 point for each correct order of 2-syllable target sequence). (2) Nonword Repetition (gap): Primary-school children were also tested with gap syllables (i.e., legal syllables and lexical tones in illegal combinations). This HKT-SpLD subtest (Ho et al., 2000) – developed for assessing primary-school children – included 14 items with three to six 'gap' syllables. The maximum score was 112, with the same scoring procedure as the non-gap nonword repetition just described. In our analyses, we assumed that syllable errors reflect cohesive representations of Cantonese syllables, whereas sub-syllable level errors probably reflect segmental representations (after Hu, 2004).

Phonological Representation Tasks

The speech-gating and lexical judgment tasks were created and administered using custom-written Matlab programs on a PC laptop. For both, the auditory stimuli were recorded in Cantonese by a male native speaker in a sound-attenuated room, digitized at a sampling rate of 44.1 kHz, and stored as '.wav' files, which were then read into Matlab. Stimuli were presented via headphones (fitted with a microphone attachment) at a comfortable sound level.

Speech Gating. Sixteen Cantonese monosyllabic concrete nouns were used as stimuli (available upon request), evenly divided between high- and low-frequency words based on a corpus of about 151,000 spoken Cantonese words collected from spontaneous adult conversations and Hong Kong radio programs[1]. Words were deemed 'high-frequency' if they occurred more than 200 times in the corpus, 'low-frequency' if they occurred fewer than 10 times. All stimulus words had a CVC composition, and had very few, if any, homophones.

Neighborhood density was estimated for each word by hand as no published database detailing such density was available in Cantonese. We used traditional definitions of neighborhood density (i.e.,

[1] At the time of testing, no published corpus of spoken Cantonese with word frequency and neighborhood density information could be obtained. Our frequency estimates were based on an online corpus developed by Prof. K. K. Luke, then at the University of Hong Kong (information can be sought from kkluke@ntu.edu.sg).

the number of words that overlap with the target word by substituting the onset, nucleus, or final phonemes; Charles-Luce & Luce, 1990), and added lexical tone to this list because Cantonese-Chinese is a tonal language. Estimates ranged from 7 (low-density) to 22 (high-density); delineation of ‘high’ and ‘low’ was based on a median split across the 16 words. Within the high- and low-frequency groups, the number of words with high- and low-density neighborhoods was balanced as far as possible – within the constraints of inherent correlation between word frequency and neighborhood density in Cantonese. Target words ranged in duration from 400 to 900 msec; high- and low-frequency words did not differ significantly in length, neither did high- and low-density neighborhood words ($ts[14]<.28$, $ps>.05$).

The speech-gating task used a forward-gating, duration-blocked procedure (Griffiths & Snowling, 2001; Grosjean, 1980). The procedure was as follows: in the first block of trials, the first gates for all 16 target words were presented in random order. This first gate was 250msec in length. The second block re-presented the first gate for each word plus one additional gate (50msec in length), again in random order (thus, each word fragment in the second block was 250msec + 50msec in length). A further 50msec gate was added to each accumulated word fragment in each successive block until all 16 words had been presented in full. Because words differed in length, a different number of gates was required for each word. Once a word had been presented in full, subsequent blocks comprised only the remaining stimuli until completion (Griffiths &Snowling, 2001). The task took approximately 12 minutes, depending on speed of response, with breaks as required.

To make the task child-friendly, the trials showed pictures of aliens who purportedly had landed on Earth and were trying to learn Cantonese. Each word segment presented was supposed to be an alien trying to say a Cantonese word, and the child had to guess the word. It might be hard at first, the experimenter would say, but it would get easier as the alien practiced and got better. The experimenter initiated each trial via keyboard-press; sitting next to the child, the experimenter was able to monitor and encourage vigilance. After 500 msec, a speech segment was presented, giving the child a maximum of 10 seconds to respond. Responses were audio-recorded onto the computer for later coding. No feedback was given during the task, but children had two practice trials using high-frequency monosyllabic concrete nouns. Children were asked afterwards to label all the pictures to confirm that they knew the target words, and all did.

Lexical Judgment Task. This task was adapted from Claessen et al.'s (2009) paradigm. Each trial presented a picture of a familiar and easily depicted object on a computer screen, simultaneously with a verbal label. The label was either correct or a subtle mispronunciation. For each picture, four mispronunciations were created by substituting a different onset, nucleus, or final phoneme, or lexical tone – all were legal Cantonese words but not correct labels for the picture. For example, a child might see a leaf and hear either [jip6] (the word for 'leaf') or [jit6] (the word for 'hot'). A fifth mispronunciation was created by changing the lexical tone of the label to yield something that was not a legal Cantonese word (stimulus list available upon request). We created mispronunciations systematically to determine the locus of potential deficits in phonological representations.

There were 12 depicted objects, each could be labeled with a familiar monosyllabic noun. Children saw each picture 10 times – five times with the correct label, and once each with its five mispronunciations. Each of the randomly-ordered 120 trials began with a picture and, 500 msec later, a spoken label (correct or incorrect) embedded in a carrier phrase in Chinese that means, 'Is this (a) ___?' Once the child had indicated 'yes' or 'no' via a keyboard-press, the experimenter initiated the next trial. Responses and reaction times (measured from the end of the stimulus) were recorded by the computer. Children were asked to respond as accurately as possible to minimize speed-accuracy trade-off. The task took approximately four minutes to complete. No feedback was given. Several practice trials were given to ensure familiarity with the task. Children were asked afterwards to label the 12 pictures to confirm that they knew the target words, and all did.

General Procedure

Children first took the Raven's nonverbal IQ test. Those with an IQ one or more standard deviations below age norm average (Raven, 1986) were excluded from further testing. Fewer than 5 children were excluded as a result. The remaining testing was done in two sessions of approximately 45 minutes each, no more than three weeks apart.

Results

Psychometric Task Results

Table 1 presents the psychometric test results. There were no significant differences in age-normed nonverbal IQ across primary grade levels, but nonverbal IQ for the K3 participants was significantly higher than for the K2 children. As expected, reading and phonological processing performance improved with age (Pearson rs$\geq$.35, ps<.01) for all measures, except nonword repetition (gap).

Phonological Representation Task Results

Speech Gating

Children's audio-recorded responses at each gate were transcribed and scored as either correct or incorrect by two trained coders. We calculated the point of correct word recognition in two ways. For 'isolation point' (IP), we used the gate when a word was first correctly identified (Metsala, 1997b). For 'acceptance point' (AP), we used the gate when a word was correctly identified without subsequent change of mind (Grosjean, 1980). IP and AP were calculated as a proportion of the overall word length for each item. Inter-rater reliability was good (intra-class correlation, R = .88 and .83 for IP and AP respectively). The data of one child (from K3) were lost due to experimental error.

Figure 1 shows isolation points (IP, upper) and acceptance points (AP, lower) for the four combinations of high- versus low-frequency and high- versus low-density neighborhood words, as a function of grade level. Collapsing across grade levels, the mean IP was .68 (SD=.08) and occurred, on average, significantly earlier than the mean AP (.73; SD=.11; t[139]=6.47, p<.0001, Cohen's d=.61).

Effects of grade level, word frequency, and neighborhood density on IP and AP were investigated via two separate three-way ANOVAs, with group as a between-child factor (K2-P4; totaling 6), and word frequency (high vs. low) and neighborhood density (high vs. low) as within-child factors. Words were classified as HF-HD, HF-LD, LF-HD and LF-LD (H=high, L=low, F=frequency, D=density). The main effect of grade for both IP and AP was significant (*F*s[5,536]>4.28, *p*s<.001, η_p^2>.38). Tukey HSD post-hoc comparisons revealed IP was significantly later for K2 than for P3 and P4 (*p*s<.05, *d*s>.39) and significantly later for K3 than for every primary grade (*p*s<.001, *d*s>.40). AP was significantly later for K3 than for P2, P3, and P4 (*p*s<.02, *d*s>.40). The main effect of word frequency for both IP and AP was also significant (*F*s[1,536]>206, *p*s<.001, η_p^2>.28), with high frequency words recognized earlier than low frequency words. The main effect of neighborhood density was also significant (*F*s[1,536]>1.72, *p*s<.001, η_p^2>.003); high-density words were recognized, on average, sooner than low-density words.

For both IP and AP, grade level did not interact with word frequency (*p*s>.05) or neighborhood density (*p*s>.8), and there was no significant 3-way interaction among them (*p*s>.7). Importantly, the 2-way interaction between word frequency and neighborhood density was significant for both IP and AP (*Fs*[1,536]>26, *ps*<.001, η_p^2>.048). To follow this up, repeated-measure ANOVAs for item and subject analysis were conducted with the four word categories (HF-HD, HF-LD, LF-HD, LF-LD) as the within-child factor. The main effect of word category was significant for both IP (*F1*[3,414]=227, *p*<.001, η_p^2=.62) and AP (*F1*[3,414]=202, *p*<.001, η_p^2=.60) in subject analysis. However, it was significant in item analysis for AP (*F2*[3,12]=5.66, *p*=.012. η_p^2=.59) but not IP (*F2*[3,12]=2.01, *p*>.16, η_p^2=.34). The composite measure *minF'* (Clark, 1973) was significant for AP (*minF'*[3,12.7]=5.50, *p*=.013) but not for IP (*minF'*[3,12.2]=1.99, *p*>.16). These results suggest that AP (acceptance point: when a word was correctly identified without subsequent change of mind) may be a more sensitive and robust measure than IP (isolation point: when a word was first correctly identified). One way to enhance generalization across items perhaps is to add items to each word category in future research. Post-hoc tests revealed AP was earliest for the HF-HD words, followed by HF-LD words, then LF-LD, and finally LF-HD words. All pairwise comparisons were significant (*p*s<.001, *d*s>.44).

Lexical Judgments

Lexical judgments were analyzed using signal detection (Macmillan & Creelman, 1991), focusing on children's ability to identify accurately pronounced words (hits) and reject mispronounced words (correct rejections), and their reaction times. Sensitivity to mispronunciation was estimated in units of d-prime (*d'*); how well an individual determined whether the stimulus did/did not match the corresponding phonological representation (with poor ability to do so likely indicating poor quality of the corresponding representation, Claessen et al., 2009). Larger value in d-prime indicates greater sensitivity. This index is useful because it is not criterion dependent. That is, it estimates sensitivity regardless of an individual's bias to respond either 'yes' or 'no'. The beta ratio estimated the criterion; the extent to which a 'yes' response was more/less probable than a 'no' response. A beta-ratio of 1 indicates no bias, less than 1 indicates a liberal 'yes' bias, and more than 1 indicates a conservative 'no' bias (Stanislaw & Todorov, 1999). Scores converting to $z=\pm3.29$ were considered outliers (corresponding to $p<.001$; Tabachnick & Fidell, 1996). If two or more scores per child were identified as outliers, the entire data set for that child was excluded. The data for five children (two K2, one K3, two P1) were thus eliminated.

Table 2 shows a trend towards ceiling as a function of grade level for hits, but not for correct rejections. Overall, mean correct rejection rate was significantly lower than mean hit rate ($F[1,134]=269$, $p<.001$, $\eta_p^2=.67$). Most children seemed to have a 'yes' bias, yielding a beta ratio of less than 1. The beta ratio correlated negatively with grade level (Table 2), suggesting that older children adopted a less stringent criterion on this task. Children across grades generally showed good sensitivity; the average score exceeded 2.6 for even the youngest children.

Lexical judgment improved with age for most measures, particularly for sensitivity, hits, and correct rejections, as indicated by their significant correlations with age (Table 2). Multivariate ANOVA treating the five task variables (hits, correct rejections, hit RT, correct-rejection RT, and sensitivity) as dependent variables, and grade level as the fixed factor, revealed a significant effect of grade level for all variables except hit RT (hits, $F[5,129]=7.05$, $p<.001$, $\eta_p^2=.22$; correct rejections,

$F[5,129]=6.85$, $p<.001$, $\eta_p^2=.21$; hit RT, $F[5,129]=1.59$, $p>.05$; correct-rejection RT, $F[5,129]=3.50$, $p=.005$, $\eta_p^2=.12$; sensitivity, $F[5,129]=10.7$, $p<.001$, $\eta_p^2=.29$). Post-hoc tests (ps<.05) revealed: K2 children had significantly fewer hits than P3 and P4 children, and K3 children had significantly fewer hits than P2, P3 and P4 children; K2 children had significantly fewer correct rejections than children in every primary-school grade; both K2 and K3 children had significantly lower sensitivity to mispronunciation than children in every primary grade; K2 children were significantly slower at making correct rejections than P3 and P4 children. Note that children in the higher grade levels were faster and more accurate than those in the lower grades (hence no evidence of speed-accuracy trade-offs). Thus, kindergarteners' lower hit and correct-rejection rates were probably not due to impulsive responding.

Correct-rejection scores suggested that mispronunciation detection depends on the within-word position of the mispronounced element (Figure 2). Mispronunciations in the initial and nucleus position yielded near perfect performance across grade levels. Detection was worse for final-phoneme substitution and legal-tone alteration, and illegal-tone mispronunciations elicited the worst performance. Repeated-measure ANOVA with mispronunciation type as the within-child variable (five levels) and grade level as the between-child variable revealed a significant main effect of mispronunciation type ($F[4,512]= 312$, $p<.001$, $\eta_p^2=.71$). Post-hoc tests revealed significant differences in the correct-rejection rate among all mispronunciation loci (ps<.001), except between initial and nucleus phonemes. Multivariate ANOVA with five mispronunciation loci as the dependent variables and grade level as the fixed factor revealed a significant main effect of grade level for all mispronunciation loci (initial, $F[5,128]=4.55$, $p=.001$, $\eta_p^2=.15$; nucleus, $F[5,128]=4.84$, $p<.001$, $\eta_p^2=.16$; final, $F[5,128]=5.21$, <.001, $\eta_p^2=.17$; legal tone, $F[5,128]=4.96$, $p<.001$, $\eta_p^2=.16$; illegal tone, $F[5,128]=2.34$, $p=.045$, $\eta_p^2=.08$). Post-hoc analyses (ps<.05) indicated K2 children made significantly fewer correct rejections than all other grade levels for both initial- and final-phoneme mispronunciations, significantly fewer than P1, P3 and P4 children for nucleus mispronunciations, and significantly fewer than P2, P3 and P4 children for legal-tone mispronunciations. K3 children made significantly fewer correct rejections for legal-tone mispronunciations than P3 and P4 children.

Nonword Repetition

Three trained coders transcribed and coded children's audio-recorded nonword repetitions (nongap). Seven recordings were missing (one K2, one K3, one P1, one P3, and three P4). Errors fell into six categories: 1. sub-syllable substitution – an element within a syllable (the onset, rime, or tone) was replaced with another; 2. sub-syllable substitution combined with syllable misordering – a sub-syllabic element was replaced with another, and that new syllable was moved to a new location within the multi-syllable word; 3. sub-syllable transposition—the onsets of two syllables were swapped (as in spoonerisms); 4. syllable substitution—a syllable was swapped for another not in the original stimulus string; 5. syllable addition—an additional syllable not in the original stimulus string was included; 6. syllable misordering—a syllable was moved to a new location within the string. The proportion of each error type as a function of the total number of errors was calculated for each child. Inter-rater reliability was acceptable, with intra-class correlations higher than .75 for all error categories. Note that inter-rater reliability was probably limited by the quality of the recording and by the difficulty agreeing on the 'correctness' of illegal tones.

Floor effects were noted for all categories of error except type 1. Accordingly, grade level effects were examined with the non-parametric Kruskal-Wallis test for each error type. The proportion of errors differed significantly between grades for syllable substitution (type 4) ($\chi^2(5)=11.2$, $p<.05$) but none of the other five types of error ($\chi^2(5)<3.67$, $ps>.60$). Post-hoc pairwise comparisons revealed that the K2 children (mean rank=82.4) committed significantly more syllable substitution errors than the P1 (mean rank=61.4; $p<.05$), P3 (mean rank=52.2, $p<.01$), and P4 participants (mean rank=57.5, $p<.05$); significant difference was also found between K3 (mean rank=75.1) and P3 children ($p<.05$). Collapsed across grade levels, sub-syllable substitution errors were by far the most common, accounting for 53.6% of all errors (mean number of errors, M=2.68). These were followed in order of prevalence by syllable substitution (17.6%, M=1.03), syllable misordering (17.0%, M=.80), sub-syllable substitution and misordering (7.5%, M=.39), sub-syllable transposition (4.0%, M=.27), and syllable addition (.3%, M=.02). A Friedman test comparing the error types (collapsed across all participants) was significant ($\chi^2(5)=307$, $p<.001$), with post-hoc tests showing significant differences ($ps<.05$) between all error types except between types 2 and 3, types 3 and 5, and types 4 and 6. Note that while the distributions for error types were generally not normal, the distribution for the overall score on nonword repetition (nongap and gap) was normal.

Relations among Phonological Representations, Phonological Processing, and Reading

Correlational Analyses

Does phonological processing mediate between phonological representations and reading, as predicted by the phonological representations hypothesis? To address this question, partial correlations among all measures of phonological representations, phonological processing and reading, controlling for the effect of age, were calculated separately for the kindergarten (K2 and K3; above the diagonal in Table 3) and primary-school children (P1-P4; below the diagonal).

For the kindergarteners, test measures for the quality of phonological representations (PR), i.e. speech gating (SG = average of IP and AP), mispronunciation sensitivity (MS) of the lexical judgment task, and the proportion of sub-syllabic errors in the nonword repetition task (NWR errors) did not significantly correlate with one another. Among the three measures, MS correlated significantly with RAN and nonword repetition, and with both Chinese word reading and one-minute reading, while SG and NWR errors did not show significant correlations with any of the phonological processing (PP) and reading tasks. Correlations between the phonological processing (PP) and reading tasks were all significant, except for nonword repetition, which correlated with neither Chinese word reading nor one-minute reading, after controlling for age.

For the primary-school children, weak but significant correlation was indicated between SG and NWR errors. None of the PR measures correlated significantly with any of the PP tasks, and only SG was significantly related to one-minute reading. Chinese word reading correlated significantly with all PP tasks, and most strongly with RAN. RAN was also the only PP measure that showed significant correlation with one-minute reading.

Multiple Mediation Analyses

To further explore the relationship between PR, PP, and reading, and to examine whether PP mediates the effect between PR and reading, multiple mediation models were tested separately for the kindergarten and primary-school children. Since the PR measures (SG, MS, and NWR errors) did not inter-correlate well, they were individually entered as the only independent variable in each model (Figure 3). Hence three analyses each were run for the kindergarten sample and the primary-school sample. The dependent variable was the composite score of the two reading tasks (Chinese word reading and one-minute reading), calculated by averaging their z-scores (Cronbach's alpha =.91). For the kindergarten group, RAN, syllable deletion, and nonword repetition (non-gap) were entered as the three proposed mediators in the model, representing the constructs of phonological retrieval, phonological awareness, and phonological short-term memory respectively. Similarly, RAN, rhyme awareness, and nonword repetition (average of the z-scores of non-gap nonword repetition and gap nonword repetition; Cronbach's alpha =.60) were the proposed mediators in the primary-school analyses. As chronological age was shown to correlate with most of the psychometric tasks in this study (Table 1), age was entered as the covariate in all the analyses to partial out its effect.

Testing a multiple mediation model allows for the determination of the *total indirect effect* attributed to all the mediators in the model, as well as teasing apart the relative magnitude of the *specific indirect effect* associated with each putative mediator in the presence of other mediators in the model (Preacher & Hayes, 2008). The bootstrapping procedure was conducted using the INDIRECT macro for SPSS developed by Preacher and Hayes (2008) to assess the indirect effects in these models. This approach is nonparametric and does not assume normality of the sampling distribution of indirect effects, and is recommended for small to moderate sample size. In this study, 10,000 bootstrap samples were generated through resampling, and the 95% bias-corrected (BC) confidence intervals (CIs) for the indirect effects were constructed based on their sampling distributions. Indirect effect was considered to be significant at the $p<.05$ level if its BC 95% CI did not include zero.

The path coefficients for the multiple mediation models and the bootstrap results for the mediation effects are presented in Tables 4 and 5 respectively. For the kindergarten group, the total indirect effect of MS on reading through the three PP measures as a whole was significant (BC 95% CI = [.01, .47]) based on bootstrapping (Table 5). Of the potential mediators, RAN was the only variable that significantly mediated the relationship between MS and reading (BC 95% CI = [.01, .34]) in the presence of other PP mediators in this model. Since the total effect of MS on reading was significant but the direct effect was not after controlling for the mediators (Table 4), the results indicated that the relation between mispronunciation sensitivity (MS) and reading ability for kindergarteners was fully mediated by phonological processing skills.

No significant mediation effects were found for any of the PR measures on reading for the primary-school children, as shown by the bootstrap results (Table 5). Interestingly, the direct effects of SG (B=1.99; $p<.01$) and NWR errors (B=-.38; $p<.05$) on reading were both found to be significant for this older group of participants in path analyses (Table 4). Additionally, RAN was a significant predictor of reading for both kindergarten and primary-school children in all the mediation models after accounting for the shared variance of all predictors (ps<.01). Phonological awareness also showed unique contribution to reading in the kindergarten sample, irrespective of the PR measure entered into the analysis (ps<.001). However, in the primary-school sample, unique contribution of phonological awareness to reading remained only for the MS model; it became insignificant in the context of the other two measures of phonological representations.

Discussion

This study set out to evaluate the phonological representations hypothesis of reading in Cantonese-Chinese speaking children learning to read a non-alphabetic script. We focused on the quality of phonological representations and its relation to phonological processing and reading skills at various ages to understand how these relations might be different as a function of development. Two accepted methods, namely lexical speech-gating and mispronunciation detection, were adapted for Cantonese to evaluate the underlying structure of phonological representations.

Group-wise Differences in Phonological Representations

According to the lexical restructuring model, phonological representations become increasingly segmented and structurally more fine-grained with development (Fowler, 1991; Metsala & Walley, 1998), and that is exactly what we found here. Speech-gating revealed that kindergartners needed longer word fragments than did third and fourth graders to identify Chinese words, suggesting a more segmented underlying structure of phonological representations among older children. Likewise in lexical judgment, primary-school children outperformed kindergarteners in judging whether words had been correctly pronounced, in reaction time for correct rejections of mispronunciations, and in the sensitivity measure. These findings converge to suggest that more accurate and refined representations of phonological information develop in early elementary years—without, it should be emphasized, the support of an alphabetic writing system based on grapheme-phoneme correspondence.

Lexical restructuring predicts stronger segmentation pressure for words that are frequently used and have many phonologically similar neighbors (Charles-Luce & Luce, 1990; Garlock, Walley, & Metsala, 2001; Metsala, 1997a). Our speech-gating data showed that from age four through ten years (K2 to P4), word recognition was, in fact, easiest for high-frequency words in high-density neighborhoods. These findings provide the first evidence that the segmentation of phonological representations in Cantonese Chinese, as in alphabetic languages, is driven by factors such as word frequency and neighborhood density.

These data also revealed an interaction between word frequency and neighborhood density, with low-frequency high-density words being the most difficult for children to recognize, regardless of age. Low-frequency words are seldom encountered and thus, for young children, not under strong pressure to segment. Should these low-frequency words also come from high-density word neighborhoods with many competitors for lexical selection, word recognition ought to be especially difficult. By contrast, high-frequency words are under greater pressure to segment because they are encountered often and need to be identified efficiently. If they are also in high-density neighborhoods, this pressure is intensified, resulting in even more efficient word recognition. This interaction between word frequency and neighborhood density replicates what Griffiths and Snowling (2001) found in English-speaking children. Together, these findings suggest that lexical restructuring occurs in a similar manner, and according to the same pressures, regardless of orthography.

Further support for the sub-syllable segmentation of phonological representations of Cantonese words came from errors in nonword repetition. Recall that syllable level errors can reflect cohesive representations of Cantonese syllables, whereas sub-syllable level errors probably reflect segmental representations. We found that sub-syllable substitution was by far the most common type of error in all grade levels, accounting for more than 50% of all errors observed. The three types of sub-syllable errors together accounted for about 65% of all errors, while syllable level errors made up the remaining 35%. These findings suggest that syllables are not the only processing unit present in the working memory of Chinese-speaking children, and that segmentation at the phoneme level probably underlies such sub-syllable errors (Hu, 2004).

One concern is that the sub-syllable level errors observed in nonword repetition may have merely resulted from children producing similar sounding real words. If so, it is not clear whether any segmental representation was involved. To check this possibility, we re-examined sub-syllable errors from about 25% of the children randomly selected in each grade. For the nonword (non-gap) task, among the re-examined errors, 6% (K2), 6% (K3), 5% (P1), 4% (P2), 14% (P3), and 7% (P4) involved real words. For the nonword (gap) task, the percentages were: 1% (P1), 0% (P2), 2% (P3), 0% (P4). The incidence of lexicalization for nonword repetition errors seems too low to cause concern.

Another concern is that we had conceptualized sub-syllable level errors as 'within a syllable,' which included lexical tone errors as well as segmental errors. The rationale was that lexical tone as well as phonemes are integral linguistic components of a Chinese syllable; indeed, all phonetic notations for a Chinese syllable include a symbol for tone along with phonemes (or onset and rimes). This was also why we included tones ('legal tone' and 'illegal tone') as possible loci, along with 'initial,' 'nucleus,' 'final,' of mispronunciation in the lexical judgment task (Figure 2). When we coded the errors in the nonword repetition tasks, we therefore coded segmental as well as lexical tone errors as sub-syllable level errors. Nevertheless, segmental errors are informative about sub-syllable segmentation. We therefore re-coded the sub-syllable errors from about 25% of the children randomly selected from each grade level. The analysis revealed that 100% of the sub-syllable errors involved segmental errors in the nonword (non-gap) repetition task for every grade (K2 to P4). For the nonword (gap) repetition task, 100% (P1), 85.5% (P2), 96.6% (P3), and 88.7% (P4) of the sub-syllable errors involved segmental errors. These results suggested that most of the sub-syllable level errors reported earlier in the main analysis should also involve segmental errors.

Even without learning written symbols that map onto Cantonese phonemes, then, Cantonese-speaking children can detect subtle mispronunciations at the phonemic level. Correct rejections of mispronunciations in the word-initial and -nucleus position are nearly perfect from age four onwards. They apparently do not rely on explicit sound-script mapping to develop phoneme-level phonological representations. Nonetheless, speech-gating and lexical judgment (or mispronunciation detection) data reveal that phonological representations in Cantonese Chinese do become increasingly accurate and well-specified from age four through ten. This result is entirely in keeping with the longstanding finding of Chaney (1992) that pre-readers of English (3-year-olds) are able to detect phoneme-level mispronunciations and are developing metalingusitic skills. Our results built on this finding to offer new insight into development after reading onset in children who read a nonalphbetic script, challenging the idea that phoneme-level specification is purely a function of orthographic 'depth'.

Relations between Phonological Representations, Phonological Processing, and Reading

The second focus for this study was the relations between phonological representations, phonological processing, and reading abilities in Cantonese, and how these may vary across grade levels during early reading development. Snowling's (2000) phonological representations hypothesis suggests that the quality of phonological representations may predict early reading abilities. Our correlational and mediation analyses for the kindergarten sample (ages four to six) provided evidence that the quality of phonological representations (specified by sensitivity to mispronunciation in the lexical judgment task) was related to the reading ability of emergent readers learning a non-alphabetic, logographic language. Moreover, this relation between phonological representations and reading was fully mediated by phonological processing skills, with rapid naming showing the strongest mediation effect.

However, such mediation was not found with the primary-school sample. Combined with the finding of direct effects of phonological representations (PR) on reading for the primary-school children, these results imply that the relation between PR and reading is no longer mediated through phonological processing skills, rapid naming in particular, but that other mediators not included in the models might be involved (Catts, Fey, Zhang, & Tomblin, 1999). Perhaps the quality of phonological representations underlying familiar and high frequency items (such as those used in RAN) is well established when children enter primary school (at around 6 years old), and hence no longer matters for the RAN tasks. However, accessing naming speed is still important for word reading in primary school independent of the phonological representations for the RAN test items, probably because rapid naming is associated with cognitive processes other than phonological access and retrieval. Indeed, RAN performance has been shown to relate to general processing speed (Powell, Stainthorp, Stuart, Garwood, & Quinlan, 2007) and visual processing speed (Stainthorp, Stuart, Powell, Quinlan, & Garwood, 2010), both of which are involved in reading.

The present study provides new insight into the underpinnings of phonological processing skills and the development of phonological representations in Cantonese Chinese, but a number of questions remain. Here we used a cross-sectional design to get a first overview of unknown developmental changes. Children were grouped by grade level, rather than chronological age, because grade levels probably matter more for our criterion variable (i.e., reading abilities). A longitudinal design tracking developmental trajectories could avoid the grade level versus age dilemma. It would also be more helpful for sorting out plausible causal directions of observed correlations. A case in point: longitudinal data would be valuable for evaluating the mediation models examined in this study. Nonetheless, this cross-sectional design uncovered likely developmental changes in phonological representations and their relations to phonological processing during early reading development. This study also gave clues as to when in development such changes occur.

A further issue concerns the types of tasks used to assess the underlying quality of phonological representations. We used speech-gating, lexical judgment, and nonword repetition errors because existing research in this area pointed strongly to their success in tapping the underlying code. However, indices of these tasks did not correlate highly. While this might point to weak construct validity, the analyses of Anthony et al. (2010) instead suggest that different tasks reflect different aspects of the phonological representation – how easily new representations are formed, how accessible they are, precision as reflected through articulation, and precision as reflected through speech perception. Probably no single task encompasses all of these. Future studies will do well to improve upon our measures (e.g., adding new measures and adding items to the speech-gating task) to uncover more robust age-related changes in the quality of phonological representations.

Summary and Conclusions

This study constitutes perhaps the first attempt to address fundamental questions about how Cantonese-Chinese-speaking children develop phonological representations and how the quality of their representations relates to higher level skills such as phonological processing skills and early reading development. In line with the lexical restructuring model, our results suggest that segmentation of phonological representations occurs at the phoneme level very early on, even for a non-alphabetic, logographic orthography. Our findings, however, do not support the idea that the ‘grain-size’ of phonological representations is determined mostly by the level of correspondence between phonology and orthography (e.g., phonemic vs. syllabic; Ziegler & Goswami, 2005).

These results further suggest that phonological processing (especially phonological retrieval) is a significant and statistically complete mediator between phonological representations (specified by sensitivity to mispronunciation in lexical judgment) and reading ability for kindergarteners in Chinese, lending plausibility to a causal link from phonological representations via phonological processing skills to reading ability. However, these relations may be developmentally limited: the mediating effect through phonological processing diminishes once phonological representations have developed to an adequate level (perhaps by first grade, according to this study), although the quality of these sound-based codes may still affect reading.

References

Anthony, J. L., Williams, J. M., Aghara, R. G., Dunkelberger, M., Novak, B. & Mukherjee, A. D. (2010). Assessment of individual differences in phonological representation. *Reading & Writing*, *23*, 969-994.

Bernthal, J. E., Greenlee, M., Eblen, R., & Marking, K. (1987). Detection of mispronunciations: A comparison of adults, normal-speaking children, and children with articulation errors. *Applied Psycholinguistics, 8*, 209-222.

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology, 95*, 153-193.

Bruno, J. L., Manis, F. R., Keating, P., Sperling, A. J., Nakamoto, J., & Seidenberg, M. S. (2007). Auditory word identification in dyslexic and normally achieving readers. *Journal of Experimental Child Psychology, 97*, 183-204.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, *3*, 331-361.

Chaney, C. (1992). Language development, metalinguistic skills, and print awareness in 3-year-old children. *Applied Psycholinguistics*, *13*, 485-514.

Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language, 17*, 205-215.

Cheung, H., Chen, H.-C., Lai, C. Y., Wong, O. C., & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. *Cognition, 81*, 227-241.

Chow, B. W.-Y., McBride-Chang, C., & Burgess, S. (2005). Phonological processing skills and early reading abilities in Hong Kong Chinese kindergarteners learning to read English as a second language. *Journal of Educational Psychology*, *97*, 81-87.

Chung, F. H. K., & Leung, M. T. (2008). Data analysis of Chinese characters in primary school corpora of Hong Kong and mainland China: Preliminary theoretical interpretations. *Clinical Linguistics & Phonetics, 22*, 379-389.

Claessen, M., Heath, S., Fletcher, J., Hogben, J., & Leitão, S. (2009). Quality of phonological representations: a window into the lexicon? *International Journal of Language & Communication Disorders, 44*, 121-144.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335-359.

Dietrich, J. A., & Brady, S. A. (2001). Phonological representations of adult poor readers: An investigation of specificity and stability. *Applied Psycholinguistics, 22*, 383-418.

Dollaghan, C. (1998). Spoken word recognition in children with and without specific language impairment. *Applied Psycholinguistics*, *19*, 193-208.

Elbro, C. (1998). When reading is 'readn' or somthn. Distinctness of phonological representations of lexical items in normal and disabled readers. *Scandinavian Journal of Psychology, 39*, 149-153.

Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, *46*, 375-384.

Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults: Evidence for deficits in nonword reading and in the phonological representation of lexical items. *Annals of dyslexia, 44*, 203-226.

Elliott, L. L., Hammer, M. A., & Evan, K. E. (1987). Perception of gated, highly familiar spoken monosyllabic nouns by children, teenagers, and older adults. *Attention, Perception, & Psychophysics*, *42*, 150-157.

Fernald, A., Swingley, D., & Pinto, J. P. (2003). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child development*, *72*, 1003-1015.

Fowler, A. E. (1991). How early phonological development might set the stage for phoneme awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 97-117). Hillsdale, N.J: Lawrence Erlbaum.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45*, 468-492.

Griffiths, Y. M., & Snowling, M. J. (2001). Auditory word identification and phonological skills in dyslexic and average readers. *Applied Psycholinguistics, 22*, 419-439.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics, 28*, 267-283.

Ho, C. S.-H., & Bryant, P. (1997a). Learning to read Chinese beyond the logographic phase. *Reading Research Quarterly, 32*, 276-289.

Ho, C. S.-H., & Bryant, P. (1997b). Phonological skills are important in learning to read Chinese. *Developmental Psychology, 33*, 946-951.

Ho, C. S.-H., Chan, D. W.-O., Tsang, S.-M., & Lee, S.-H. (2000). *The Hong Kong Test of Specific Learning Difficulties in Reading and Writing*. Hong Kong: Hong Kong Specific Learning Difficulties Research Team.

Ho, C. S.-H., Leung, M. T., & Cheung, H. (2011). Early difficulties of Chinese preschoolers at familial risk for dyslexia: deficits in oral language, phonological processing skills, and print-related skills. *Dyslexia*, *17*(2), 143-164.

Ho, C. S.-H., Ng, T.-T., & Ng, W.-K. (2003). A 'radical' approach to reading development in Chinese: The role of semantic radicals and phonetic radicals. *Journal of Literacy Research, 35*, 849-878.

Horlyck, S., Reid, A., & Burnham, D. (2012). The relationship between learning to read and language-specific speech perception: Maturation versus experience. *Scientific Studies of Reading*, *16*, 218-239.

Hu, C. F. (2004). The development of phonological representations among Chinese-speaking children. *Taiwan Journal of Linguistics, 2*, 103-130.

Hu, C. F., & Catts, H. W. (1998). The role of phonological processing in early reading ability: What we can learn from Chinese. *Scientific Studies of Reading, 2*, 55-79.

Huang, H. S., & Hanley, J. R. (1997). A longitudinal study of phonological awareness, visual skills, and Chinese reading acquisition among first-graders in Taiwan. *International Journal of Behavioral Development, 20*, 249-268.

Jusczyk, P. W. (1997). Finding and remembering words: Some beginnings by English learning infants. *Current Directions in Psychological Science, 6*, 170-174.

Loureiroa, C. S., Bragab, L. W., Souzab, L. N., Filhob, G. N., Queirozb, E., & Dellatolasc, G. (2004). Degree of illiteracy and phonological and metaphonological skills in unschooled adults. *Brain and Language, 89*, 499-502.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

McBride-Chang, C., Bialystok, E., Chong, K. K., & Li, Y. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology, 89*, 93-111.

McBride-Chang, C., & Ho, C. S.-H. (2005). Predictors of beginning reading in Chinese and English: A 2-year longitudinal study of Chinese kindergartners. *Scientific Studies of Reading*, *9*, 117-144.

McBride-Chang, C., & Kail, R. V. (2002). Cross-cultural similarities in the predictors of reading acquisition. *Child Development, 73*, 1392-1407.

Metsala, J. L. (1997a). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition, 25*, 47-56.

Metsala, J. L. (1997b). Spoken word recognition in reading disabled children. *Journal of Educational Psychology, 89*, 159-169.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word Recognition in Beginning Literacy*. N.J. Erlbaum: Hillsdale.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24*, 45-64.

Perfetti, C., Cao, F., & Booth, J. (2013). Specialization and universals in the development of reading skill: How Chinese research informs a universal science of reading. *Scientific Studies of Reading*, *17*, 5-21.

Powell, D., Stainthorp, R., Stuart, M., Garwood, H., & Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology*, *98*(1), 46-68.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, *40*, 879-891.

Purdy, S. C., Farrington, D. R., Moran, C. A., Chard, L. L., & Hodgson, S. A. (2002). A parental questionnaire to evaluate children's Auditory Behavior in Everyday Life (ABEL). *American Journal of Audiology*, *11*, 72-82.

Raven, J. C. (1986). *Hong Kong Supplement Guide to the Standard Progressive Matrices*. H.K.: HKSAR Government, Education Department.

Share, D. L. (1995). Phonological recoding and self-teaching: *sine qua non* of reading acquisition. *Cognition*, *55*, 151-218.

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development*, *74*, 27-47.

Snowling, M. J. (2000). *Dyslexia*: Blackwell Publishing.

Stackhouse, J., & Wells, B. (1997). How do speech and language problems affect literacy development? In C. Hulme & M. J. Snowling (Eds.), *Dyslexia: Biology, cognition and intervention* (pp. 182-211). London: Whurr Publishers Ltd.

Stainthorp, R., Stuart, M., Powell, D., Quinlan, P., & Garwood, H. (2010). Visual processing deficits in children with slow RAN performance. *Scientific Studies of Reading*, *14*(3), 266-292.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Beahvior Research Methods, Instruments & Computers, 31*, 137-149.

Swan, D., & Goswami, U. (1997a). Picture naming deficits in developmental dyslexia: The phonological representations hypothesis. *Brain and Language, 56*, 334-353.

Swan, D., & Goswami, U. (1997b). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *Journal of experimental child psychology*, *66*, 18-41.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using Multivariate Statistics* (3rd ed.). New York: Harper Collins College Publishers.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192-212.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33*, 468-479.

Walley, A. C., & Metsala, J. L. (1990). The growth of lexical constraints on spoken word recognition. *Attention, Perception, & Psychophysics*, *47*, 267-280.

Walley, A. C., Michela, V. L., & Wood, D. R. (1995). The gating paradigm: Effects of presentation format on spoken word recognition by children and adults. *Attention, Perception, & Psychophysics*, *57*, 343-351.

Wesseling, R., & Reitsma, P. (2001). Preschool phonological representations and development of reading skills. *Annals of Dyslexia, 51*, 203-229.

Wong, A. M.-Y., Kidd, J. C., Ho, C. S.-H., & Au, T. K. (2010). Characterizing the overlap between SLI and dyslexia in Chinese: The role of phonology and beyond. *Scientific Studies of Reading*, *14*, 30-57.

Yeung, P.-S., Ho, C. S.-H., Chik, P. P.-M., Lo, L.-Y., Luan, H., Chan, D. W., & Chung, K. K.-H. (2011). Reading and spelling Chinese among beginning readers: What skills make a difference? *Scientific Study of Reading, 15*, 285-313.

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

Table 1
Psychometric task results for kindergarten (K2 and K3) and primary-school (P1 to P4) children.

Test		**K2 (N=28)**	**K3 (N=25)**	**Pearson *r* with Age**	**Comparison**
IQ (Raven's Color)	*Mean*	114.29	120.32	0.41**	t(51)=-2.13*
	SD	11.33	9.24		
One Minute Reading	*Mean*	11.25	21.44	0.48***	t(51)=-3.42**
(max.=80)	*SD*	11.06	10.59		
Chinese Single Character	*Mean*	13.96	24.20	0.55***	t(51)=-3.82***
Reading *(max.=35)*	*SD*	11.43	7.39		
Syllable Deletion	*Mean*	7.36	10.00	0.41**	t(51)=-3.44**
(max.=12)	*SD*	3.38	1.94		
RAN objects	*Mean*	29.73	23.91	-0.52***	t(51)=3.23**
	SD	7.57	5.13		
Nonword Repetition	*Mean*	43.61	47.52	0.40**	t(51)=-1.91
(non-gap)	*SD*	8.56	5.90		
(max.=63)					

Test		**P1 (N=22)**	**P2 (N=21)**	**P3 (N=22)**	**P4 (N=22)**	**Pearson *r* with Age**	**Comparison**
IQ (Raven's	*Mean*	114.82	116.67	115.45	115.18	0.001	F[3,83]=0.09
Standard)	*SD*	11.72	13.27	14.20	11.08		
One Minute Reading	*Mean*	58.77	65.05	77.32	79.82	0.49***	F[3,83]=7.98***
(max.=80)	*SD*	18.26	14.18	18.73	14.53		
Chinese Word	*Mean*	67.50	83.19	96.55	116.50	0.58***	F[3,83]=13.80***
Reading (max.=150)	*SD*	25.67	26.61	29.01	23.20		
Rhyme Awareness	*Mean*	13.23	13.57	15.14	16.32	0.37***	F[3,83]=4.72**
(max.=18)	*SD*	3.74	3.19	3.27	1.81		
RAN digits	*Mean*	21.49	19.81	16.16	14.84	-0.53***	F[3,83]=10.88***
	SD	5.61	4.53	3.80	3.31		
Nonword Repetition	*Mean*	52.36	55.14	56.95	56.23	0.35**	F[3,83]=4.04*
(non-gap)	*SD*	5.21	5.87	3.53	3.88		
(max.=63)							
Nonword	*Mean*	78.91	82.05	85.14	84.09	0.17	F[3,83]=0.79
Repetition (gap)	*SD*	17.72	13.83	14.49	10.99		
(max.=112)							

*p<.05; **p<.01; ***p<.001. All scores are raw scores, except for IQ, which is given in standard scores based on a mean of 100 and a standard deviation of 15.

Table 2.
Descriptive statistics for the lexical judgment task across grade levels.

Measure		**K2 (N=26)**	**K3 (N=24)**	**P1 (N=20)**	**P2 (N=21)**	**P3 (N=22)**	**P4 (N=22)**	**Pearson *r* with Age**
Hits	*Mean*	56.85	56.17	58.00	58.43	59.18	59.32	0.42***
(max=60)	*SD*	3.08	3.12	2.51	1.99	0.73	0.84	
CR	*Mean*	47.73	50.58	52.60	52.24	53.68	53.77	0.41***
(max=60)	*SD*	7.95	3.55	2.37	3.82	2.06	1.69	
Hit RT	*Mean*	1.54	1.32	1.75	1.40	1.27	1.31	-0.06
(seconds)	*SD*	0.49	0.45	0.88	0.67	0.72	0.74	
CR RT	*Mean*	2.03	1.82	1.74	1.56	1.32	1.40	-.29**
(seconds)	*SD*	0.74	0.67	0.76	0.64	0.65	0.74	
Sensitivity	*Mean*	2.68	2.67	3.15	3.20	3.45	3.49	0.51***
(d')	*SD*	0.81	0.51	0.55	0.47	0.35	0.28	
Beta Ratio	*Mean*	0.39	0.53	0.35	0.31	0.23	0.22	-0.32***
	SD	0.28	0.36	0.27	0.27	0.09	0.13	

**$p<.01$; ***$p<.001$. CR = correct rejections; RT = reaction time

Table 3.
Partial correlations among measures of phonological representations, phonological processing, and reading after controlling for chronological age.

Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1. SG	-	.07	-.05	.14	-.03	-.20	-	-.12	-.12	-.13
2. MS	.01	-	-.20	-.29*	.22	.31*	-	.33*	.30*	.33*
3. NWR errors	-.22*	.15	-	.18	.17	.26	-	-.03	-.10	-.07
4. RAN	.10	-.18	.04	-	.03	-.08	-	-.40**	-.38**	-.41**
5. PA	.07	.04	-.11	-.18	-	.28	-	.40**	.45**	.44**
6. NWR (non-gap)	.02	.05	.03	-.29**	.17	-	-	.13	.10	.12
7. NWR (gap)	-.05	.02	.12	-.19	.25*	.44***	-	-	-	-
8. OMR	.26*	-.08	-.18	-.39***	.14	.11	.19	-	.86***	.97***
9. CWR	.00	.07	-.15	-.57***	.34**	.26*	.22*	.21	-	.96***
10. Reading	.17	-.01	-.21	-.62***	.31**	.24*	.27*	.79***	.77***	-

**p*<.05; ***p*<.01; ****p*<.001. Kindergarten (*N* = 46) and primary-school (*N* = 78) correlations are respectively above and below the diagonal. SG = average of IP and AP of the speech gating task; MS = mispronunciation sensitivity in the lexical judgment task; NWR errors = proportion of sub-syllabic errors in the nonword repetition task; PA = phonological awareness task; NWR = nonword repetition; OMR = one-minute reading; CWR = Chinese word reading; Reading = average of the z-scores of Chinese word reading and one-minute reading.

Table 4.
Unstandardized path coefficients (*B*) and standard errors (*SE*) for the multiple mediation models on the relationship between quality of phonological representations, phonological processing and reading, after controlling for age. Speech gating (SG), mispronunciation sensitivity in lexical judgment (MS), and proportion of sub-syllabic errors in nonword repetition (NWR errors) were the independent variable in three separate models. Symbols *a1*, *a2*, *a3*, *b1*, *b2*, *b3*, *c*, and *c'* represent the paths of the mediation models illustrated in Figure 3. *p<.05; **p<.01; ***p<.001.

Kindergarten: K2-K3 (*N*=46)

		SG		MS		NWR errors	
Path		***B***	***SE***	***B***	***SE***	***B***	***SE***
Total Effect (*c*)		-1.47	1.41	.39*	.17	-.19	.60
Direct Effect (*c'*)		-.78	1.19	.17	.16	-.22	.51
Mediators:							
RAN	a_1	11.17	10.33	-2.56	1.28	4.73	4.37
	b_1	-.06***	.02	-.05**	.02	-.06**	.02
PA	a_2	-2.22	4.84	.91	.61	2.64	2.02
	b_2	.14***	.03	.13***	.04	.14***	.04
NWR	a_3	-19.57	11.76	3.18*	1.41	9.07	4.90
	b_3	-.01	.01	-.01	.02	-.01	.01

Primary-school: P1-P4 (*N*=78)

		SG		MS		NWR errors	
Path		***B***	***SE***	***B***	***SE***	***B***	***SE***
Total Effect (*c*)		1.66	.85	.003	.17	-.44	.23
Direct Effect (*c'*)		1.99**	.65	-.18	.13	-.38*	.18
Mediators:							
RAN	a_1	4.86	5.69	-1.85	1.11	.58	1.58
	b_1	-.09***	.01	-.09***	.01	-.08***	.01
PA	a_2	3.05	4.01	.38	.78	-1.07	1.10
	b_2	.03	.02	.04*	.02	.04	.02
NWR	a_3	-.09	1.06	.06	.21	.23	.29
	b_3	.09	.07	.08	.07	.10	.07

Table 5.
Bootstrap results for the indirect effects of phonological representations on reading through phonological processing skills.

Kindergarten: K2-K3 (*N*=46)

	SG			MS			NWR errors		
	Point	BC 95%CI		Point	BC 95% CI		Point	BC 95% CI	
Indirect Effect	Estimate (SE)	Lower	Upper	Estimate (SE)	Lower	Upper	Estimate (SE)	Lower	Upper
Total	-.73 (.89)	-2.67	.94	.22 (.12)	.01	.47	.02 (.44)	-.77	1.02
RAN	-.62 (.58)	-1.95	.37	.13 (.08)	.01	.34	-.29 (.25)	-.97	.07
PA	-.34 (.68)	-1.80	.93	.12 (.10)	-.04	.34	.41 (.35)	-.15	1.24
NWR	.23 (.37)	-.14	1.64	-.03 (.05)	-.19	.03	-.10 (.17)	-.60	.11

Primary-school: P1-P4 (*N*=78)

	SG			MS			NWR errors		
	Point	BC 95%CI		Point	BC 95% CI		Point	BC 95% CI	
Indirect Effect	Estimate (SE)	Lower	Upper	Estimate (SE)	Lower	Upper	Estimate (SE)	Lower	Upper
Total	-.30 (.54)	-1.37	.77	.18 (.13)	-.06	.46	-.06 (.13)	-.31	.22
RAN	-.37 (.45)	-1.28	.51	.16 (.10)	-.02	.39	-.05 (.12)	-.29	.17
PA	.08 (.17)	-.15	.61	.01 (.04)	-.04	.12	-.04 (.04)	-.15	.01
NWR	-.01 (.11)	-.33	.17	.00 (.03)	-.03	.09	.02 (.04)	-.02	.17

Figure 1.
Speech-gating performance for each word type by group. Isolation Point (IP) is given above, Acceptance Point (AP) is given below. Error bars indicate SEM.

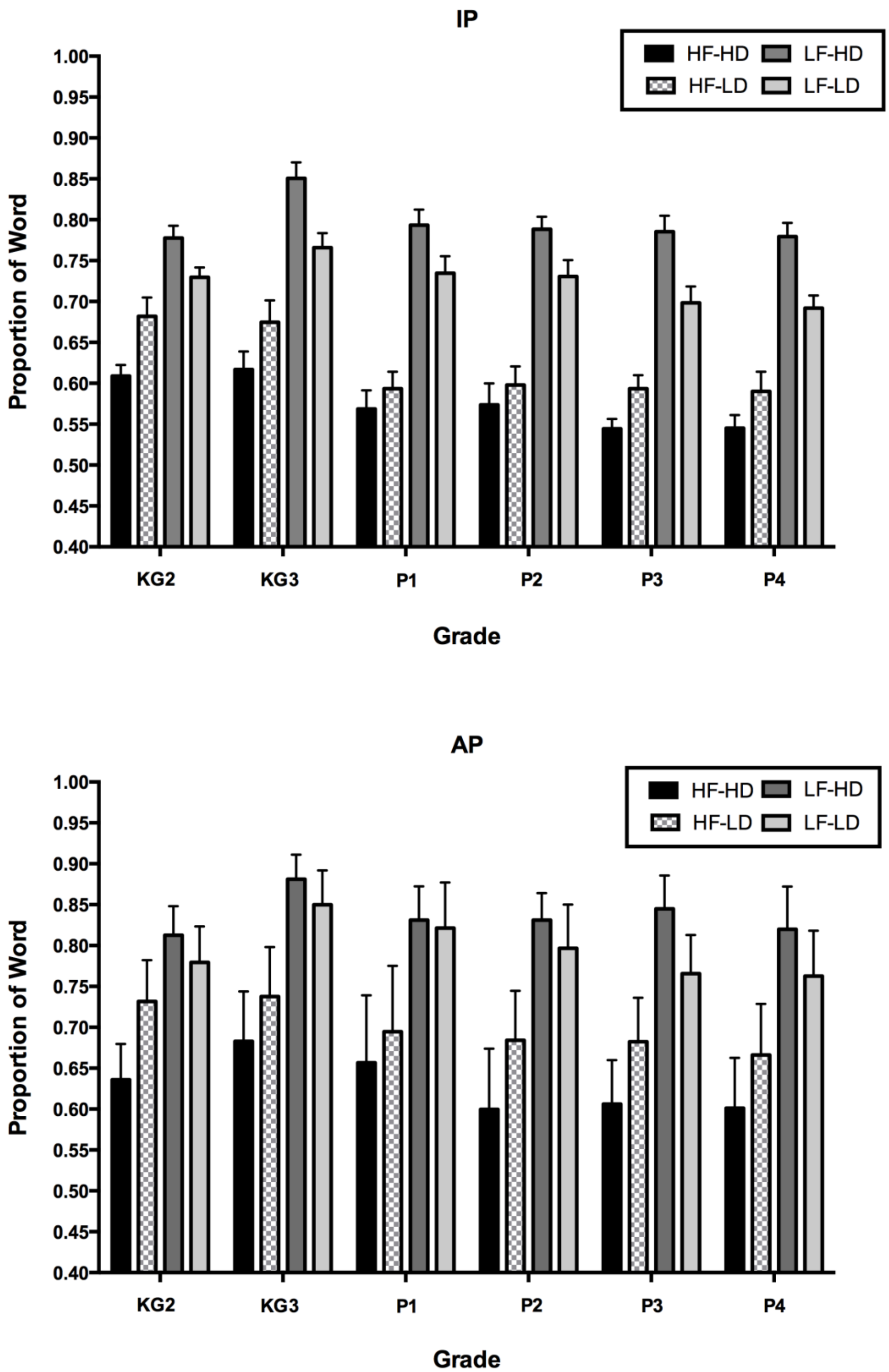

Figure 2
Number of correct rejections on the lexical judgment task for the variety of mispronunciations, across grade levels. Error bars indicate SEM.

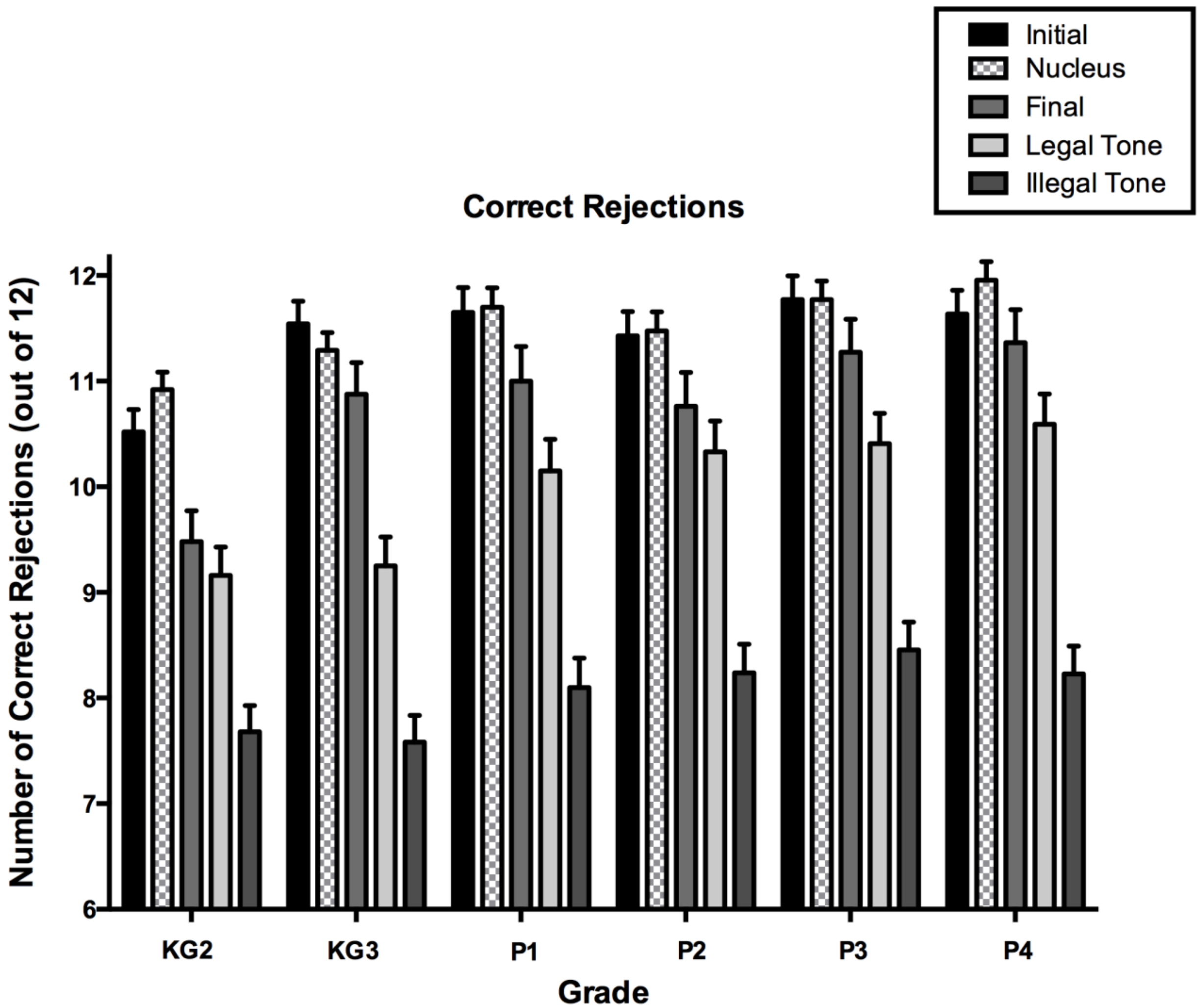

Figure 3.
Multiple mediation model on the relationship between quality of phonological representations (PR), phonological processing (PP), and reading. PR measures (SG, MS, and NWR errors) were each entered as the only independent variable in separate analysis. The top figure depicts the total effect of PR on reading (path *c*). The bottom figure illustrates both the direct effect of PR on reading (path *c'*) and the indirect effects (a_1b_1, a_2b_2, and a_3b_3) of PR on reading via the PP mediators.