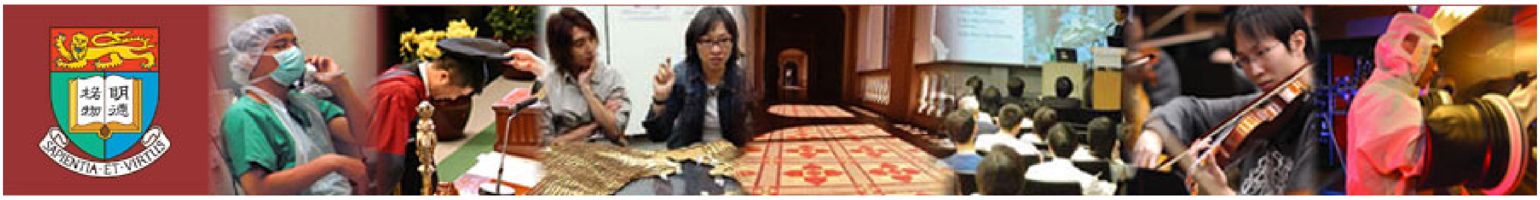


Title	The effect of word structure on the processing of Chinese two-character compound words and its acquisition in Hong Kong school-aged children
Other Contributor(s)	University of Hong Kong.
Author(s)	Lam, Man-ching, Janice; 林文晶
Citation	
Issued Date	2004
URL	http://hdl.handle.net/10722/48778
Rights	Creative Commons: Attribution 3.0 Hong Kong License

The Effect of Word Structure on the Processing of Chinese Two-character Compound Words and its Acquisition in Hong Kong School-aged Children

Lam Man Ching, Janice

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, May 7, 2004

ABSTRACT

This study investigated the hypothesis that main component characters which contributed more to the whole-word meaning was used as an access code in reading Chinese compound words. Compound words of three structure types: modifier, supplement and coordinative structures were employed to testify the hypothesis. The interaction between word structure and component character frequency in Hong Kong school-aged children and its change across grades were also investigated. Primary 1, 3 and 5 readers were recruited to participate in a reading aloud task using stimuli with three levels of word structures and four levels of component character frequency. Results showed that main component character frequency effect was present for all three types of word structures, and the use of main component character as a reading strategy was found in Primary 3 and 5 children in their reading of low-frequency compound words. It was proposed that representations of low-frequency Chinese compounds were stored in morphologically decomposed form and that word structure information had a role in the processing of compound words.

INTRODUCTION

Chinese is a logographic orthography where most characters are morphemes in their own right. Up to 89% of Chinese characters represent unique morphemes (Ku & Anderson, 2003). Besides, statistical analyses of the language showed that around 80% of Chinese words are polymorphemic (Taft, Liu & Zhu, 1999); therefore, they are regarded as multi-character words. These multi-character words are mainly formed by three main ways: affixation (i.e. the use of prefixes and suffixes to form words like 第一 [the first] and 竹子 [bamboo]), reduplication (i.e. doubling of words or syllables, like 媽媽 [mother]), and compounding (Matthews & Yip, 1994). Compounding refers to the combination of two or more independent morphemes to form a single word, for example, a two-character compound word 手錶 (hand-watch, [wrist-watch]), where the character 手 means “hand” and 錶 means “watch”. Both characters can combine with other characters to form different words. For example 手 [hand] can combine with 巾 [towel] to form 手巾 (hand-towel, [handkerchief]). 錶 [watch] can combine with 懷 [chest] to form 懷錶 (chest-watch, [pocket-watch]. This kind of combination is so productive that about 73.6% of the words in Chinese are two-character compound words (Zhou, Shu, Bi & Shi, 1999). According to Zhu (1982), Chinese compounds can be syntactically categorized into five different types with respect to the structural relationship between component characters, namely modifier compound (偏正), verb-object compound (述賓), supplement compound (述補), subject-predicate compound (主謂), and coordinative compound (聯合). According to Zhang & Peng (1992), two-character coordinative words can be viewed as words having two component characters which are of equal importance to the whole-word meaning, such as 奇怪 (strange-weird, [straight/weird]); whilst component characters of modifier compounds, such as 飛機 (fly-machine, [aeroplane]), and supplement words, such as 改良 (correct-good, [improve]), are not equally important to the whole-word meaning: the second characters of modifier words and the first characters of supplement words are the main morphemes, while the others are the modifier or supplement morphemes respectively. One the other hand, affixed words only constitute 3.63% of all Chinese words (Beijing Language College, 1986, as cited in Zhang & Peng, 1992).

The circumstance of English words is very different from Chinese one. English includes a large number of affixed words (e.g. inflected words and derived words) (Peng, Liu, & Wang, 1999). Inflectional morphology in English determines a root word's aspect, tense, number, etc; while derived morphology takes the role in changing the part of speech of roots. As a result, much research of the word representation and processing focused on derived and inflected words. In view of the fact that relatively few studies addressed the issue of how compound words were represented, Juhasz, Starr, Inhoff & Placke (2003) conducted a series of experiments on English compounds using lexical decision task, naming task, and gaze durations. The English compounds incorporated in their study were formed by joining two free lexemes (i.e. a compound's constitute words), for example "beefsteak" and "blackmail". They found an ending lexeme effect with relatively short lexical decision time, naming latencies and gaze durations when a compound's ending lexeme was a high-frequency word. They explained this result by stating that it was not the position of lexeme within the full compound form but its contribution to compound meaning that defined its effect (Starr et al., 2000, as cited in Juhasz et al., 2003); thus an ending lexeme effect would emerge in compound word when the meaning of the full compound was defined by the ending lexeme, while a beginning lexeme effect would emerge in compound word when the beginning lexeme defined compound word meaning. Since they found that ending lexemes of most stimuli generally defined the compound meanings, they suggested that the ending lexeme was used as an access code to locate the meaning of full compound word, and morphological decomposition did take place for English compound word during reading. A post-examination of Juhasz et al.'s (2003) stimuli also found that 28 out of 40 stimuli (70% in total) had the main lexemes at the ending position within the compound words. In other words, lexemes that were more important in contributing the meaning of compound words assumed a privileged role during compound recognition.

The phenomenon of using the more important morpheme to access full compound form was also found in Chinese compound word in Zhang & Peng's (1992) study. Zhang & Peng (1992) found that, when word frequency was controlled, frequency of the two component characters of coordinative words significantly affected the lexical decision latencies of stimuli. However, only the second character' frequency of modifier compound had an effect on the response time to stimuli. When character frequency was controlled, word frequency determined the latencies of two-character words. Based on the finding that both the word and character's frequency affected the access of Chinese words, they proposed that all two-character compound words were stored in decomposed form, and the information of word structure determined their relationship. They explained the difference between coordinative and modifier compounds by assuming that, for different word structures, the importance of individual component characters to the whole-word meaning affected the access of compound words. In a coordinative compound word, the two component characters were equally important in conveying the whole-word meaning (i.e. both were main characters); therefore, either the first character or the second character could work as the "head" of the character's network. Both equally important heads could positively associate with the coordinative compound. To retrieve correct semantic information, both characters would be activated. In a modifier word, the second character was the main character conveying the meaning of the word; thus only the second character could act as the head in the lexicon. The second character would be activated and the first character would be inhibited by the word structure information. The strength between these two characters depended on the frequency of the compound word, where the higher frequency two-character word had a stronger association between component characters than lower one; therefore, high-frequency word with strong association between its component characters would be activated to reach its threshold earlier. Furthermore, component characters with higher frequency would be activated stronger than those with lower frequency. The facilitatory and inhibitory effects operating both word and character frequencies with respected to different word structures explained Zhang & Peng's (1992) results of main character frequency effect and whole-word frequency effect. However, Taft, Huang, & Zhu (as cited in Taft, Liu, & Zhu, 1999) failed to find any difference between coordinative and modifier words in a post hoc analysis of Zhang & Peng's data and argued that differences observed was invalid.

Taft & Zhu (1994) (as cited in Taft, Liu & Zhu, 1999) instead suggested that morpheme and whole-word were represented at separated levels, which was against Zhang & Peng's (1992) postulation that characters were stored in the lexicon as a network. According to their multilevel interactive-activation framework (Taft, Liu & Zhu, 1999), Chinese word processing was a multilevel interactive-active system where activation units were hierarchically organized in levels corresponding to words, morphemes, and submorphemic components; information had to pass from a lower level to a higher level. Compound words were recognized by passing activation up to the meaning through the orthographic levels, including stroke, radical, character and compound word levels. The strength of connection between units would be influenced by the frequency with which that connection was used and thus, words of higher frequency were activated through stronger links at a higher speech than low-frequency words.

In order to investigate the word processing of Chinese compound words, the frequency of and the structural relationship between their component characters were manipulated in the present study. Only low-frequency words were used in the current study so as to ensure that component characters would be accessed and the effects of word structure (relationship between the two characters) would be manifested in the present experiment. Both the first and the second characters were manipulated as either high- or low-frequency, while three word structures, namely modifier, supplement and coordinative structures, were incorporated. The reason for adding the supplement structure in the current study on top of the two used by Zhang & Peng's (1992) was that the supplement compound word has its main character at the first position and whereas the modifier word has its main character at the second position. The incorporation of supplement compound would extend Zhang & Peng's (1992) findings by showing that main character effect could appear for the first character position alone without the involvement of the second character, and the second position within a compound word could be inhibited by word structure information. Besides, both Juhasz et al.'s (2003) and Zhang & Peng's (1992) studies demonstrating main lexeme/character effects used adults as subjects, and no previous study had worked on the acquisition of word structure awareness. Therefore, Primary 1, 3 and 5 children were recruited in the present study so as to investigate the developmental trend of the word structure effect from poor readers (Primary 1 children) to advance readers (Primary 5 children). Reading aloud task with error rate and pattern measurement will be adopted in the current study so that both quantitative and qualitative error analysis would be allowed.

The current study pursued two goals: first, to investigate whether main character frequency effect obtained in Zhang & Peng's (1992) study occurs during the word reading of the three structure types; second, to determine the change of the interaction of word structure and component character frequency in Hong Kong school-aged children across grades.

METHOD

Subjects

A total of 130 Cantonese-speaking children were recruited from a pool of Primary 1, 3 and 5 children from a local primary school to enter the initial stage of the study. All subjects were born in Hong Kong and studied in local kindergartens. They were administered the Raven's Standard Progressive Matrices (Raven, 1986) and the Chinese word reading sub-test of the Hong Kong Test of Specific Learning Difficulties in Reading and Writing (Ho, Chan & E.D., HKSAR Government, 2000). A total of 90 students, 30 children from each grade, who obtained scores within the 5^{th} and the 95^{th} percentile in both tests were randomly selected to enter the second stage of the present study. ANOVA was run to ensure that reading test scores and intelligence test scores of subjects from each of the grades was not significantly different. Subjects' ages and performances of the two tests were shown in Table 1.

Table 1
Ages, intelligence and reading tests scores of the three groups of subjects

Grade	Age		Scaled intelligence scores	Scaled reading scores
	Range	Mean	Mean percentile (S.D.)	Mean percentile (S.D.)
P1	6;01-7;04	6;07	67.8 (15.6)	51.5 (1.74)
P3	8;01-9;11	8;08	55.9 (9.57)	60.0 (1.91)
P5	10;01-11;06	10;08	58.1 (11.8)	50.4 (1.87)

Materials

Three sets of 120 two-character compound words were selected from the Hong Kong Corpus of Primary School Chinese (Leung & Lee, 2002) as stimuli for subjects in Primary 1, 3 and 5 respectively. The corpus contained two-character words from primary-school textbooks of the subjects of Chinese and General Studies, and stated their corresponding cumulative whole-word frequencies and component character frequencies with respect to each grade. The whole-word and component character frequencies were used to reflect children's accumulative exposure to Chinese words and characters since the entry of primary schools.

The two-character words in the corpus for each grade were divided into three frequency groups: high, mid, and low. All two-character words with the cumulative frequency of one were identified as low-frequency words, which occupied 46% of the total number of different words (NDW) in Primary 1, 43% of total NDW in Primary 3, and 40% of total NDW in Primary 5. 8.3% of the total number of different words (NDW) with the highest frequency values in Primary

1, 13.4% of total NDW in Primary 3, and 18.6% in Primary 5 of total NDW were defined as high-frequency words. The corpus of Chinese characters was also divided into three sets according to the component character frequencies, each of which contained approximately one-third of the total number of different characters (NDC). The set with the highest character frequencies was defined as high-frequency characters (33% of total NDC in Primary 1, 33% of total NDC in Primary 3, 33.5% of total NDC in Primary 5); whilst the set with lowest character frequencies was identified as low-frequency characters (37.4% of total NDC in Primary 1, 34% of total NDC in Primary 3, 34% of total NDC in Primary 5). The character frequencies of the first character and the second character of the words were manipulated to form four frequency combinations: both first and second characters were of high-frequency (HH), both first and second characters were of low-frequency (LL), high-frequency first character and low-frequency second character (HL), low-frequency first character and high-frequency second character (LH).

Apart from whole-word frequency and component character frequency, the word structures of stimuli were also controlled. Three types of word structures were incorporated into the present experiment, namely modifier, supplement, and coordinative compound words. The two-character compound words in the corpus were classified into different structure types according to the criteria adapted from Zhu (1982). First of all, for modifier words e.g. 乳齒 (milk-teeth, [milk-teeth]), (i) the first character had to modify or restrict the meaning of the second character; and (ii) the first character was either an adjective (words that could be modified by the word “很” [very] and could not carry object) or noun (represented persons, events, places, or time). All the modifier words included in the word lists were nouns. Secondly, for the supplement compound word e.g. 削弱 (reduce-weak, [weaken]), (i) the first character was a verb (words could not be modified by the word "很" [very] or could carry object); and (ii) the second character was the consequence or status of the action bought about by the first character; and (iii) the characters "得" or "不" could be added between the first and second characters. Only those supplement words that acted as verbs were included. Thirdly, for the coordinative compound words e.g. 光明 (light-bright, [bright]), (i) the first and the second characters were parallel in meaning and importance; and (ii) the part of speech of both character were the same. The coordinative words in the word lists included nouns, verbs, and adjectives.

To summarise, stimuli for each grade were divided into 12 experimental conditions according to two within-group factors: three levels of word structures (modifier, supplement, and coordinative compounds) and four levels of component character frequency (HH, HL, LH, LL). Ten compound words were selected randomly from the corpus for each condition. Examples were listed in Table 2.

Table 2
Examples of stimuli in all experimental conditions for three subject groups

	Word Structure											
	Modifier words				Supplement words				Coordinative words			
	Character frequency				Character frequency				Character frequency			
Grade	HH	HL	LH	LL	HH	HL	LH	LL	HH	HL	LH	LL
P1	歌聲	魚缸	謊話	恆齒	照亮	放鬆	漂白	剪斷	孩童	站立	閱讀	危險
P3	表皮	石階	墨鏡	硬幣	加深	喝醉	擴大	淋濕	菜蔬	土壤	軌道	刑罰
P5	香煙	果醬	讒言	迴廊	拉斷	吹熄	削弱	嚼碎	病痛	反叛	惦掛	牢獄

All stimuli were randomized. Each stimulus was presented in black with size 90 of "biau kai" font (標楷體) at the centre of one-fourth of an A4 size plain white paper. They were bound into a loose-leaf booklet for each grade.

Procedure

Subjects were tested individually. Each subject was seated in a quiet room in their primary school facing an experimenter. Participants were firstly informed about the audio-recording of their performances and were told to read out 120 stimuli of their grades, which were presented one at a time in form of a loose-leaf booklet, without time limitation. Subjects' productions were recorded manually on scoring sheets and instrumentally with MD discs. All experimenters of the reading aloud task in the present study were undergraduate students majoring in speech and hearing sciences of the University of Hong Kong.

Measurement

Subjects' performances were manually recorded character by character on scoring sheets on a mean-time basis. Target compound words were regarded as correct when subjects produced both characters of those words correctly and the numbers of correctly produced word were scored. All incorrect productions were transcribed orthographically. If there was no suitable orthographic substitute, IPA transcription was used. Audio-recording was used for double-checking the transcription of error production. Error pattern was used in the qualitative analysis.

Design and data analysis

There were three independent factors: grade (three levels), word structure (three levels), and component character frequency (four levels). Grade was the between group factor, while word structure and component character frequency were within-group factors. The dependent factor was the subjects' mean score of word reading performance. The mean score of percentage correct in each experimental condition was calculated, and then entered into a 3 (grade) x 3 (word structure) x 4 (character frequency) three-way ANOVA with repeated measures.

RESULTS

Effects of grade, word structure and component character frequency on the naming accuracy were shown in Table 3.

Table 3

Mean and standard deviation of the subject's percentage correct

	Overall percentage score (%)			Mean percentage score (%) Component character frequency			
Grade	Mean (S.D.)			HH	HL	LH	LL
P1	50.4 (2.93)	Word structure	Modifier	89.0	51.0	45.7	54.7
			Supplement	88.0	28.7	23.7	11.7
			Coordinative	80.3	62.0	39.3	30.7
P3	67.0 (2.51)	Word structure	Modifier	96.7	62.8	80.0	53.3
			Supplement	96.3	60.0	49.0	45.0
			Coordinative	95.3	64.7	61.0	40.3
P5	74.1 (2.40)	Word structure	Modifier	97.7	86.0	60.0	89.0
			Supplement	99.7	81.7	56.3	46.3
			Coordinative	96.3	64.0	67.0	44.7

A 3-way ANOVA with repeated measures was administered. The main effect of grade was statistically significant, $F(2, 87) = 36.09$, $p<.001$. The differences among grades that contributed to the main effect were investigated with a Tukey HSD post hoc test and all pair-wise comparisons were significant, $p<.050$. This suggested that reading accuracy significantly increased across grades (as seen in Table 3). Since the whole-word frequency and component character frequency were controlled across grades, the result could not be attributed to the more frequent occurrence of the characters and compound words by higher grade then lower grade readers, and a more likely explanation was that higher grade readers acquired better reading skills.

The main effect of word structure was statistically significant, F(2, 174) = 182.49, p<.001. All pair-wise comparisons were significant in the Tukey HDS post hoc test, p<.001. Modifier words (M) were more accurately read than supplement (S) and coordinative words (C), while the performance of coordinative words were better than that of supplement words (i.e. M>C>S).

Significant main effect of component character frequency was found, F(3, 261) = 604.85, p<.001. All pair-wise comparisons were significant in the Tukey HDS post hoc test, p<.001. Subjects' performance decreased across the component character frequency condition HH>HL>LH>LL by investigating the means table. This finding agreed with the results reported by Taft, Huang & Zhu (1994) (cited in Peng, Liu & Wang, 1999).

Significant interaction effect between word structure and component character frequency, F(6, 522) = .000, p<.001, was found. The interaction was further investigated with the Tukey test and it shown that subjects' performance decreased significantly across the component character frequency levels in the order of HH>HL>LH>LL when the compound words are supplement and coordinative ones, p<.001. For modifier word, HH component character frequency performed significantly better than HL, LH and LL levels, p>.001; while differences among other component character frequency levels were insignificant. These results was elusive and appeared not to support Zhang & Peng's (1992) hypothesis that frequency effect would only observed in the main characters of compound words. However, main character effect did occur when results of the 3-way interaction were considered.

Three-way interaction among all factors was statistically significant, F(12, 522) = 24.93; p<.001. The Tukey analysis was done to investigate the factors contributing to the significant interaction. Figure 1 displayed the performance of three subject groups as the functions of word structure and component character frequency in the reading aloud task.

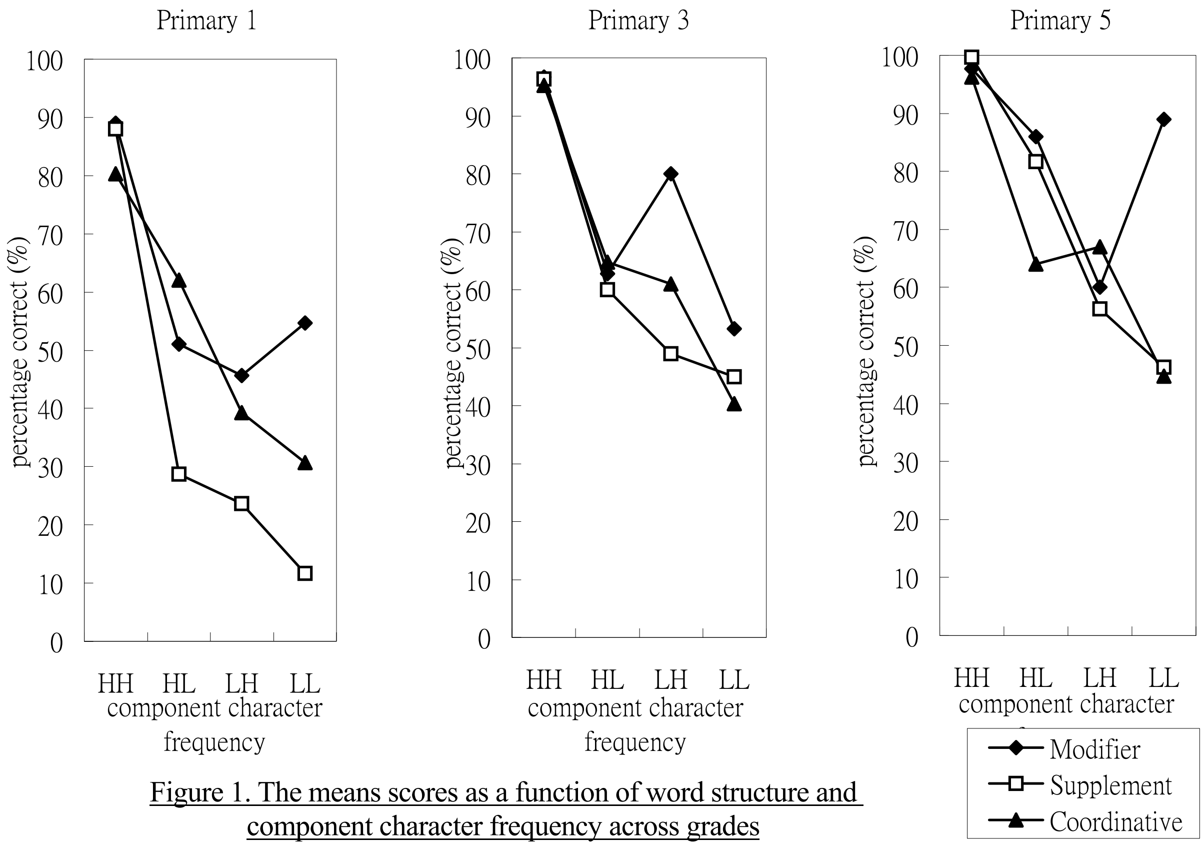


Figure 1. The means scores as a function of word structure and component character frequency across grades

The performance differences in reading all word structures with HH component character frequency were not significant in all grades. Besides, there was no significant difference of the three word structures with HH conditions between Primary 3 and 5.

In the modifier word condition, the pair-wise comparisons in Tukey test revealed that Primary 1 children read modifier words with HH component character frequency significantly better than HL, LH and LL, p<.001; while the difference among HL, LH and LL conditions were insignificant (i.e. HH>HL≈LH≈LL). Primary 3 readers performed similarly poor in modifier words with HL and LL conditions than in HH and LH conditions (p<.001), with accuracy in HH conditions was significantly higher than LH condition (p<.001) (i.e. HH>LH>HL≈LL). This suggested that character frequency effect was observed at the second position in a modifier word but first character frequency effect was only obtained when the second character was of high frequency. For Primary 5 readers, performance in reading modifier words significantly decreased from component character frequency conditions of HH to HL and to LL ($p<.020$). However, it was interesting to noted they performance in LL conditions was statistically similar to that of HH and HL conditions ($p>.1$) (i.e. HH>HL>LH, while LL≈HH and LL≈HL). The finding showed that main character effect for modifier words was present in Primary 3 children.

In the supplement word condition, Primary 1 children demonstrated significant differences between HH and HL ($p<.001$), as well as LH and LL conditions ($p<.050$), with HH was better than HL and LH was better than LL; whilst no significant difference between HL and LH component character frequency was found (i.e. HH>HL≈LH>LL). For Grade 3 and 5, both groups of readers performed similarly in LH and LL conditions, but with the accuracy significantly poorer than those in HH and HL conditions ($p<.001$) (i.e. HH>HL>LH≈LL). This showed that character frequency effect was observed at the first position in a modifier word but second character frequency effect was only obtained when the first character was of high frequency. The finding demonstrated that main character effect for supplement words was present in Primary 3 and 5 children.

In the coordinative word condition, Primary 1 readers performed equally poor in coordinative words with LH and LL conditions than in HH and HL conditions ($p<.001$), with the accuracy in HH conditions significantly higher than HL condition ($p>.001$) (i.e.HH>HL>LH≈LL). For both Primary 3 and 5 children, they demonstrated significant differences between HH and HL

($p<.001$), as well as LH and LL conditions ($p<.001$), with HH better than HL and LH better than LL; whilst no significant difference between HL and LH component character frequency was found (i.e. HH>HL≈LH>LL). This revealed that component character frequency effects were observed at both position of characters in the coordinative words. The finding demonstrated that main character effect for coordinative words was present in Primary 3 and 5 children.

In summary, significant main effects of grade, word structure and component character frequency were found. Besides, significant two-way interaction effects between word structure and component character frequency, and three-way interaction of grade, word structure and component character frequency were also found.

DISCUSSION

The results of the present study showed that low-frequency Chinese compound words were morphologically decomposed in the lexicon during reading, since the effect of component character frequency was significant with the pattern of HH>HL>LH>LL. This finding did not agree with that obtained by Peng, Liu, & Wang (1999). They found that component character frequency effect only existed in high-frequency words but not in low-frequency words, so they postulated that high-frequency words tended to be represented as decomposed components, whereas low-frequency words tended to be represented in unitary fashion. Therefore, Peng et al. (1999) proposed another model named Inter/Intra Connection (IIC) model. In this model, both whole-words and morphemes were represented at the same level and there were intra-level connections between morphemes and whole-word units. The strength of connections was determined by the whole-word frequency and the semantic relationship between the meaning of morphemes and words. The information from the orthographic level mapped directly onto access representation through inter-level connections where both words and morphemes might be activated. When a word was visually presented, both morphemes units and word units in the access representation would receive activation from a lower visual input level. Word units would receive activation from morpheme units at the same level simultaneously. Since the connections between high-frequency words were stronger than those between low-frequency words, the impact from morpheme units of high-frequency words would be larger than that of low-frequency words. This model help explaining their results that component character frequency effect existed only in high-frequency words but not in low-frequency ones. However, the present experiment found a component character frequency effect for low-frequency words, so it was hypothesized that low-frequency words tended to be represented in a decomposed way, contradicting to Peng et al.'s (1999) conclusion that low-frequency words were represented as a whole in lexicon.

The main character frequency effect in modifier words (i.e. HH>LH>HL≈LL) was found in Primary 3 children while main character effects in both supplement words (i.e.HH>HL>LH≈LL) and coordinative words (i.e. HH>HL≈LH>LL) were obtained in Primary 3 and 5 children. Concerning the performance of Primary 3 children in reading modifier words, significant frequency effect was always found at the second character no matter the first character frequency was high or low, since HH>HL and LH>LL; however no frequency effect at the first character was obtained when the second character were of low-frequency condition, since HL≈LL. First character frequency effect, HH>LH, was found only when the frequency of the second character was held constant at a high level. This suggested that the second characters of modifier words played a more important role in compound word processing. This phenomenon was attributable to the morphological structure of modifier words. In modifier words, the two characters were not of equal importance. The second character was the main character while the first character was just a modifier of the second character, which, provided less clues to whole-word meaning. Since the result was consistent with the study of Zhang & Peng (1992) which adopted a lexical decision task, it was assumed that, although reading aloud task was used in the current study, the morphological structure information would facilitate the second character activation and inhibit the first character activation. Besides, the main character frequency effect emerged in coordinative words in Primary 3 and Primary 5 readers in the way that effect present in both characters, as seen from the result that HH>HL≈LH>LL, and it was also in agreement with Zhang & Peng's (1992) study. It can be explained by the fact that the two characters of the coordinative words were of equal importance and either the first or the second character could be used to access the full compound form. Thus, both characters expressed their role in reading of coordinative compound words. Moreover, the reading score of supplement words, a type of compound words where the first characters were the main characters, in Primary 3 and 5 children also revealed the main character effect. Significant frequency effect was obtained for the first characters of supplement words (i.e. HH>LH & HL>LL but LH≈LL). This substantiated Zhang & Peng's (1992) study which only involved modifier and coordinative compounds and supported the argument that main character assumed a privileged role during compound reading.

As seen from the fact that frequency of a character only had an effect when represented a morpheme that was semantically important to the meaning of the word (i.e. the second character of modifier words, the first character of supplement words, and both characters of coordinative words), it would imply that relationship between component characters had a major influence on the storage and processing of compound words. In other words, access representation was a mixed one (i.e. whole-words and morphemes were represented at the same level) and both facilitatory and inhibitory links between the two component characters concerning the activation of individual characters could be inferred. The information of word structure determined whether the link was facilitatory or inhibitory; the sum of facilitatory and inhibitory activations determined which characters could express its role. This conclusion did not agree with Zhu & Taft's (1994, as cited in Taft, Liu & Zhu, 1999) model which considered words and characters were at two different levels of representation with some connections between levels.

The results in the present study demonstrated a developmental effect of word structure in reading of compound words across grades. Children in Primary 3 demonstrated main character frequency effects in all the investigated word structures, Primary 5 children showed effects in supplement and coordinative words. Although the elusive behaviour of modifier words in Primary 5 readers did not show main character effect as expected previously, the examination of the materials used in this experiment would provide methodological explanation. It was found that for the modifier words in HL condition, there were four out of tem stimuli where the pronunciation of the phonetic radicals of the low-frequency character were regular to the pronunciation of corresponding characters. As the first character was a high-frequency character and the second low-frequency character contained a radical that carrying the phonological information of the whole character, subjects had a higher chance to read modifier word with HL conditions correctly and thus an unexpectedly high percentage accuracy in the modifier words with HL condition in Primary 5 children resulted. This situation also occurred in Primary 5 modifier words with LL condition, where three out of ten stimuli were containing radicals that were phonologically regular to the whole character, hence giving rise to another unexpectedly high reading accuracy. The lack of control in the phonological regularity in modifier words with HL and LL conditions in Primary 5 wordlist caused the disobedience to the expected pattern of main character effect. For the stimuli of other word structures, this uneven distribution did not occur more than once in any condition and thus the effect of regularity on the results of other structures were insignificant. Therefore, it was suspected that children in Primary 5 also read low-frequency modifier words with the use of main character, despite the methodological bias introduced by the phonologically regular stimuli. This claim was further supported by the error pattern analysis of the Primary 5 data (discussed later). It was concluded that Primary 3 and 5 children had developed morphological awareness of word structure and would use main character as a strategy of compound word reading.

Pattern of main character effect was not observed in Primary 1 subjects. It was concluded that the use of main character as a reading strategy in compound word reading had not yet been developed in Primary 1 readers; whilst advanced readers like Primary 3 and Primary 5 children had a higher morphological awareness than Primary 1 graders. This development of morphological awareness of word structure went along with the significant increase in the reading performance across grades as seen in the main effect, which reflected that the reading ability of children also increased with grades. Since both whole-word frequency and component character frequency were controlled for each grade in the current study, the significant increase in reading performance across grade was probably due to the more advance use of reading strategy but not their accumulative exposure to the targeted compound words.

Though different word structures did not affect the reading performance of Primary 1 children in the form of main character frequency effect, significant interaction effect between grades and word structure was still observed. Primary 1 readers performed significantly better in modifier words, followed by coordinative words and supplement words the worst. The significant differences among three word structures were not contributed by the main character effect since the pattern of word structures across different component character frequency conditions deviated from the expected one. In other words, the source of significant difference came from factors brought about by word structures other than main character. In order to find out the factor, an investigation into the proportion of compound words with different structures was done. Token frequency of word structures of all two-character compound words with high and low whole-word frequency were calculated. Results are showed in Table 4.

Table 4
Token frequency of word structures in the corpus of two-character compound words

	Token frequency (%)		
	Primary 1	Primary 3	Primary 5
Modifier words	33.7	29.2	27.8
Supplement words	3.40	2.94	3.03
Coordinative words	19.1	19.0	23.3
Others	43.8	48.9	45.9

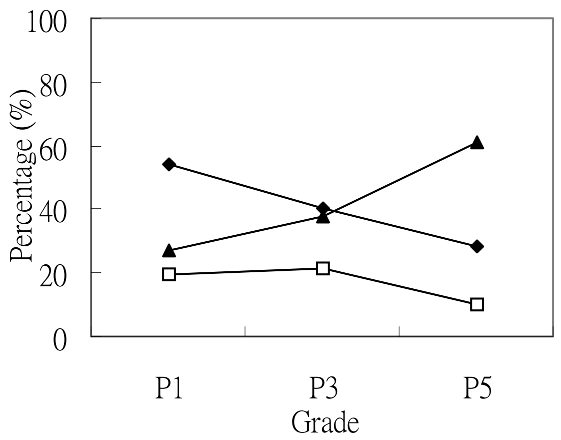
According to the result as showed in Table 4, Primary 1 children exposed more to modifier words (one-third of all the high- and low-frequency two-character words) when compared to the other two word structures under investigation. Coordinative words ranked second and supplement words last. This order of word structure distribution (i.e. M>C>S) in the corpus of Primary 1 children was consistent with the corresponding accuracy in the reading aloud task. That is, when a Primary 1 reader, whose reading strategies were not as effective as advance reader, encountered a low-frequency or unfamiliar words, he/she would tend to read the compound word as a modifier word; if mapping of that word into a modifier structure failed, he would then attempt to map the word into a coordinative structure and lastly a supplement. As a result, modifier words were more readily to be read correctly than coordinative and supplement words. This implied that the interaction effect between Primary 1 condition and word structures might came from the distribution of word structures in children's lexicon, where compound words with structure of higher proportion tended to be read more accurately than lower ones.

Apart from the quantitative analysis as stated above, the analysis of subjects' error provided further evidence to show the effect of word structures and component character frequency on the reading of two-character compound words in the primary school children of Hong Kong across grades. Errors produced by all subjects were firstly grouped together and classified into substitution errors and no response. Substitution errors were further sub-categorized into (i) whole-word level errors, which referred to the use of a semantically related word that was activated along with the target at the word level processing to replace the target; (ii) character level errors, which were errors resulted from the competition of the target with the related entries, such as phonologically, orthographically and semantically similar units, at the component character level, and (iii) errors involving both word and character levels, which arose when readers processed the compound words into a decompositional fashion and retrieved erroneous characters/whole-word at the whole-word level. Table 5 displayed the different types and number of errors produced by three groups of subjects.

Table 5
Different types and number of errors produced by three groups of subjects

		Primary 1		Primary 3		Primary 5	
Type of Errors		**No.**	**% in total**	**no.**	**% in total**	**no.**	**% in total**
Substitution errors		799	44.7 %	627	52.8 %	580	62.1 %
No response		987	55.3 %	560	47.2 %	354	37.9 %
Total no. of errors		1786		1187		934	
Type of Substitution Errors		**No.**	**% in sub error**	**no.**	**% in sub error**	**no.**	**% in sub error**
Whole-word level access errors		4	0.50 %	19	3.00 %	9	1.60 %
Character level access errors	Phonological	83	10.4 %	114	18.2 %	108	18.6 %
	Orthographical	206	25.8 %	157	25.0 %	176	30.3 %
	Semantic	16	2.00 %	11	1.80 %	11	1.90 %
Errors involved both levels	Transposition	56	7.00 %	80	12.8 %	21	3.60 %
	Induced production	105	13.1 %	96	15.3 %	74	12.8 %
Others		330	41.3 %	187	29.8 %	171	29.5 %
Type of Induced production		**No.**	**% in ind error**	**no.**	**% in ind error**	**no.**	**% in ind error**
Known character at the first position	Modifier substitution	28	53.8 %	15	40.5 %	14	28.6 %
	Supplement substitution	10	19.2 %	8	21.6 %	5	10.2 %
	Coordinative substitution	14	26.9 %	14	37.8 %	30	61.2 %
Known character at the second position	Modifier substitution	33	62.3 %	18	30.5 %	6	24.0 %
	Supplement substitution	6	11.3 %	12	20.3 %	4	16.0 %
	Coordinative substitution	14	26.4 %	29	49.2 %	15	60.0 %

The error pattern analysis focused on discussing errors involving both character and whole-word levels because it was the type of errors which illustrated the morphological relationship between character within a whole-word. There were two subtypes of errors involving both levels, namely transposition errors and induced errors. Transposition errors occurred when a reader analysed a compound word, for example 睡房 (sleep-room, [bedroom]), character by character and tried to access the representation of 睡 [sleep] and 房 [room] individually at the component character level; however, since the character 房 [room] was a low-frequency character and thus the child failed to find such representation at the character level. As a result, a high-frequency whole-word representation containing that character, such as 房屋 (room-house, [housing]), was instead activated at the word level. Once the activation at the whole-word level fed back to the component character level, the targeted character in the activated word, i.e. 房 [room], was transposed to another position and thus the character 屋 [house] was actually retrieved at the component character level. A non-word error of 睡屋 (sleep-house) produced. However, the morphological relationship between characters of the produced error was not clearly specified. Hence, another kind of errors which also involved both character and word levels processing were thus utilised and named as induced errors. For this type of error to happen, reader also decomposed the target compound word, for example 山洞 (mountain-hole, [cave]), into its component characters 山 [mountain] and 洞 [hole] during the reading process. However, since the character 洞 [hole] was of low-frequency, no representation could be found at the character level. Reader could not activate whole-word representation containing the character 洞 [hole] either at the word level processing. As a result, the reader could only depend on the high-frequency character 山 [mountain] to figure out the whole-word target. The known character fed its activation to the units representing whole-word that contained this character, such as 山村 (mountain-estate, [village]), and thus 山村 (mountain-estate, [village]) was produced as a response to stimulus 山洞 (mountain-hole, [cave]). This subtype of errors revealed readers' preference of morphological structures when only one character at a particular position within the stimuli words was accessed at the component character level. Thereby, if readers used main characters of corresponding word structures during the reading of two-character compound words, it was expected that, when the unfamiliar character was at the first position and the second was a known character, the second character became the main character and this morphological information would have favoured a modifier or coordinative substitute. On the other way round, when the known character was at the first position and the second was an unfamiliar character, the first character became the main character and this morphological information would have favoured a supplement or coordinative substitutes. Table 5 also showed subjects' preferences for particular word structures in induced errors when known characters were either at first or second positions across grades.



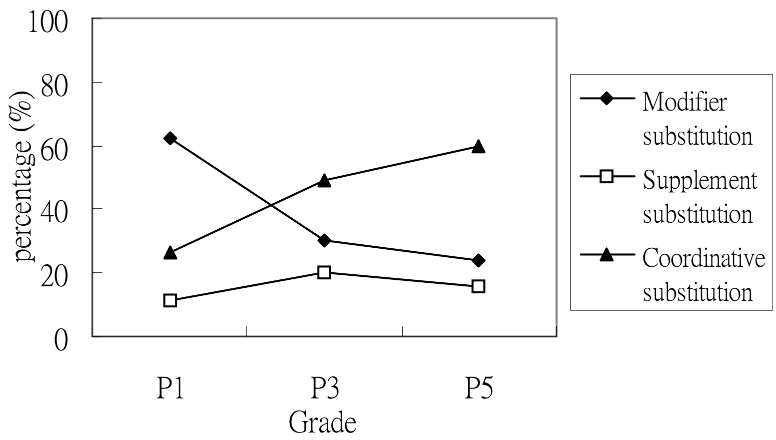


Figure 2b. Known character at the second character position

As seen from Figure 2a and Figure 2b, no matter which positions the known characters located, preference for supplement words was the least among all three structures in all grades. This might be due to limited distribution of supplement words in all grades. When the main character was at the first position of a stimulus word, there was an increase in preference for a coordinative structure and a decrease in preference for modifier structure from Primary 3 to Primary 5; whilst preference for coordinative words also significantly increase from Primary 3 to 5 when the main character was at the second position of the target compound word. These patterns of preference for word structures were supportive to the proposition that, for those who were able to use main character in compound word reading, advanced readers (i.e. Primary 5 children) used this reading strategy more readily than less advanced readers (i.e. Primary 3 children). For example, when a Primary 5 reader encountered a new word and was required to read it out, he/she would map a familiar compound word containing that known character at the corresponding position because compound words contributed a large proportion in his/her lexicon. If the known character was at the first position, it became a main character, and the reader would probably went for a coordinative compound word containing that known character at the correct position, because both character positions of coordinative compounds could accommodate main characters. Supplement compounds, where the first character position could also accommodate a main character, would not be chosen since its proportion in the reader's lexicon was significantly small. Therefore, if a coordinative compound could not be mapped, a modifier compound would be activated instead of supplement compound. On the other hand, if the known character was at the second position, it became a main character and both modifier and coordinative words would be mapped, since both types of words could accommodate main character at the second position. Coordinative compound would be considered earlier than modifier compound because the morphological structure information of the same grammatical class of the characters in coordinative words gave this type of compound a privilege.

For subjects in Primary 1, whichever positions the main characters located, the preference for word structure was always favourable to modifier, following by coordinative and supplement words in a descending order. This was another piece of evidence to the claim that distribution of word structures had an effect on reading two-character compound words in Primary 1 children, who had not yet developed the main character reading strategy. Since Primary 3 and 5 readers were developing their morphological awareness of main characters in different structures, the effect of word structure distribution was not as robust as that in Primary 1, and it interacted with the main character frequency effect. However, the interaction between the distribution effect and main character frequency effect must be considered tentative.

Taken together, the findings of the present study proposed that representations of low-frequency Chinese compounds were stored in morphologically decomposed form; and characters were stored in the lexicon as a character network. Word structure information determined the relationship between the component characters and influenced word processing in the way that compound words were accessed via their main characters. Besides, it was suggested that the use of main characters in corresponding structures for lexical access started to develop in Primary 3 and 5 children as a reading strategy of low-frequency compound words. Reading performances of two-character compound words of Primary 1 children were affected by the proportion of different word structures they exposed. Modifier structure being the most favourable one, followed by coordinative and supplement structures.

ACKNOWLEDGEMENT

I would like to give my most sincere thanks to my supervisor, Dr. Leung Man Tak, for his valuable guidance, and to my dear classmates who helped me with my data collection in this study. I would also like to deliver my grateful appreciation to Arki Wong, Fanny Lam, Florence Sung, Grace Leung, and Shum Wing Man for providing emotional support throughout the year.

REFERENCES

Ho, C. S-H., Chan D. W-O. & Education Department, HKSAR Government. (2002). *The Hong Kong Test of Specific Learning Difficulties in Reading and Writing*. H.K.: E.D., HKSAR Government.

Juhasz, B.J., Starr, M.S., Inhoff, A.W. & Placke, L. (2003). The effect of morphology on the processing of compound words: evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, 94, 223-244.

Ku, Y.-M., & Anderson, R.C. (2003). Development of morphological awareness in Chinese and English. *Reading and Writing: An interdisciplinary Journal*, 16, 399-422.

Leung, M.T. & Lee W.Y. (2002). *Hong Kong Corpus of Primary School Chinese*. Paper presented at the 9th Meeting of the International Clinical Phonetics and Linguistics Association, Hong Kong.

Matthews, S. & Yip, V. (1994). *Cantonese: A comprehensive Grammar*. London: Routledge.

Peng, D.L. & Ding G.C. (1997). Zhong wen shuang zi ci de biao zheng yu Jia gong. *Xin li ke xue*, 20, 4, 294-297

Peng, D., Liu, Y., & Wang, C. (1999). How is access representation organized? The relation of polymorphemic words and their morphemes in Chinese. In J. Wang, A.W. Inhoff & H.-C. Chen (Eds)., *Reading Chinese Script: A Cognitive Analysis*. New Jersey: Lawrence Erlbaum Associates.

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

Taft, M., Liu, Y. & Zhu, X. (1999). Morphemic processing in reading Chinese. In J. Wang, A.W. Inhoff & H.-C. Chen (Eds)., *Reading Chinese Script: A Cognitive Analysis*. New Jersey: Lawrence Erlbaum Associates.

Zhang, B.Y. & Peng, D.L. (1992). Decomposed storage of Chinese words. In H.-C. Chen & O.J.L. Tzeng (Eds), *Language Processing in Chinese*. Amsterdam: Elsevier Science Publishers.

Zhou, X., Shuh, H., Bi, Y. & Shi, D. (1999). Is There Phonologically Mediated Access to Lexical Semantics in Reading Chinese? In J. Wang, A.W. Inhoff & H.-C. Chen (Eds)., *Reading Chinese Script: A Cognitive Analysis*. New Jersey: Lawrence Erlbaum Associates.

Zhu, D.X. (1982). *Yu fa jiang yi*. Beijing : Shang wu yin shu guan.