



306416751X

Word Recognition and Reading in Chinese

Yi-Ping Chen

Green College, University of Oxford



(A thesis submitted for the degree of Doctor of Philosophy)
Hilary term, 1993

To My Daughter,

Lucia Y.R. Chen

Contents

Long Abstract

Chapter 1. The Chinese Language	1
Chinese speech.....	1
1) Dialects and the common language (Mandarin).....	2
2) Homophony.....	4
3) Tones.....	5
Written Chinese and its reforms.....	6
1) From literary style to colloquial style.....	6
2) From sophisticated characters to simplified characters.....	8
3) Phonetic representation of Chinese characters (Pin-Yin).....	9
4) The Chinese lexicon-single character words and multiple characters CI.....	10
5) The sorting systems of the Chinese lexicon.....	13
Chapter 2. Chinese Orthography	15
Orthographic knowledge and word recognition.....	16
1) Orthographic units of written words.....	16
2) Rules relevant to reading.....	17
Orthographic knowledge of Chinese words.....	17
1) The orthographic units of Chinese words: <i>the smallest stroke-patterns</i>	17
2) Two functional distinct orthographic components of ZI: <i>the lexical radicals and the non-radical components</i>	21
3) Three types of rules in Chinese.....	24
<i>a) The traditional (descriptive) rules of word-formation</i>	
<i>b) Rules of lexical decomposition: Part-to-Whole Correspondences</i>	
<i>c) Transcoding rules: Derivation and Analogy</i>	
Comparison between Chinese and English orthography.....	32
Chapter 3. Theoretical Issues and Methodological Concerns	34
Theoretical issues.....	34
Methodological concerns.....	35
1) Background of experiments.....	35
<i>a) Simultaneous Same-Different comparison paradigm</i>	
<i>b) Subjects who participated in this study</i>	
<i>c) Stimuli and language statistics employed in this study</i>	

2) Statistical analyses.....	37
<i>a) Variables used for assessments</i>	
<i>b) Speed-accuracy trade-off</i>	
<i>c) Overall analyses</i>	
<i>d) Further analyses</i>	
Chapter 4. The Visual Analysis of Chinese Script.....	40
Paradigm and Rationale.....	40
<i>Experiment 1: Visual comparison of pairs of Chinese words.....</i>	<i>42</i>
Introduction.....	42
Method.....	43
Results.....	45
1) A unit-dissimilarity effect	
Discussion.....	48
<i>Experiment 2: Visual comparisons of Chinese words, pseudowords and nonwords.....</i>	<i>48</i>
Introduction.....	48
Method.....	52
Results.....	58
1) Word type effect on both judgments	
2) Effects of unit-complexity and stroke-complexity	
3) Effect of unit-dissimilarity	
Discussion.....	68
<i>Experiment 3: Visual comparisons of Chinese script by non-Chinese readers.....</i>	<i>69</i>
Introduction.....	69
Method.....	71
Results.....	71
1) Visual familiarity and judgments effects	
2) Same judgments: unit-complexity and stroke-complexity effects	
3) Different judgments: unit-dissimilarity	
4) Comparison of performance between Chinese and non-Chinese readers	
<i>a) Speed of performance</i>	
<i>b) Fast 'Different' performance</i>	
<i>c) Same judgments: Unit-complexity and stroke-complexity</i>	
<i>d) Different judgments: Unit-dissimilarity</i>	
Discussion.....	81
General discussion.....	82
1) The functional orthographic unit of Chinese words: <i>The stroke-pattern but</i>	

not the individual stroke.....82

2) The pseudowords advantage in Chinese: *The importance of the lexical radical and its positional regularity*.....84

3) Word superiority over pseudowords in Chinese: *"Words" rather than "characters"*84

4) Two types of word superiority effect in Chinese: *Mutual interactions between units and words*.....85

Conclusions.....86

Chapter 5. Phonological and Semantic Processing.....88

Introduction.....88

Background & Rationale.....89

Experiment 4: A pilot study of the phonological and the semantic comparisons of Chinese words.....94

Introduction.....94

Method.....94

Results.....99

Reaction time analyses

Error rate analyses

 1) *Opposite effects of radical and non-radical matching on 'Different' judgments of the semantic and the phonological comparisons*

Discussion.....105

General methods for Experiments 5-10.....106

 1) Subjects.....106

 2) Apparatus.....106

 3) Paradigms and procedures.....106

Chapter 6. From Script to Sound in Chinese.....109

Theoretical issues.....109

Experiment 5: Phonological comparisons of Chinese words.....109

 Introduction.....109

 Stimuli and conditions.....110

 Result112

 1) *Synonym effect on Different judgments*

 Discussion.....114

Experiment 6: Phonological comparison of Chinese words with matching component.....115

Introduction.....	115
1) Rule-defined phonetic component.....	115
2) Three types of frequency indices of a component in the lexicon.....	116
<i>b) Radical frequency index (RFI)</i>	
<i>c) Congruency index</i>	
<i>d) Consistency index</i>	
Stimuli and conditions.....	118
Results.....	121
1) Same and Different judgments	
2) Different judgments: effects of Radical and Non-Radical Match	
3) Same judgements: effects of different non-radical component match	
4) The indices of Radical frequency, Congruency and Consistency of a component: effect on phonological processing	
Discussion.....	132
<i>Experiment 7: Phonological comparisons of word-pseudoword pairs.....</i>	134
Introduction.....	134
Stimuli and conditions.....	135
Results.....	137
1) Effects of the radical and the non-radical matching with word-pseudoword pairs	
Discussion.....	139
General discussion.....	140
1) A direct route from script to sound.....	140
2) The rule-governed lexical decomposition process.....	141
Conclusions.....	143
Chapter 7. From Script to Meaning in Chinese.....	145
Theoretical issues.....	145
<i>Experiment 8: The semantic comparison of pairs of Chinese words.....</i>	<i>145</i>
Introduction.....	145
Stimulus Conditions.....	146
Results.....	148
1) Homophone effect on Different judgments	
Discussion.....	150
<i>Experiment 9: Semantic comparisons of Chinese words with matching component.....</i>	<i>151</i>
Introduction.....	150
Stimuli and conditions.....	153

Results.....	154
1) 'Same' and 'Different' judgments	
2) Different judgments	
3) Same judgments	
4) Influence of radical frequency index upon the semantic comparisons	
Discussion.....	161
<i>Experiment 10: The semantic comparison of word-pseudoword pairs</i>	164
Introduction.....	164
Results.....	164
1) Word-Radical and Word-Component match in word-pseudoword pairs	
2) Graphic-Radical and Graphic-Component match in word-pseudoword pairs	
Discussion.....	168
General discussion and Conclusions.....	168
1) A direct route from script to meaning	
2) The functional role of the lexical radical	
Chapter 8. Reading two-character CI in Chinese	171
Introduction.....	171
<i>Experiment 11: Lexical decisions with two-character CI in Chinese</i>	171
Introduction.....	171
Background & Rationale.....	171
1) Paradigm: the lexical decision of two-character CI in Chinese.....	171
2) Reading strategies.....	177
Method.....	177
Results.....	181
1) CI-type effect.....	181
2) CI frequency and concreteness effects.....	183
3) Pseudo-CI: analysis of main stimulus-categories.....	186
4) CI-homophones: effect of tone and base-CI frequency.....	189
5) Component migration effects (Lexical decomposition).....	189
6) The speed and accuracy of lexical decision and its association with reading strategy indices.....	193
7) Reading strategies: Interassociations among orthographic, semantic and phonological processing indices.....	197
Discussion.....	197
1) Lexical decision with two-character CI.....	197

a) <i>Orthographic-lexical processing: effects of CI-frequency.</i>	
b) <i>Phonological recoding</i>	
c) <i>Lexical decomposition processes</i>	
2) Relationship between different reading strategies and their implications for reading efficiency.....	205
Conclusions.....	207
Chapter 9. Summary and Implications.....	209
Summary of the empirical results.....	209
1) The visual analysis of Chinese scripts.....	209
2) Semantic and phonological processing of Chinese words.....	210
a) <i>The direct routes from script to sound and from script to meaning</i>	
b) <i>Rule-governed lexical decomposition processes</i>	
3) Word frequency effects in reading Chinese.....	213
4) Reading and word recognition in Chinese.....	214
5) Reading strategies and reading efficiency in Chinese.....	215
Toward a word recognition model of Chinese.....	216
1) Interface between the visual pattern and its linguistic domain.....	218
2) Tree-like structure for the organization of the words and the units.....	219
The implication of the present study.....	221
1) On experimental studies-Some methodological concerns.....	221
a) <i>Language statistics on word-stimuli</i>	
b) <i>Orthographic attributes of word-stimuli</i>	
c) <i>More stimuli</i>	
2) On neuropsychological studies on dyslexia in Chinese.....	223
a) <i>Component-associated phonological reading errors-"Half-word" reading and reversed reading errors in Chinese</i>	
b) <i>Word-associated and CI-associated semantic errors in Chinese</i>	
c) <i>Reading tests to study dyslexia in Chinese</i>	
3) On acquisition of Chinese words.....	226
a) <i>A new analytic method for acquisition of Chinese words-the importance of recognition of lexical radicals</i>	
References.....	228
Appendix I: Critical stimulus pairs in Experiments 6 & 9.....	241
Appendix II: Two-character CI in three semantic categories in Experiment 11.....	244

Acknowledgements

There are so many who have helped me, in one way or another, in the course of this research, and to whom the thanks are due.

I would like to thank both my supervisors, Drs John Marshall and Alan Allport, for their conscientious guidance, support and warmth of the friendship throughout the course of this work. In particular, I am greatly indebted to Dr. Allport from whom I have learned more than just psychology.

I am most grateful to my former supervisor, Professor Ruth Campbell, whose warmth and support gave me the courage and inspiration to start this research. Her valuable discussion at the early stage of this research is also greatly appreciated.

In various phases of this study, I had useful discussion with many colleagues in the Department of Experimental Psychology. My special thanks to Dr. Michael Treisman for the useful discussion in the early stage of this research. My thanks also go to Drs Dahlia Zaidel and Eran Zaidel for their valuable discussions and supports.

At last but not least, I would like to place on record my special gratitude to my husband, Zhengming, for his continuous support and understanding with good humour during this research and the course of writing this thesis.

The arrival of my daughter, Lucia, during the period of this research, brought me great joy and inspiration. Her tenderness and love make this work so rewarding and to her this thesis is dedicated.

This research is sponsored by Sino-British Friendship Scholarship.

Word Recognition and Reading in Chinese

Yi-Ping Chen

Green College, University of Oxford
(A thesis submitted for the degree of Doctor of Philosophy)
Hilary term, 1993

Abstract

An important issue in contemporary studies of reading is whether (and, if so, how) different writing systems affect the cognitive processes involved in reading. The current view is that reading in non-alphabetic and alphabetic orthographies may involve similar cognitive processes. Starting from current word recognition models, established for alphabetic languages, the present thesis provides a theoretical and empirical analysis of word recognition and reading in Chinese.

The thesis first proposes that the orthographic units of Chinese words are the smallest-stroke-patterns, rather than individual strokes as conventionally thought. Over 90% of Chinese words are compounds (called ZI in Chinese), constructed from the combination of primitive ideographs. The two major orthographic components, termed the *lexical radical* and the *non-radical component* are essential components of such compound words. Secondly, this thesis proposes two orthographic rules called the *Part-to-Whole* Phonological correspondence rule and semantic correspondence rule. These rules specify that the lexical radical and the non-radical component of a word are the semantic and phonetic components, respectively. Thirdly, the thesis proposes that pronunciation of Chinese may be achieved via the rule-defined phonetic component, according to two phonological recoding rules called "*derivation*" and "*analogy*".

The empirical analysis (ten experiments) explores visual, phonological and semantic processing of Chinese words (single characters) through simple comparison tasks. Visual matching of Chinese words shows that the orthographic units used by skilled native readers are stroke-patterns, not the individual strokes. As in English, a direct route from print to sound and another from print to meaning also exist in Chinese. Subjects (skilled Chinese readers) also showed a selective attentional bias toward the rule-defined semantic and phonetic components, respectively, in semantic and phonological comparison tasks. This attentional bias was stronger for the non-radical component, in phonological comparison tasks. This was confirmed by lexical decision on two-word phrases (called CI). The attentional bias in the lexical decision task was asymmetrical, and was found only on the non-radical component. It is proposed that lexical decomposition may involve segmenting the non-radical component away from the lexical radical, while leaving the lexical radical as a fixed base.

In an approach toward reading larger units, the last experiment investigated lexical decision with two-character CI. This experiment investigated not only the effects of word-frequency, concreteness and homophony but also their differential effects in individual subjects, as indices of reading strategies. The lexical decision performance of skilled Chinese readers was significantly correlated with individual subject's susceptibility to frequency and concreteness effects but not with homophone effect. As a reading strategy, a bias towards semantic recoding is correlated with faster performance while a phonological strategy is correlated with slower performance. These two strategies themselves had an inverse association. Lexical decomposition as an optional strategy was significantly associated with the phonological strategy.

Finally, the present thesis develops an analytic account of word recognition and reading in Chinese. Implications are drawn for both experimental studies of normal reading and neuropsychological studies of dyslexia in Chinese.

Word Recognition and Reading in Chinese

Yi-Ping Chen

Green College, University of Oxford

(A thesis submitted for the degree of Doctor of Philosophy)

Hilary term, 1993

Long Abstract

An important issue in contemporary studies of reading is whether (and, if so, how) different writing systems affect the cognitive processes involved in reading. The current view is that reading both non-alphabetic and alphabetic words may involve similar cognitive processes (*Hung & Tzeng, 1981; Flores d'Arcais, 1992*). However, cross-orthographic studies on the issue of the "same" or "different" processes by no means fully answer the question of how words written in non-alphabetic scripts (such as Chinese) are read. This thesis provides a theoretical and empirical analysis of word recognition and reading in Chinese. In order to test the wider applicability of current word recognition models, established on alphabetic languages (mainly English), this thesis examines a number of aspects of reading Chinese words (single-character compound words). In contrast to the wholistic view of reading Chinese words, the thesis develops an analytic account of word (and two-word CI) recognition, involving rule-based lexical decomposition of compound (single-character) words.

According to current connectionist ideas, word recognition involves the activation of orthographic, semantic and phonological information in parallel over a set of primitive representational units (*Seidenberg & McClelland, 1989*). At a different, more abstract level of analysis, the recognition and naming of written words may be described in terms of 'rules', subject to different writing systems, such as the Grapheme-Phoneme Correspondence rules in English. In examining such approaches toward word recognition with Chinese words, a key problem is

that the orthographic structure of single words (characters) has not yet been specified for Chinese orthography. The present thesis therefore begins by characterising the orthographic knowledge that may be relevant to reading.

Chinese words are called WEN and ZI in Chinese, indicating two classes. In fact, only 5% or so of Chinese words are the so-called ideographs (WEN), made up from individual strokes. The vast majority of Chinese words (ZI) were constructed from combination of the set of primitive ideographs (WEN). The structures of these WEN thus forms a level of orthographic units intermediate between words and individual strokes. These units can be called stroke-patterns. Thus the present thesis proposes that the orthographic constituents of Chinese words should be considered these stroke-patterns, rather than the individual strokes, as is often thought to be the case. The present study is concerned primarily with ZI, i.e., the vast majority of Chinese words. All ZI can be divided into two principal orthographic bodies-the lexical radical and the remainder: *the 'non-radical component'*. The lexical radical is the essential component of all ZI. However, the status of a given component as a lexical radical can only be defined in the word context, according not only to its form but also its position in the different words that contain it.

Secondly, this thesis classifies three types of rules in Chinese orthography: 1) etymological rules for the composition of ZI; 2) Rules of Part-to-Whole Semantic and Phonological Correspondences; 3) Recoding rules for pronunciation of Chinese words, via the non-radical component. The etymological rules that are responsible for the composition of more than 90% of Chinese words (the compound ZI) are the ideocompound and ideophonetic rules. The Part-to-Whole Correspondence rules specify the semantic and phonetic components of Chinese words as the lexical radical and the non-radical component respectively, with words formed by the ideocompound rule as phonologically irregular words according to the Part-to-Whole phonological correspondence rule. Pronunciation

of a word via the rule-defined phonetic component (i.e., the non-radical component) can be achieved through "derivation" or "analogy", which are accordingly called recoding rules for pronunciation via the non-radical component.

A total of eleven experiments is reported in this thesis. Experiments 1 to 10 used Same-Different judgments to simultaneously presented word-pairs. This task requires subjects to judge whether the paired Chinese characters (which may also be pseudowords and nonwords) are the "Same" or "Different" in their visual form, or pronunciation, or meaning. All experiments (except Experiments 1 and 4) used computerised tasks with multi-factorial control of stimuli. In some of these experiments, not only the same subjects but also the same stimuli were used in different tasks. Thus, the different cognitive processes involved in word recognition could be examined orthogonally.

Experiments 1 to 3 examined visual comparison between pairs of Chinese words. The pilot experiment, (Experiment 1) showed an effect of the number of mismatching features, quantified in terms of the orthographic units defined in this study (the 'unit-dissimilarity' effect). Experiment 2 used not only words but also pseudowords and nonwords. In addition to the unit-dissimilarity in 'Different' pairs, visual complexity of 'Same' pairs was also manipulated orthogonally, in terms of both the numbers of units and the numbers of strokes (called 'stroke-complexity' and 'unit-complexity', respectively). In Experiment 2, characters consisting of either 2 units or 3 units can have the same or different numbers of strokes. Comparison of 2-unit characters was found to be significantly faster than 3-unit characters, even when the numbers of strokes in 2-unit and 3-unit characters was identical. However, there was no significant difference in either speed or accuracy of comparison between characters with different numbers of strokes, in either 2-unit or 3-unit characters. This suggests that the orthographic unit used for visual word recognition (in compound words, that is, in the large

majority of Chinese words) is the stroke-pattern rather than the individual strokes, at least for skilled readers of Chinese (such as the subjects in this study).

As in English, recognition of orthographic units of Chinese words also benefits from a word context (the word-unit effect). Evidence for this is that the unit-based effects (unit-complexity and unit-dissimilarity) were stronger in comparisons between real words than between pseudowords or nonwords. Moreover, words were also judged generally more quickly and accurately than the pseudowords, and pseudowords were faster and more accurate than nonwords. These two types of word superiority effects reflect the mutual interactions between recognition of orthographic units and of words, as proposed by some interactive models (*e.g.*, *McClelland & Rumelhart, 1981*). The demonstration of both types of word superiority effects with single-character Chinese words suggest some functional parallelism between the orthographic units of Chinese words and the letters of English words, in word recognition. It also suggests a wider applicability of models characterised by the mutual interaction between orthographic units and words (*e.g.*, *McClelland & Rumelhart, 1981*). A pseudoword advantage over nonwords, between which the sole difference is the position of the lexical radical, is also shown in Chinese. This finding indicates that recognition of the orthographic units of Chinese words (such as the lexical radicals) is highly sensitive to positional information. Experiment 3, with exactly the same procedure and stimuli as Experiment 2, used non-Chinese readers, for whom Chinese words are merely graphic patterns. After quite limited practice (240 pairs or so), the unit-dissimilarity effect reached significance, implying that the graphic patterns of the orthographic units of Chinese words can be used even by non-Chinese (native English) subjects, on the sole basis of visual familiarity. These results suggest that the structure of the orthographic units may be salient as graphic forms, independent of their underlying linguistic features.

Experiments 5 to 10 explored semantic and phonological comparisons in three paired experiments. The first paired experiments (Experiments 5 & 8) demonstrated that semantic relatedness (synonymy) between two words to be compared had no significant effects on their phonological comparison (Experiment 5). Likewise, relatedness in pronunciation (homophony) had no significant effects on semantic comparison (Experiment 8). These results suggest that, as in English, there are direct routes between orthography and phonology as well as between orthography and semantics in Chinese.

The second pair of experiments (Experiments 6 & 9), used physical matching on either the lexical radical or on the non-radical component, as distractors, to bias subjects toward false 'Same' responses. Subjects showed the opposite selective attentional biases in semantic and phonological judgment tasks, respectively, even though the actual stimuli employed in the different tasks were identical. They were significantly slower, and had higher false 'same' response rates, in phonological comparison of word-pairs when the non-radical component was the distractor (the Word-Component match condition) than when the lexical radical component was the distractor (the Word-Radical match condition) (Experiment 6). In contrast, the same subjects were significantly biased by the lexical radical component and not by the non-radical component as the distractor, in semantic comparisons (Experiment 9). The opposite effects of the lexical radical and the non-radical component as the distractor, respectively, were also shown, on phonological and semantic comparison tasks, using word-pseudoword pairs (Experiments 7 & 10). Therefore the matching radical and non-radical components function quite differently as distractors in semantic vs. phonological tasks.

These results are in accord with the semantic and the phonological rules of Part-to-Whole correspondences proposed in this study, which specify the lexical radical and non-radical component as semantic and phonological components

respectively. The non-radical component was equally effective as a distractor in phonological comparisons when the word was an ideocompound word, in which both components (including the non-radical component) are semantic components, by the etymological ideocompound rule. Therefore, the rules relevant to, or used by subjects, in the recognition of written words appear to be the orthographic rules, which describe the regular orthographic structure of written words, rather than etymological rules, which define the function of these orthographic components in terms of their word-formation origins. Furthermore, the present study demonstrated that the influence of a matching component as a distractor, in the phonological judgment task, is a function of the frequency-index of that component as a lexical radical in all the words contain it (called the radical frequency index).

In an approach toward reading larger units, the last experiment (Experiment 11) investigated lexical decisions with two-word phrases (called CI in Chinese). This experiment had two aims. First, it investigated whether or how the frequency, concreteness, and homophony of two-character CI, and of the components of the single-character words, may affect lexical decisions on two-character CI, when a specific linguistic task such as semantic or phonological judgment is not required. Firstly, lexical decision on two-character CI was affected by manipulation of the components of single-character words. The position of the same components in single-character words, regular or irregular, had significant effects on the lexical decision task. However, the position of the same set of radicals in the two-character CI, in terms of beginning or middle of the two-character CI as a whole, had no significant effects in this task. These findings suggest that reading larger units in Chinese (e.g., two-character CI) is based on the processing of individual characters. The individual character is the basic reading unit in Chinese. It should be called a 'word' in English rather than a 'character'. Secondly, lexical decision with two-character CI were found to be significantly influenced by CI-frequency

and concreteness but not homophony of the CI. Consistent with recent studies of word frequency effects in English, performance on CI-homophones was not significantly influenced by their base-CI frequency.

A further analysis of the data in this experiment investigated the association between reading strategies and reading efficiency, as well as the associations among different reading strategies. Individual bias in attending to different information in the CI (i.e., different reading strategies) were measured by the magnitude of corresponding experimental effects in individual subjects. Orthographic-lexical frequency dependence was indexed by the effects of CI-frequency for individual subjects. Semantic and phonological strategies were indexed respectively by the magnitude of the concreteness effect, and of the CI-homophone effect, in individual subjects. Dependence on lexical decomposition was indexed by the radical-migration and non-radical-migration effects, in individual subjects. Finally, reading efficiency was measured by the latency and accuracy of lexical decision on 180 real CI, for each subject. The semantic strategy was found to correlate with faster performance in the lexical decision task. By contrast, the phonological strategy was found to correlate with slower performance in the lexical decision task. A lexical decomposition strategy was also associated with slower performance. Hence, in skilled (silent) adult reading in Chinese (at least for CI), subjects biased towards orthographic-lexical processing and semantic recoding may be more efficient readers. However, caution should be exercised in generalising any of these results to the acquisition of reading skill in children, or to other reading tasks.

As to associations between reading strategies themselves, an inverse association between the semantic and phonological strategies was found. Moreover, the strategy of lexical decomposition, indexed by the component-migration effect, was significantly associated with a phonological strategy. Lexical decomposition appeared to affect the lexical radical and the non-radical

component asymmetrically. That is, perceptual recombination of a component of one word with the other whole word, to form a 'illusory conjunction' pseudo-CI, occurred in respect of the non-radical components, but not at all in respect of the lexical radical. The lexical decomposition processes in Chinese may be asymmetrically focused on the non-radical component and it is only important for phonological recoding.

The present study suggests the wider applicability of models for word recognition established on English. Nonetheless, there are also orthography-specific features of word recognition in Chinese. In semantic and phonological recoding tasks, there appears to be a selective attentional bias in skilled readers of Chinese, respectively, toward the rule-defined semantic and phonetic component—the lexical radical and the non-radical component in compound words, depending on a lexical decomposition process. In the lexical decision task, lexical decomposition affected the lexical radicals and the non-radical components asymmetrically. This decomposition process may involve segmenting the non-radical component as a whole away from the lexical radical, leaving the lexical radical fixed as the base. This kind of lexical decomposition process may be unhelpful for semantic recoding.

The analytic account of word recognition in Chinese, developed in this thesis, may serve as a theoretical basis for a Chinese version of parallel distributed processing models for word recognition, in the future. There are some important implications both for experimental studies of normal reading and for neuropsychological studies of dyslexia in Chinese. The results also offer some support for a new analytic method for learning to read Chinese.

Chapter 1

The Chinese language

Introduction

The subject of this thesis is the orthographic recognition of Chinese words. Since most readers of this thesis will not be native readers of Chinese, it is first necessary to introduce the Chinese language.

Chinese, which is spoken as a first language by a larger number of people than any other language, also possesses venerable and extensive literatures. Chinese has played a role in east Asia comparable to that of Latin and Greek in Europe. The Chinese language has already been widely discussed in English (*e.g.* *Karlgren, 1971; Martin, 1972, Wang, 1971, 1973; Henderson, 1982; Tzeng and Hung, 1980, 1981; Hung and Tzeng, 1981; Hoosain, 1991; Flores d' Arcais, 1992, Hu & Huang, 1990*). This chapter accordingly provides a general rather than a comprehensive introduction to Chinese, including both spoken and written Chinese. The description of spoken Chinese focuses specifically on those features which are distinctive for Chinese. The introduction to written Chinese centres on the reforms of written Chinese, which have not yet been fully described or taken account of, in the current English literature.

Chinese speech

To the ears of western speakers, Chinese speech may sound rather simple and melodious. The basic speech unit in Chinese is the syllable, which can be a single vowel (V) or one consonant plus vowel (CV). There are no consonant clusters (CC) like [sp, bl,]. Three other features contrast with English. These are mutually unintelligible dialects, gross homophony and tones.

1) Dialects and the Common language (Mandarin)

There are eight major dialects in spoken Chinese. The vocabulary items (typically combinations of multiple characters) used in the different dialects for expressing the same concept or object can be totally different, so that different dialects can be mutually unintelligible. Within one dialect, there may also be different accents. Like different accents in English, however, these are usually mutually intelligible. Despite the eight dialects and many accents in Chinese speech, a standard spoken language has existed in China for more than 2000 years. Thus, according to early Chinese literature, the great master Confucius (250 B.C.) taught his philosophy to his students in "elegant language" [jia]3 [jian]2, the "Queen's English " of that time. This standard speech has been known as "elegant language" 雅言 [jia]3 [jian]2, or "official language" 官话 [guan]1 [hua]4 in ancient times, as the "national language" 国语 [guo]2 [yu]3 in Taiwan, as 华语 [hua]2 [yu]3 used in overseas Chinese society and in the latest version (as the) "common language" 普通话 , [pu]3 [tong]1 [hua]4, in mainland China at present. Due to obvious political reasons, the standard speech has always been based on the local dialect of the capital. Given that Beijing was the capital for the YUANG, MING and QIN dynasties, the standard language of Chinese has been based on the northern dialect, and Beijing accent, for the last eight centuries. Apart from political reasons, the literature written in the Beijing dialect may also have played an important role in promoting the northern dialect to the standard speech of Chinese. The conventional translation in English of standard Chinese speech is *Mandarin*. This, however, may well be a historical mistake, since *Mandarin*, which should refer to the native speech of Mandarin, is taken to refer to the native speech of the majority Han Chinese. Mandarin, who were not Chinese, conquered China and established the last empire in China-the QIN-(1664-1912). However, they acquired the Chinese language and were assimilated by the Chinese.

In this thesis, standard spoken Chinese will be referred to as the "common language" [PU-TONG-HUA]. According to a recent survey, 70 per cent of the population speak the northern dialect, i.e., 70 per cent of the Chinese population use the common language as their native tongue. However, most overseas Chinese speak the Canton dialect. In 1955, for the purposes of education and official communication, the national language committee of the People's Republic of China stipulated the northern dialect, and Beijing accent, as the "common language", [pu] [tong] [hua], the standard spoken language for China. Since then, PU-TONG-HUA has been widely used in education, the mass media and formal communication. The common language PU-TONG-HUA has 22 consonants and 38 vowels and 1200 syllables; it is the simplest of the eight dialects according to one recent linguistic study (*Huang & Hu, 1990*).

As for the relationship between spoken and written Chinese, in fact only the northern dialect, i.e., the common language, has a written format that corresponds directly to this spoken form, particularly as regards colloquial language. The vocabularies (i.e., the combinations of multiple characters) in other dialects, if written down to correspond with their pronunciation, would be either visually meaningless or have a different meaning from the spoken vocabulary. For instance, 洗手 WASH HAND is pronounced [xi]3 [shou]3 in the northern dialect whereas in Shanghai dialect it is pronounced [da] [sou] (tone in Shanghai dialect is insignificant). The written two-character CI, 洗手 when read aloud in PU-TONG-HUA would be [xi]3 [shou]3, identical to the pronunciation in northern dialect. Yet the Shanghai dialect [xi] [sou] for WASH HAND, if written down as pronounced, would be 打手, which means BEAT HAND. Hence, vocabularies in dialects other than the northern dialect (the common language) do not have a written format which corresponds directly to the spoken form. However, it must be emphasized that the difference among vocabulary items used in different dialects are primarily in the combinations of multiple characters. The

pronunciation of the individual written words in the different dialects is a matter of accents, that is, variation in phoneme values, over a limited range of distinctive features. For instance, WASH HAND in Beijing speech is [xi]3 [shou]3 and in Shanghai speech is [da] [sou]. In reading aloud the written word 洗手, Beijing speakers would read [xi]3 [shou]3 while Shanghai speakers would read [xi] [sou]; the two forms differ merely in the systematic change of accents. Further evidence for this claim can be seen in the fact that literate Chinese who speak different dialects can appreciate equally well the melodious tones and rhyme of Tang and Song dynasty poems, which were written about 900 years ago. Hence, the influence of dialects in reading is by no means as important as it is in speech.

2) Homophony

Words distinct in meaning but which sound exactly the same are known as "homophones". Homophony in Chinese has been considered particularly problematic for reading comprehension. For instance, Chao (1968) noticed 116 distinct characters under the syllable [si] in Goodrich's pocket dictionary, and he actually constructed a written Chinese text using 161 homophones, which if read aloud would consist of repetitions of the one sound [si]. One can imagine the difficulty of reading comprehension for such text. However, the homophone problem has been somewhat exaggerated in the literature. Firstly, the number of homophones such as [si] may be overestimated due to inconsistencies between character lexicons in the different dictionaries. In the XIN-HUA dictionary, the most popular and authoritative dictionary for contemporary Chinese in China, only 39 distinct characters can be found under the syllable [si], including all four different tones (*Xin Hua*, 1979). Some characters included as homophones of [si] in the Goodrich dictionary must be archaic, even for readers with large vocabularies. The number of characters in practical use in modern Chinese is in fact no more than 7000, according to a recent survey by the Chinese National Language Committee (*Hung & Hu*, 1990). Based on the XIN-HUA dictionary

(1979), the largest number of homophones for individual characters is the pronunciation [jiao], under which there are 52 distinct characters, including all four different tones. Most individual characters usually have no more than 10 homophones including all four tones, and fewer than five, excluding words pronounced with different tones. Secondly, the effect of homophony between individual characters becomes still less important in practice, since modern spoken and written vocabularies are usually based on combinations of two or three characters. As a matter of fact, homophones of these two-character CI, e.g., 丹唇 RED LIP [dan]1 [shuen]2 and 单纯 SIMPLE [dan]1 [shuen]2 are far fewer than those of single characters, and are probably no more common than English homophones such as KNEW and NEW. Hence, homophones are unlikely to constitute a particular problem either in Chinese speech or in the reading of modern texts in Chinese.

3) Tones

Tone is a prominent feature of Chinese speech (PU-TONG-HUA). The number of tones varies from one dialect to another. For instance, in Canton dialect there are six tones, while in Shanghai dialect tone is almost absent. In PU-TONG-HUA, there are four tones. They are conventionally represented by the symbols –, /, ∨, and \, indicating flat, up, down-up and down, to designate the tones 1, 2, 3, and 4. The most peculiar tone is the third one, which falls and then rises. No equivalent phonetic feature can be found in spoken English. However, the second is similar to the rising tone used at the end of questions in English, while the fourth tone is similar to simple stress in an English word. The first tone is simply neutral (without distinctive intonation). Words with the same pronunciation but different tones can be totally different words. For instance, the pronunciation [shi]1 in the first tone can be 师 MASTER, in the second [shi]2 can be 食 FOOD, in the third [shi]3 can be 屎 FAECES, and in the last [shi]4 can be 事 THING. In theory, tones can be used to distinguish the different words with the same pronunciation.

However, in real speech, the distinction between tones is not so important, since spoken vocabularies are usually combinations of two or three characters. For instance, [lao]3 [shi]1(老师) and [lian]2 [shi]2 (粮食) will be spoken for MASTER and FOOD respectively instead of 师 [shi]1 and 食 [shi]2, which only differ in their tones. The relatively insignificant role of tone in speech may also be deduced from the fact that tone is absent in some spoken dialects of Chinese, such as the Shanghai dialect, without causing problems for oral communication in that dialect.

In summary, although dialects, homophones and tones appear to be prominent features of Chinese speech, their influence upon speech and reading is not unique to Chinese.

Written Chinese and its reforms

The vertical columns of characters, which have to be read from right column to left, as usually described in the English literature, no longer represent modern Chinese writing, at least in the People's Republic of China. Written Chinese reforms started at the beginning of this century. However, it was not until 1955 that the planned reform of written Chinese had been systematically carried out all over China (except Taiwan). The costs and benefits of this reform need not be discussed here. However, to understand the modern Chinese used in the People's Republic of China it is necessarily to give a brief introduction to the reform of written Chinese. This introduction to the written Chinese reform centres on the following four aspects:

1) From literary style to colloquial style

The written form of individual Chinese words has been remarkably constant. Accordingly, the Chinese reader can understand the meaning of words used in DAO text, written in 600 B.C. Nonetheless, the structure of writing in ancient

different from speech in terms of both vocabulary and grammatical structure. This literary style of writing was called 文言 [wen] [yan]. In WEN YAN, the individual characters were used as independent words in a specific grammatical structure. The literary style of writing was not a direct record of speech and it could be mastered by only a small segment of the population. However, another writing-style called 白话文 [bai] [hua] [wen], meaning "plain or colloquial" has also existed in China for at least 800 years. The colloquial style of writing is a more direct record of speech; more exactly, it transcribes the common language, PU-TONG-HUA. As in speech, the vocabularies used in the colloquial style of writing are also combinations of two or three characters. The colloquial style of writing was initially used only when writing entertainments (e.g, novels). Given the concern with increasing the level of literacy, the colloquial style of writing, which is obviously easier to acquire, became more and more popular. In 1918, the "new culture movement" promoted the colloquial style of writing for formal literature, including poetry. Since then, the colloquial style of writing has been formally adopted in Chinese writing. At present, writing in the literary style is extremely rare and it is only used for special purposes. Most literary works written in the literary style have to be translated into colloquial style in order to be understood by contemporary Chinese readers. Writing in the literary style is no longer conventionally taught in schools, although the reading of literary style text still is.

The arrangement of the written characters in text was also changed at the same time as the colloquial style of writing was introduced. The traditional vertical columns are now more frequently replaced by horizontal rows. The use of horizontal rows, from left to right, started in 1919 and has applied to all newspapers and printed books in China since 1955 (*Hung & Hu, 1990*). Given that single characters are printed in the same size and (square) shape, they are suitable for both horizontal and vertical printing. For some newspapers and magazines, especially for those published in more traditional Chinese societies

such as overseas Chinese communities and in Taiwan, the vertical columns are still used. For vertical columns, reading is still from right to left at present.

2) From sophisticated to simplified characters

Simplification is the natural evolution of Chinese orthography. The early evolution of pictographic script to simplified ideographic script in Chinese, e.g., from Tiger 𧇧 to 虎 is already a simplification. In ancient times, sophisticated and simplified versions of the same words existed in parallel for different purposes (practical and aesthetic, for example). Sometimes, sophisticated and simplified words were regarded as different calligraphies: for example, 'NO' is written as 無 in the KAI version and as 无 in the QI version. A simplified format, originally called 俗体 *plain version*, was very popular in the Tang dynasty and almost every subsequent period in history. However, it was not until 1956 that the planned and more thoroughgoing simplification of Chinese words was undertaken in the People's Republic of China. Given that a standard norm of simplified characters had been announced to replace the different versions of simplified characters that were already in public use, the simplification should be regarded as a standardization of Chinese words. Current forms of simplified characters in China are based on the norms for 2238 words, announced by the National Language Reform Committee in 1964. The simplification is rule-governed, and was put into practice gradually (*Huang & Hu, 1990*). Most importantly, instead of changing the form of individual characters, the simplification involved a relatively small number of components, called radicals, which recur frequently in different characters. For instance, among the 2238 standard simplified characters, 1754 are generated from just 132 simplified components (the so-called radicals). That is, among the 2238 simplified characters, 1754 are different from the sophisticated form only in one component such as 達 and 达. The simplified and sophisticated forms of Chinese characters are, as it were, mutually intelligible. Thus literate Chinese, who learned to read and write with the sophisticated form of Chinese

characters before 1956, do not need to go back to school in order to learn the characters in the simplified form.

3) Phonetic representations of characters-[Pin] [yin]

From many English words (so-called regular words), one can obtain the sound of the word by simply assembling the sounds of the constituent letters according to rule ('regular' words). Letters are usually learned before words (or at least, before any substantial vocabulary of written words is acquired). However, in Chinese, no such set of components is acquired prior to the acquisition of words. The pronunciation of Chinese words is taught entirely by "look and say". Many attempts have been made to construct a relatively small set of sound symbols, so that the sound of Chinese words can be obtained by assembling a small number of components. In history, there have been at least four proposals. They are Welsh phonetics, Roman letters, simple ancient Chinese characters, and 拼音 [pin]1 [yin]1 ASSEMBLING SOUND. The most influential are Welsh phonetics and PIN-YIN. The system that used ancient Chinese Characters such as ㄅ, somewhat like Kana, called [zhu]4 [yin]1 [zi]4 [mu]3 注音字母 in Chinese, has only been used in Taiwan. The Roman letters are no longer used at the present time.

In fact, the use of alphabetic letters to represent the sounds of Chinese words began with foreign missionaries, for the purpose of teaching the Chinese to read the Christian Bible, about 1610. The most influential proposal was Welsh phonetics, proposed by Thomas Welsh, a British ambassador to China. It was used in the first instance to teach not only English but also other European diplomats to speak Chinese. Therefore the sound of the letters is based on both the pronunciation of English and of other European languages. The Welsh proposal was used to represent Chinese personal names and place names in both Chinese and western literature until 1958. In some current English literature it is still used to represent the sound of Chinese. In 1957, PIN YIN, a set of 39 alphabetic letters

and their compounds (e.g., ch, zh, eng), was announced by the Chinese language reform committee. It aimed to promote the common language Pu Tong Hua in speech and reading. The sounds of letters in Pin-Yin correspond approximately to phonemes. Since 1958, Pin-Yin has been widely taught in schools all over China. In mainland China it is taught independently of the characters and usually before the acquisition of characters. In Chinese literature, the names of foreign peoples and districts are now transcribed in Pin-yin. Most English literature now also tends to use Pin-yin instead of Welsh phonetics to represent Chinese place names or specific terms. For instance, Beijing in PIN-YIN, instead of Peking, is now used to represent the capital of China in most English texts. Pin-yin is now widely used to represent the sound symbols of Chinese words. However, it must be born in mind that PIN-YIN represents merely the abstract sound structure of Chinese characters. It cannot be regarded as an alternative written form of Chinese for general use. Written texts in Chinese publications still use characters, but not PIN YIN. Even the combination of characters and alphabetic letters such as, e.g., 喝茶 [he cha] for DRINK TEA, is not acceptable in Chinese, except for commercial use. The function of PIN YIN for most Chinese readers is mainly to learn the sound of new words or to find standard Pu Tong Hua pronunciations of known words in the Chinese dictionary.

4) The Chinese lexicon-single character ZI and multiple character CI

The lexicon in Chinese includes the single-character words, called 文 WEN (simple characters) and 字 ZI (compound characters) and the combination of two characters (sometimes three characters) 词 called [ci]². In English writings, the single character word is sometimes described as a single "morpheme" rather than a "word" (*Hung & Tzeng, 1981*). However, the term "word" is also sometimes used to refer both to single characters and to combinations of characters, CI. This has given rise to confusions such as the conflicting claims that all Chinese "words" are monosyllabic (single characters) but that modern Chinese "words" (CI) are

usually polysyllabic. However the idea that the individual character corresponds to the "morpheme" rather than the "word" can be contested, for at least two reasons. Firstly, the "morpheme" is by definition the smallest meaningful unit in language. However, (as discussed further in Chapter 2), the smallest meaningful unit can be (and often is) smaller than the individual character: that is, in all compound characters (ZI), which make up over 90% of the characters. Secondly, individual characters can be used as independent "words" in the literary style of writing. Although, in the modern colloquial style of writing, the vocabulary is more frequently formed by the combinations of characters, or CI, both single character words (ZI) and multiple-character CI for the same concept can exist in parallel in the lexicon. They are used in speech and writing as different styles, literary or colloquial. For instance, SUN can be expressed as 日 [ri]4, or as 太 阳 [tai]4 [yang]2; OBJECT can be expressed as 物 [wu]4 or as 东 西 [dong]1 [xi]1; RUN can be 跑 [pao]3 or 跑 步 [pao]3 [bu]4. In CI, the constituent characters are both spoken and written independently. For instance, CI for COMPUTER in Chinese is 计 算 机 where the constituent characters individually mean CALCULATE COUNT MACHINE respectively, and are pronounced [ji]4 [suan]4 [ji]1. The meaning of this three-character CI is the combination of the meaning of all three characters, while its pronunciation is the pronunciation of the individual characters in sequence with no change or blending of the sounds of individual characters. To some extent, the CI is comparable to the phrase in English. Most of the combinations have become more or less fixed expressions in speech and writing. For example, 东 西 [dong]1 [xi]1 for OBJECT (literally EAST WEST) is a common fixed vocabulary item in Chinese. The meaning of the combination cannot always be understood from the literal meanings of the constituent characters. An English translation of the word "CI" is not currently available. In the present study, the term "word" will be used to refer to single characters only, without further specification.

As for the size of the Chinese lexicon, it should be taken to include both single-character words and two-character or three-character CI. According to some authorities, the number of words (single characters) can be as many as 80,000. However, the core lexicon for practical usage is far smaller than this legendary number. Many efforts have been made to select a core lexicon based on the most useful characters. For instance, "A thousand word textbook for peasants" was published in 1908. In 1921, Chen Huo-Qin first selected 4261 commonly used characters from the lexicon of Chinese, by using statistical analysis of the frequency of characters in modern Chinese writing. In 1951, 2000 common characters were stipulated as the normative or essential literary level in China. The latest "contemporary norm for Chinese words", which was published by the National Language Committee in 1988, has selected 7,000 characters as the core lexicon for education, for word-processing of Chinese characters and for literature search index systems. The most popular dictionary in mainland China, the Xin-Hua dictionary, contains 11,000 characters, which includes all 7,000 characters from the 1988 norms (*Hung & Hu, 1990*).

Generated from combinations of single-character words, the number of CI which are available in speech and writing is almost uncountable. The number of CI increases as new concepts are invented. In Chinese, new concepts are expressed by constructing new CI from existing characters. For instance, when the automobile was invented, a new word had to be created (or an old one, CAR, adopted in use) in English. In Chinese, the existing two characters 汽 车 STEAM VEHICLE were used together to form the new two-character CI for automobile. A similar way of creating new words in English is to construct compound words such as ARMCHAIR or BLACKBOARD. However, word-compounding in English is limited in application, and in length; long compounds are rare.

Although CI consist of two or three characters, each character is written independently, in the same size, with no hyphen between them. The acquisition of CI (perhaps similar to learning new compound words in English) may be easier than that of truly new words in English, because the constituents of CI in Chinese are known words. The pronunciation of CI is simply the pronunciation of the individual constituent characters independently. The meaning of many CI may be obtained from the combination of the meanings of constituent words, such as CALCULATE COUNT MACHINE for COMPUTER. However, this is not always the case (as in EAST WEST for OBJECT). A typical character can often occur in more than five different CI. For instance, 寒 COLD can occur in 寒冷 COLD, 贫寒 POOR, 寒 颤 FEAR, 寒 心 DISAPPOINTED, 寒 战 SHIVER, 寒 伦 DISGRACEFUL. Thus the acquisition of literacy in Chinese involves learning only 2000-4000 single-character words, since these 2000 or 4000 characters could occur in over 10,000 CI available for speech and writing in Chinese.

5) The sorting systems of the Chinese lexicon

There are four ways of sorting or classifying single-character words in Chinese dictionaries. These are, respectively, the radical index, the PIN YIN index, the four corner code index and the stroke-shape index (*Hung & Hu, 1990*). The radical index is the most useful. It has been used by almost every Chinese dictionary from the ancient KAN XI (A.D.1776) to the present XIN HUA (1979). The radical index consists of 189 radicals, according to the XIN-HUA dictionary (1979) and this number is usually constant among different dictionaries. The PIN YIN index is usually used in parallel with the radical index in dictionaries for reference to pronunciation. In the XIN HUA dictionary, characters with the same pronunciation are arranged together in a PIN-YIN index. A word with unknown pronunciation can be found with the radical list. In the radical list (*e.g., XIN HUA, 1979*), the 189 radicals are ordered according to the number of strokes they

contain. Under each radical, words with the same radical are ordered according to the number of strokes in the remainder of the word excluding the radical (i.e., in the non-radical component as a whole). For example, for the word 怡 JOY/[yi]2, its radical 忄 appears in the three-stroke-radical category, and the non-radical component 台 is counted as 5 strokes. Therefore, 怡 can be found in the 5-stroke category under the 忄 radical. The four corner code consists of four Arabic numbers, which is arbitrarily used to represent the different shapes of strokes. On the four corner index, words can be found from the shapes of the strokes at the four corners of the words. This system is now seldom used, at least in mainland China. The stroke-shape index is only used for special purposes, such as to input Chinese words into a computer system.

Chapter 2

Chinese Orthography

Written Chinese words are compact stroke-patterns. These patterns are written in squares, each of the same size, with no spaces or punctuation marks to separate them. To Chinese eyes, this square pattern, a written esperanto, is a word and also far more than just a word. To the eyes of alphabetic readers, the square pattern is conspicuously complex; it is radically different from any alphabetic "word" and indeed looks rather like a picture. Chinese orthography has accordingly been regarded as pictography, ideography and lexigraphy. It is usually regarded as a salient example of a writing system that is radically different from such alphabetic systems as English. With regard to the neuropsychology of reading, *"Brains may be similar from one language to another, but orthographies are certainly not"* (Coltheart, Patterson and Marshall, 1980). Before addressing the issue of whether the same neuropsychological processes are employed in reading Chinese and in reading English, it is reasonable to ask the question: To what extent, and in what ways, are words written in Chinese radically different from words written in English? To address this question, an introduction to Chinese orthography should centre on knowledge of the writing system as a component of the word recognition system, in Chinese and in English. This introduction, which has direct relevance to the main study, should also be regarded as an introduction to certain descriptive terms and concepts that pervade the present study.

Before turning to an introduction to the orthographic structure of Chinese, it may be useful to review briefly some current issues concerning the relationship between orthographic structure and word recognition.

Orthographic knowledge and word recognition

1) Orthographic units of written words

Many current theoretical frameworks for word recognition in the study of English orthography are letter-mediated accounts (*e.g.*, Morton, 1969; Coltheart, 1981; McClelland and Rumelhart, 1981; Henderson, 1982; Besner, Coltheart and Davelaar, 1984; Humphreys, Evett, Quinlan and Besner, 1987; Rayner & Pollatsek, 1987). A recent connectionist model for word recognition does not even include representational units for "words" (Seidenberg and McClelland, 1989), although this lacuna is debatable (Besner, McCann, Twilley and Seergobin, 1990; Seidenberg & McClelland, 1990). According to this connectionist model, word recognition involves interactive processing between three basic codes, orthographic, semantic and phonological. Each of these codes is represented as a pattern of activation distributed over a number of primitive representational units.

Explicit orthographic knowledge of letters is commonplace in English, since the acquisition of literacy in English usually begins with the learning of individual letters. However, the constituents of Chinese words are neither conventionally taught in school nor even consistently and systematically classified. Hence, in extending these letter-mediated models to reading Chinese (if another very different model is not necessary), the first fundamental question is: What is the equivalent of the "letter" in Chinese? As indicated by the widely used term "radical" in describing Chinese words, word constituents do seem to exist in Chinese orthography. However, the radical is usually classified according to its function, *e.g.*, the so-called semantic versus phonetic radicals. The orthographic structure of the radical has not yet been clearly and consistently classified. The classification of the basic structure of orthographic units in Chinese words thus becomes the first task of this thesis.

2) Rules relevant to reading

Written words are linguistic structures as well as visual patterns. As pointed out by Henderson (1987), current studies on word recognition have been characterised by the view that word recognition is the interface between the perceptual and linguistic domains (Henderson, 1987). In order to read, one must have acquired knowledge (or rules) of writing system. The best known example of the concept of orthographic rules in the study of word recognition concerns the grapheme-phoneme correspondences (GPC) of alphabetic orthographies. With regard to the idea of rules in reading, as pointed out by Glushko (1979), one should distinguish, " (a) rules as linguistic descriptions; (b) rules as knowledge of language structure; (c) rules as procedures or mechanisms of pronunciation". In order to address questions about the importance of rules in reading Chinese, the rules themselves should first be appropriately defined.

The remainder of this chapter offers an introduction to knowledge of Chinese words, which focuses on the clarification of the basic orthographic units and the three types of rules in Chinese orthography.

Orthographic knowledge of Chinese words

1) The Orthographic unit of Chinese words—the smallest stroke-pattern

Chinese words, each in a compact square pattern of strokes, are a salient example of a non-alphabetic writing system. In contrast to alphabetic words, words in Chinese are not conventionally spelt by a set of components, like the letters in English. However, this need not mean that such components are absent in Chinese. In fact, the principal word constituents in Chinese—analogueous to letters—are usually thought of as the individual strokes, as described by Martin (1972): "Wang (1971) perceptively notes that the average number of strokes of a character, reduced from six or seven to four or five in the new simplified shapes,

is now about the same number of letters in an English word, so that there is obvious parallelism in information content between the stroke and the letter". However, there are problems for this identification of the strokes as the significant orthographic constituents. Firstly, it must be born in mind that the segmentation of individual strokes in Chinese words is not always directly visible, and very often it is a matter of writing conventions. For example, a square □, which as a graphic pattern might be counted as four strokes, is actually counted as three strokes 丨, 丿, 一 in Chinese. Second, and more importantly, even when the individual strokes are graphically distinct, they always appear as elements in a fixed stroke-pattern (a constant spatial design) which occupy specific positions, such as three dots in a vertical line (always on the left side) or four dots in a horizontal line (always at the bottom) of the different words in which they occur, e.g., 江 河 泪 海 , 热 烈 杰 然 .

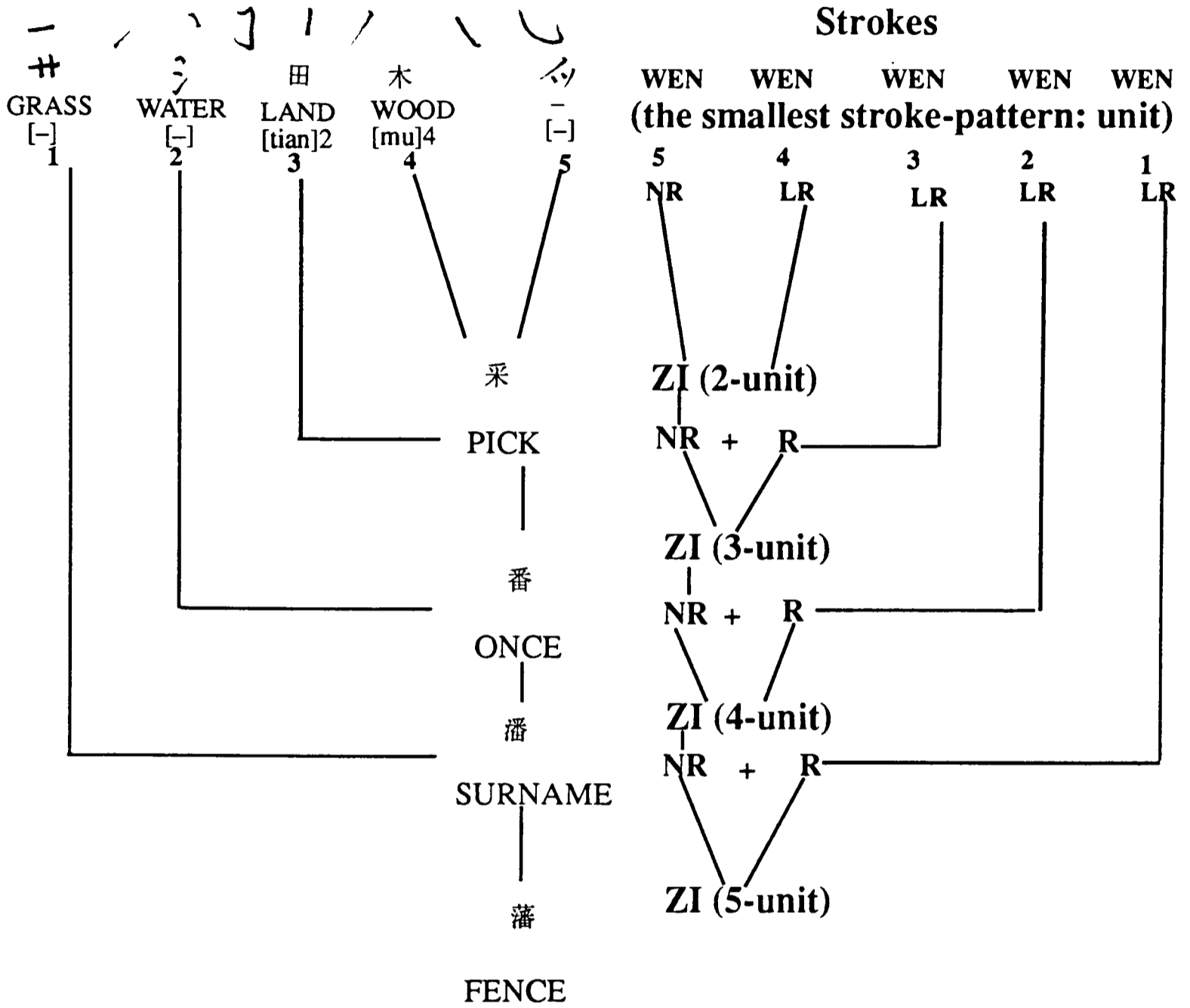
Third, from the etymological point of view, in Chinese, single-character words are called 文 and 字 [wen]2 [zi]4. These two terms indicate two distinctive categories of written words, with a different constituent structure (*Hu & Huang, 1990*). Figure 2 illustrates how words are formed hierarchically by individual strokes and by stroke-patterns.

As seen in Figure 2.1, about 3500 years ago or even earlier, the first generation of Chinese words (Figure 2.1) were formed out of individual strokes (*Hu & Huang, 1990*). These may be called pictographs or ideographs (e.g., 山

MOUNTAIN/[shan]1, 日 SUN/[ri]4, 月 MOON/[yue]4, 鸟 BIRD/[niao]3). These simple ideographs are called [wen]2 in Chinese (*Hu & Huang, 1990*). However, an important fact that should not be ignored is that these so-called ideographic or pictographic words only account for 5% or so of Chinese words in modern usage. More importantly, these simple ideographs are not only used in isolation as words, but are also used as the components in later-formed, compound Chinese words, called [zi]4. According to etymological studies, well

The formation of Chinese words—WEN & ZI

WEN ZI



Note: unit refers to the orthographic unit introduced in this study

LR: Lexical radical, NR: Non-radical Component

[-]: no phonetic value; -: meaningless

Figure 2.1. Illustration of formation of Chinese words—WEN & ZI

over 90% of current Chinese words are ZI. They were formed at a later time, with the primitive simple ideographs (WEN) as their constituents. This account was set out by Xu Shen in his book 说文解字 [shuo]1 [wen]2 [jian]3 [zi]4 'The description of WEN and analyses of ZI', (about 25 A.D.), which has been acknowledged as the most authoritative book on etymology for more than a thousand years (Hung & Hu, 1990).

WEN, the building-blocks of ZI, are the pictographic or ideographic symbols, not unlike the ideographic symbols familiar to alphabetic readers, such as \$ (dollar), % (percent), & (and), etc.. In contrast to the letters of English, WEN symbols were created in isolation as words, although they were later used as components in many more words. In isolation they are both pronounceable and meaningful, i.e., they have full lexical status. However, some of the ideographs in isolation are no longer used as words in modern Chinese. They are only used as components in words (e.g., 𠂇 in Figure 2). The structure of WEN is called "single body" (Hu & Hung, 1990). This is the smallest stroke-pattern, in that further decomposition would yield only individual strokes. For instance, 石 STONE/[shi]2 is already the smallest stroke-pattern despite the fact that it contains (or appears to contain) the even smaller pattern 凵 MOUTH/[kou]3. However, the decomposition of 石 would yield 凵 plus the individual strokes 丿, which do not form a legitimate, independent stroke-pattern in Chinese.

These smallest stroke-patterns are thus proposed as the basic orthographic units of all words in Chinese. (in what follows, the term 'unit' will be used, for brevity.) In this classification, the WEN are single-unit words, while Z I are words with more than one unit—usually from two to five units in the simplified version. Each unit itself is usually written in a square pattern. The composite spatial arrangement of the units in ZI are of three main types: left-right (明), top-bottom (宝) and outside-inside (国). Note that the top-bottom and

left-right arrangement can overlap. The left or right units can be further subdivided into top and bottom components (别 , 海). Similarly, the top or the bottom units can have left and right components (骂 , 想). Although each major component (left or right, top or bottom) may contain unequal numbers of subordinate units or stroke-patterns, they usually occupy roughly equal areas in standard written format (e.g., KAI format).

2) Two functionally distinct orthographic constituents of ZI: *the lexical radical and the non-radical components*

The present study is primarily concerned with the 90% or more of Chinese words that contain more than one unit (i.e., ZI). As seen in Figure 2 , ZI are formed by two major components. One is usually a word as a whole (sometimes called the stem). Another component, indicating the semantic category of the word, is an essential component sometimes called the "*signifier*", "*radical*" or "*semantic radical*" in the English literature. From KAN XI (A.D.776), the classical Chinese dictionary, to XIN HUA (1979), the contemporary Chinese dictionary, words are organized according to their radicals (other methods such as the four corner stroke code are seldom used in mainland China). The radicals are called 偏/[bian]1, 旁/[ban]2, 部/[bu]4, 首/[shou]3 in Chinese, meaning SIDE, COMPONENT and PART, HEAD, indicating their spatial positions. There are 189 radicals used as indices in the XIN HUA dictionary (1979), reduced from the traditional 214 in KAN-XI, including 8 single-stroke radicals for words consisting of only one stroke-pattern (WEN). These indices are ordered according to the number of strokes they have. Words are usually listed under just one index, with few exceptions. Although the set of indices listed in the dictionary is not conventionally taught in school, they should be familiar to most Chinese readers through using a dictionary. (The radical-index dictionary is the most commonly used one in mainland China). Much as the initial letter in an English word at the beginning of a sentence is written in uppercase, so components used as indices can

can also have a special format (called the radical format in this study). The radical format can differ from the non-radical format (i.e., the format for the same components when used as words in isolation or as the non-radical component of a word) as much as "A" and "a" or as little as "C" and "c". For instance, the component 刀 (also a word in isolation meaning KNIFE/[dao]1) is written as 刀 (word format) in isolation or as a non-radical component (as in MURMUR/[dao]2). However, it is written as 刂 when it is used as the index of the words 剥 PEEL/[pao]1, 割 CUT/[ke]1, 刘 SURNAME/[liu]2. Note that the position of this component 刂, when used as a dictionary index, is also fixed on the right side of the different words in which it occurs. Each radical used as a dictionary index has its own fixed position, which is called the "radical position" in this study. For different radicals, the radical position can be the left (e.g., 氵 in 江 河 泪 海) right (e.g., 刂 in 剥 割 刘 刚) top (e.g., 宀 in 宝 家 守 宋) or bottom (e.g., ...in 热, 烈, 杰, 然). The radical position can also be found in the dictionary (*XIN HUA, 1979*). For instance, the radical position for 彳 is on the left, since under 彳, it is on the left side of most words in which it occurs. Although in some cases, a component may be a radical in one word and a non-radical component in another, the same component used as a radical or as a non-radical component always occurs in a different position (and sometimes a different format) depending on its radical or non-radical status. For instance, 马 is listed as an index in the dictionary. It appears as a radical in 驴 [lu]2 DONKEY, but as a non-radical component in 妈'[ma]1 MOTHER, since the radical position of 马 should be on the left side.

To distinguish from the general (and somewhat unsatisfactory) term "radical", the term "lexical radical" is proposed here in the present study. "Lexical radical" refers to any of the 189 indices listed in the Chinese dictionary *when these components occur in the radical position (and radical format) in a word*. It is important to emphasize that lexical radicals can only be defined in word context,

not in isolation. Moreover, the positional regularity of a lexical radical (its radical position in different words) is generally an all or none rule, in contrast to single letter positional frequencies in English, which are a matter of more or less. In Chinese single-character words, orthographic regularity is in fact determined by the positional regularity of the lexical radicals. If a left radical (彳) is placed on the right side of a word, it yields an illegal word (i.e., a nonword) in Chinese (e.g., 彳亍).

For all ZI (i.e., words with more than one unit), the "*non-radical component*" is defined simply as all the remainder of the word, excluding the lexical radical. Thus the non-radical component may include more than one orthographic unit (see Figure 2.1). The distinction between the lexical radical and the non-radical component, as the names indicate, is thus determined by the structure and position of the lexical radical.

Further classifications can be made according to the lexical status of both lexical radical and non-radical components. Despite the fact that the orthographic units of Chinese words (ZI), can themselves be words in isolation, and have full lexical status, some orthographic units exist whose sound and sometimes meaning are no longer available in contemporary Chinese, due to language change and variation. These orthographic units are merely graphic symbols in modern Chinese, e.g., 金, 爻. These orthographic units may be called graphic-components in contrast to word-components. Hence, in the present study, lexical radicals are further sub-classified into word-radicals and graphic-radicals, and the non-radical components are further sub-classified into word-components and graphic-components, respectively.

The lexical radical and non-radical components are not only distinctive orthographic bodies but are also used to represent different linguistic information, which can be defined by rules, as follows.

2) Three types of rules in Chinese

a) *The traditional (descriptive) rules of word-formation*

Traditionally, there are six rules of word-formation called 六书 [liu]4[shu]1. These six rules are considered to account for the construction of all Chinese words. Accordingly, Chinese words have been classified into six traditional categories, namely, 象形 *Imitative* [xang]4 [xing]2, 指事 *Indicative*, [zhi]3 [shi]4, 会义 *Ideocompound* [hui]4 [yi]4, 形声 *Ideophonetic* [xing]2 [shen]1, 假借 [jian]3 [jie]4 *Loan* and 转注 [zhuan]3 [zhu]4 *Transformation*. In the English literature (Martin, 1972; Henderson, 1982), these six categories are sometimes reduced to five, namely, *pictography, ideography, compound ideographs, phonetic compounds and loans*, with the exclusion of the *transformation* category.

In fact both the (*semantic*) *transformation* category and the (*phonetic*) *loan* categories are methods of *using* words, rather than rules of composition (Hu & Huang, 1990). The use of *Phonetic loans* applies to the situation where words exist in speech but have (as yet) no corresponding written form. A word with the same sound (i.e., a homophone), is then loaned as the written format for the existing word in speech. Phonetic loans are thus homophonic words. The *semantic transformation* category refers to words with the same lexical radical, (hence members of the same semantic field). For instance, 耄、耆、耆、耆、耆、耆, all have the lexical radical OLD/ [lao]老. All these words mean OLD PEOPLE. These words are called 转注 ZHUAN-ZHU (*transformations*), and are to some extent synonyms (Hu & Huang, 1990).

The formation of all Chinese words, including both WEN and ZI, can thus be described in terms of four rules. They are the rules of *imitative, indicative, ideocompound and ideophonetic* formation. The most authoritative book (shuo-WEN-jian-ZI), used only these four rules to describe and analyse the structure,

meaning and sound of WEN and ZI. In that book, Xu-Shen divided words into the two classes WEN and ZI. He then described the meaning and sound of WEN in terms of the *imitative and indicative* rules, whereas he analysed the meaning and sound of ZI in terms of the ideocompound and ideophonetic rules.

The *imitative* rule refers to the archaic way of forming written symbols by a schematic 'picture' of the real objects. For instance, 日 SUN, 水 WATER, 雨

RAIN, 山 MOUNTAIN. The *indicative* rule refers to the formation of written symbols which can *abstractly* indicate the meaning, such as 上 (dot up a line) UP [shang]4, 下 (dot down a line) DOWN [xia]4. This is perhaps similar to the construction of (western) traffic signs ↑ . The *imitative and indicative* rules are in practice difficult to distinguish from each other. Both of them may be called *ideographic* rules, which are responsible for the formation of WEN (single-unit Chinese words) from individual strokes. In contrast, the composition of ZI (the 90% or more of Chinese words that contain more than one units) was based on the two rules known as *ideocompound and ideophonetic*.

The ideocompound rule refers to a culture-specific logic whereby the aggregation of the meaning of two components produces a new word. This is not unlike compounding in English, e.g., BLACKBOARD, ARMCHAIR. In Chinese, ideocompound words were formed by using two components which, in isolation, are also words (WEN). For instance,

好 = 女 + 子	(GOOD/[hao]3=WOMAN/[nu]3 + CHILD/[zi]3),
男 = 田 + 力	(MAN/[nan]2=LAND/[di]4 + POWER/[li]4),
明 = 日 + 月	(BRIGHT/[ming]2= SUN/[ri]4+ MOON/[yue]4),
安 = 宀 + 女	(SETTLED/[an]1=ROOM + WOMAN/[nu]3)

As indicated in above examples, both components, i.e., the constituents of ideocompound words, have a semantic but not a phonological association with the words. In ideocompound words, no orthographic unit was used to represent the pronunciation.

The ideophonetic rule refers to the composition of words by using one component (the lexical radical) to represent the word's approximate semantic category or semantic field, and another component, either simple or compound (i.e., the non-radical component) to represent its pronunciation. For instance,

村 = 木 + 寸	(VILLAGE/[cun]= WOOD/[mu] + INCH/[cun])
宾 = 宀 + 兵	(GUEST/[bin]=ROOM + SOLDIER/[bin])
抖 = 扌 + 斗	(SHAKE/[dou]= HAND+ STRUGGLE/[dou]),
肝 = 月 + 干	(LIVER/[gan]=FLESH + DO/[gan])

As can be seen one of these examples above, the non-radical component INCH/[cun] only has phonological but no semantic association with the word it forms (VILLAGE). By contrast, the lexical radical WOOD/[mu] has a (culture-specific) semantic association but no phonetic relationship to the word (VILLAGE/[cun]) that it forms.

b) Rules of lexical decomposition: Part-to-Whole Correspondences

Chinese orthography has been regarded as the "worst case" regarding the regularity of spelling-sound correspondences, with such regular orthographies as Serbo-Croatian, Japanese Kana, and Italian at the other end, and English orthography somewhere in the middle (e.g., Seidenberg, 1985a). However, what is characterised as regular or irregular depends very much on how the regularity is defined. In English (and other alphabetic languages), phonological regularity is usually defined in terms of Grapheme-Phoneme Correspondence (GPC) rules. The GPC rules in fact define the regularity of the pronunciations of *components* (graphemes). In English, orthographic units (letters or letter groups) can have more than one pronunciation in different words. For instance, letter group EA in English can be pronounced [i:] or [e] in different words (e.g., TEA or BREAK). In contrast to English, the pronunciation of a component in Chinese has only one phonetic value, which is never changed in different words. (For example, 台/[tai] in 抬/[tai], 胎/[tai], 苔/[tai]). Otherwise the phonetic value of the

component is null, e.g., Graphic-component. In this case, the pronunciation of the graphic-component via analogy with different words in which it occurs may be same or different (see later about phonological consistency). Therefore with respect to the regularity of the pronunciation of the *component*, Chinese should not be considered as the worst case of irregularity. On the contrary, it is arguably a highly regular orthography.

However, GPC rules mapping *multiple* orthographic components of a word onto *multiple* phonetic components (as in alphabetic scripts) do not apply to the pronunciation of a Chinese written word. Most Chinese words have more than one component. They differ from alphabetic scripts in that the pronunciation in Chinese specified by the pronunciation of a simple component, as well as the pronunciation of the word as a whole, is always a single syllable in Chinese. Accordingly pronunciations of a whole word, derived componentially (i.e., via lexical decomposition), cannot be achieved by assembling the phonetic value of each component, but only by derivation from just *one* component. This component is conventionally called the phonetic component (e.g., *Martin, 1972; Wieger, 1965; Seidenberg, 1985a, Butterworth and Yin, 1991; Flores d'Arcais, 1992*).

The so-called phonetic component was defined by the ideophonetic rule of word-formation, described above. The "phonetic" component is thus constituted by the non-radical component. In contrast to English orthography, which specifies (or encodes) both semantic and phonological information simultaneously across the whole letter array (*Chomsky, 1968*), in Chinese semantic and phonological information are encoded in separate orthographic components. (The following five points apply *only* to compound words or ZI).

1) When any component of a compound word specifies pronunciation (as in all ideophonetic compounds), this component is the *non-radical component* as a

whole. That is, the non-radical component, as a unitary orthographic constituent, encodes or specifies the syllabic pronunciation of the whole word. Accordingly, stated as a simple rule, *the phonetic component is the non-radical component*. This relationship will be called the rule of Part-to-Whole phonological correspondence (the 'phonological rule' for short).

2) In all compound words, one component always has a purely semantic function, and never encodes any phonetic value. This component is the *lexical radical*. (However, unlike the rule of Part-to-Whole phonological correspondence, the lexical radical cannot be said to specify the whole meaning of the word, but only certain features, or semantic associations, of its meaning, cf. the example of [cun]1/VILLAGE, above; the non-radical component is thus generally needed as well, to fully specify the word's meaning.) The important regularity, however, is that the lexical radical is a purely semantic component. Stated as a simple rule, *the (uniquely) semantic component is the lexical radical*. This will be called the rule of Part-to-Whole semantic correspondence (the 'semantic rule' for short).

3) It is important to emphasize that the (semantic or phonetic) function of a given orthographic component cannot be consistently determined from the identity of the component itself (i.e., from its internal 'content' or form). For instance, 平 [ping]2/PEACE is the phonetic component of 苹 [ping]2/APPLE, which is an ideophonic compound. However, 平 / [ping]2 is a semantic component (at least in terms of etymology) in 秤 [cheng]4/SCALES, which is an ideocompound, and it has no phonetic function in that word. The semantic or phonetic function of a component can be specified solely in terms of orthographic structure of the whole word.

4) The orthographic structure of a word—and hence the appropriate lexical decomposition needed for the component-based derivation of phonetic and

semantic transcoding of that word—is hierarchically organized. The first logical step in lexical decomposition is to identify, and locate, the lexical radical. Then, and only then, can the non-radical component be identified, namely as the rest of the word excluding the lexical radical. The two rules stated above, the Part-to-Whole phonological correspondence rule and the semantic correspondence rule, are thus rules pertaining to lexical decomposition, not rules of script-to-sound or script-to-meaning) transcoding. Lexical decomposition according to these rules is a necessary precondition for the application of any further component-based transcoding processes.

5) These two rules (the 'phonological rule' and the 'semantic rule') apply strictly to ideophonic compounds, which represent about 39% of all written words in the lexicon (*Zhou, 1978*). The 'semantic rule' applies also to ideocompounds (though the lexical radical is clearly not the only semantic component in ideocompounds). However, if the phonological rule is taken to define the regular orthographic structure of Chinese written words, it follows that all ideocompounds are *irregular*. A question of interest in this thesis is whether either, or both, of these rules have 'psychological reality'. Are they in fact implicitly applied by skilled readers of Chinese? If they are, there should be clearly identifiable behavioural costs of processing *irregular* words.

The phonetic and semantic rules for lexical decomposition, above, specify the regular structure of the phonetic and semantic components, *not* their etymologically-defined functions. However, it is the latter that underlies the conventional terminology of "semantic radical" and "phonetic radical" (*e.g., Wieger, 1965; Martin, 1972; Seidenberg, 1985; Butterworth and Yin, 1991; Flores d'Arcais, 1992*). Clarification of the regular structure of the semantic and phonetic components may be crucial for studies of the distinct functional roles of these components of Chinese words in word recognition, since selection of experimental stimuli must be based on a proper orthographic definition of the rule-based components. A serious problem that might undermine the recent important

studies, concerning the roles of the semantic and phonetic component, both with normal subjects (*Seidenberg, 1985a; Flores d'Arcais, 1991*) and with neuropsychological patients (*Butterworth and Yin, 1991*) is that the structure of the so-called "semantic radical" and "phonetic radical" was seriously under-specified. The distinction between semantic and phonetic radical was based either on their etymology (*Butterworth & Yin, 1991*) or simply on relative position; that is, components on the left side of words were regarded as "semantic radicals" while a components on the right side were considered as "phonetic radicals" (*Flores d'Arcais, 1991*). The left-right classification for semantic and phonetic components of Chinese words is obviously oversimplified and incomplete. It cannot accommodate words with top-bottom or outside-inside design, nor those with right-radicals. As stated before, different lexical radicals may appear in any position in a word. The semantic and phonetic rules of lexical decomposition, defined above, specify that "the semantic component is the lexical radical of a word" and "the phonetic component is the non-radical component of a word". The lexical radical and non-radical component can only be distinguished on the basis of the structure and position of the lexical radical. These rules of lexical decomposition require no etymological knowledge and are therefore far less ambiguous, and less complex, than the distinction between the "semantic radical" and the "phonetic radical" based on the postulated etymological origin of these components.

c) Transcoding rules: Derivation and Analogy

Pronunciation of written Chinese words by means of lexical decomposition thus involves at least three steps. (i) finding the lexical radical; (ii) segmenting the non-radical component from the lexical radical; and (iii) transcoding the non-radical component.

The possible transcoding procedures for pronouncing Chinese words (i.e., the third step) depend on the nature of the phonological relationship between the non-radical component and the words containing it. In Chinese, there are two types of relationship, which may be classified as congruency and consistency.

Phonological congruency refers to the congruency of pronunciation of a word with the pronunciation of its own non-radical component in isolation.

Phonological consistency refers to the consistency of the pronunciation among different words with the same component (as to English). In Chinese, this component must be the rule-defined phonetic component, (i.e., the non-radical component).

Consistency of pronunciation is usually incongruent with the pronunciation of the common non-radical component in isolation among different words. For instance, 怡 JOY and 贻 LEAVE BEHIND have a consistent pronunciation [yi]², which is incongruent with the common pronunciation of the non-radical component 台 STAGE [tai]² in isolation. Furthermore, words with the same non-radical component can have more than one consistent pronunciation. For instance, 捡 PICK UP and 检 CHECK, 脸 FACE, 敛 WITHDRAW have the same non-radical component 佥, 捡 and 检 are pronounced [jian]³ but 脸 and 敛 are pronounced [lian]³. Hence, [lian] & [jian] are two pronunciations for words with the component 佥. Thus where there is phonological congruency between a word and its non-radical component in isolation, the pronunciation of the word as a whole is simply indicated by the pronunciation of its own non-radical component in isolation. In this case, pronunciation of whole words via the non-radical components is *derived*. However, if there is no phonological congruency between a word as a whole and its own non-radical component in isolation, or the non-radical component in isolation is unpronounceable (i.e., is a Graphic-component), pronunciation of the whole word may be achieved via *analogy* with other words containing the same non-radical component, since there is a high level of phonological consistency among words with the same non-

radical components, in Chinese orthography, as described above. These two possible procedures for pronouncing Chinese words by use of lexical decomposition are therefore called *Derivation* and *Analogy*. In contrast to the decomposition rule that specifies the Part-to-Whole correspondence, these two procedures should be considered as processing rules or transcoding rules, applicable after lexical decomposition. However, it is important to point out that both these Derivation and Analogy in Chinese are based on just one rule of lexical decomposition, the Part-to-Whole phonological correspondence rule, since they are based on the (Part-to-Whole) rule-defined phonetic component, i.e., the non-radical component.

Table 2.1 compares Chinese and English orthography with regard to those aspects of orthographic knowledge that are relevant to word recognition.

Table 2.1 Comparisons between Chinese and English orthography

	Chinese	English
Orthographic unit		
Etymological origin	Primitive words	Abstract symbol
Orthographic structure	Stroke-pattern	Letter
Unique orthographic unit	Lexical radical	None
Orthographic regularity		
	Position of lexical radical	Single letter positional frequencies
Pronounceability	Pronounceability of the non-radical component	Bigram and trigram freq. Boss (?)
Part-to-Whole Correspondence		
Phonology	Non-radical component	Body (?)
Semantics	Lexical radicals	Morpheme (?)
Procedures for lexical decomposition		
	Derivation from one component	Assembly of many components
	Analogy from other words	Analogy from other words

From these comparisons in Table 2.1 between Chinese and English orthography, it appears that Chinese and English orthography are not as radically different as they are often thought to be. This may offer some encouragement for a future, universal theoretical framework for word recognition, which, however, will need to accommodate also orthography-specific and language-specific aspects. The limited available evidence from studying non-alphabetic languages (notably Japanese, *Paradis, Hagiwara, and Hildebrandt, 1985*) seems to suggest that current reading theories, established through the study of English orthography, are a first step in the right direction.

The present thesis is concerned with the recognition of Chinese words, in particular the 90% and more of Chinese words with more than one unit (ZI). So far, we have considered orthographic knowledge basically *a priori*. However, the psychological issue addressed in the present thesis is: How far and in what ways does such knowledge exert an influence upon reading Chinese. To answer this issue, in the following experimental studies with Chinese words, different aspects of orthographic knowledge will be examined in turn. The next chapter summarises the main theoretical issues and introduces the methodological background of the present study.

Chapter 3

Theoretical issues and Methodological Concerns

Theoretical issues

For a general overview, the issues to be addressed in the present study can be listed as follows:

- 1.) *What, if anything, is the orthographic unit of Chinese words, corresponding to the letter in alphabetic writing systems?*
- 2.) *Is there any interactive processing between the recognition of orthographic units and the recognition of words in Chinese? Do the word superiority effects also exist in the recognition of Chinese words?*
- 3.) *Does the pseudoword advantage over nonwords also exist in reading Chinese words?*
- 4.) *What is the role of semantic mediation in the phonological processing of Chinese words? Is there a direct route from orthography to phonology in reading Chinese?*
- 5.) *What is the role of the phonological recoding in the semantic processing of Chinese words?*
- 6.) *Can the two distinctive orthographic components—the lexical radicals and non-radical components—be recognized by native Chinese readers in accordance with the rules of Part-to-Whole Correspondence proposed in this study?*
- 7.) *How does the phonological congruency and consistency of words affect the phonological processing of Chinese words?*
- 8.) *Can word recognition in Chinese be adequately described by current word recognition models for English (and other alphabetic orthographies)?*
- 9.) *How are the orthographic, phonological and semantic recodings of written words integrated in reading Chinese?*
- 10.) *What is the relationship between reading efficiency and reading skill in Chinese?*

Methodological Concerns

1) Background of experiments

a) Simultaneous Same-Different comparison paradigm

Instead of such paradigms as lexical decision and naming, which are more commonly used in the study of word recognition, simultaneous Same-Different tasks (with different criteria) were used in the experiments of the present study. The basic experimental task is simply to judge whether pairs of Chinese scripts (which may be words or pseudowords or nonwords) are the Same or Different in *visual form, in pronunciation, or in meaning*, in different experiments. Such comparisons within specified domains (visual, or phonological, or semantic) may tap the specific component processes involved in reading words more directly or more selectively than do other more familiar tasks (e.g., naming). In principle, comparison in one domain (e.g., visual form), can be completed without necessarily involving processing from other domains, such as meaning or pronunciation. The lexical decision task (deciding whether a target is a word or not) can involve visual familiarity and/or semantic and phonological processing (*Besner & McCann, 1987*). The use of naming tasks to investigate phonological processing in Chinese can be problematic; they demand phonological *output* as well as recoding. The naming accuracy and latency of Chinese words may be influenced by subjects naming the word in her/his native dialect and then translating that form into the standard pronunciation of Chinese. Moreover, due to certain culture-specific factors, it is very difficult to ask native Chinese speakers to name pseudowords even when those pseudowords are described (by the experimenter) as 'unknown' words. However, in silent reading, it is not uncommon that native readers of Chinese would simply guess the pronunciation of unknown words. The phonological comparison paradigm attempts to

circumvent these problems by requiring phonological recoding but not requiring articulation and overt phonetic output. More importantly, by adopting the same comparison paradigm, the orthographic, semantic and phonological processing of Chinese words can be examined coherently and comparatively; comparison paradigms enable the use not only of the same subjects but also of the same stimuli in the study of semantic and phonological processing.

The very first experiment on the visual analysis of Chinese words and the exploratory experiment with semantic and phonological comparisons were run on a tachistoscope. All the subsequent main experiments were carried out with computer-presentation. Computer programmes were designed to control the entire set of experiments and data management therefrom. This made it possible to use twice as many stimuli (360) in about a third of the time (compared with the original tachistoscopic version). It also made it possible to use Chinese words in their standard printed format. Moreover, pseudowords and illegal nonwords can be constructed in the same printed format as words with a programme designed to be connected with a Chinese word processor.

b) Subjects who participated in this study

All subjects were skilled native speakers and readers of Chinese. Most of them were native speakers of northern dialects (i.e., the common language) and all of them can speak the common language, at least for oral reading and formal communication. All took their first degree in China and were reading for doctorates at the University of Oxford. No subject was reading psychology.

c) Stimuli and language statistics employed in this study

The stimuli used in all experiments were in simplified Chinese scripts. Apart from words, pseudowords and nonwords were also constructed and used as the stimuli. In all computerised experiments, words, pseudowords and nonwords

were presented in the standard printed KAI format. Words were selected from the Xin-Hua dictionary, published by the Shan-wu press in Shanghai (1979). Pseudowords and nonwords were constructed by changing the 16 x 24 dot pattern in the lexicon of ET (1981), a Chinese word processing package. Moreover, the visual complexity of the Chinese scripts were carefully balanced in terms of orthographic units and the positions of the unique orthographic units (the lexical radicals).

The word frequencies of all word stimuli in the present study were obtained from the most up to date and authoritative word frequency count, published by Beijing university (1988). More importantly, the present study also included linguistic statistics on the various frequencies of subcomponents such as radical frequency index, congruency frequency index and consistency frequency index in the dictionary (*Xin-Hua*, 1979). Each measurement will be described in detail in the individual experiments.

2) Statistical analyses

The terminology and statistical methods described below were used consistently in all experiments reported in this thesis. They will not be described for the individual experiments, unless any changes were made.

a) Variables used for the assessment:

Both the reaction time (RT) of correct responses and error rates were recorded. Error rates were usually very low, and in these cases, statistical analyses (especially the overall analyses) were performed on RTs only. In some further analyses, however, both RT and error rates were used. RTs and error rates were calculated both by subjects and by-items as follows:

Mean RT and error rates (by subjects): these refer to the mean value of reaction times and error rates of all stimuli in each condition for individual subjects.

Mean RT and error rates (By items): these refer to the mean value of reaction times and error rates of all subjects in responding to individual stimuli (items) in each condition.

b) Speed-accuracy trade offs

Before further statistical analyses, simple correlations and regressions of RTs and corresponding error rates for each subject (by-subject data) and each item (by-items data) were always conducted. The results of these regression tests are only reported when there is evidence of a significant speed-accuracy trade off. Thus RTs and error rates were always treated as two independent variables for assessment, unless otherwise specified.

c) Overall analyses

Analysis of variance (ANOVA) and analysis of variance and co-variance (ACOVA) were used (as overall analyses) to examine the general experimental effects and their interactions. ANOVA and ACOVA were used in within-subject designs with repeated measures.

d) Further analyses

Further analyses involved the examination of the different RTs or error rates between specific conditions. For multiple comparisons, the Newman-Keuls Q test was usually used, while for paired comparisons, the paired t test was used. For analysis of error rates, only if significant Skewness or Kurtosis was found were corresponding distribution free tests (such as the Friedman and the Wilcoxon signed rank test) used for multiple and paired comparisons, respectively. Other statistical analyses, such as simple and multiple regression tests, were also used for the purposes described in individual experiments. Statistical analyses were conducted by using the BMDP programme 2V for both ANOVA and ANCOVA

on the VAX terminal and the Statview programme on the Macintosh for other analyses.

Chapter 4

The visual analysis of Chinese characters

Word recognition in reading involves analysis of the visual, semantic and phonological forms of the written words. The studies reported in this chapter focus on the visual processing of Chinese script. In the introduction to Chinese orthography, the classification of the orthographic units of Chinese words was developed on the basis of relevant etymological knowledge. The orthographic units of Chinese words are defined in the present study as *the smallest stroke-patterns that can recur consistently and independently as constituents in different words, such that further decomposition of such a stroke-pattern will yield single strokes*. Unlike letters in English, the orthographic units of Chinese words, i.e., these smallest stroke-patterns, are not conventionally taught in the acquisition of literacy. Here the relevant psychological issue is: whether the orthographic units defined in this study can be implicitly recognized and used by native readers of Chinese. To address this issue, three experiments on visual comparisons of Chinese characters were conducted.

Paradigm and Rationale

The paradigm used in the three experiments reported in this chapter is that of simultaneous Same-Different judgments. The simultaneous 'Same'-'Different' (S-D) task was widely used particularly in the 1970s, in many studies of the visual analysis of English words (*e.g., Eichelman 1970; Egeth and Blecker, 1971; Henderson, 1974; Barron and Pittenger, 1974*). However, the cognitive processes underlying the S-D task are not yet entirely clear. The earlier theories such as Bamber's "fast 'Same' identitor" and Krueger's "noise-operator" (on 'Different' pairs) (*Bamber, 1969; Krueger, 1975; 1978*) were developed to

accommodate in particular the so-called fast 'Same' phenomenon in the S-D task. Given the extent of parallel processing in other domains, it may not be implausible to consider also that perceptual matching involves matching processes in parallel on the components of stimuli to be compared. Once a mismatching component is detected, the whole matching process can be terminated. If, however, a mismatching component is not found, the whole matching processes may repeat until it has been completed for all constituent components. In this case, one would expect that speed and/or accuracy of judgments on pairs of stimuli with no mismatching components (i.e., the 'Same' pairs), would be affected by the total number of constituent components, i.e., by visual complexity. Judgments of the 'Different' pairs may be affected by the number of mismatching components, termed Unit-Dissimilarity in the present study. Such parallel processing for perceptual matching would predict primarily a Fast-'Different' phenomenon rather than fast-'Same' judgments. However, it may also be able to accommodate the Fast-'Same' phenomenon, since decision processes in the comparison task may also be affected by verification or re-check procedures, either concurrently with the matching process or after it is completed. The latter may be particularly important when the paired stimuli to be compared differ along more than one dimension. In multidimensional 'Same'-'Different' (S-D) tasks, the stimuli to be compared differ in more than one dimension, e.g., in both colour and form, while the task itself may only require the comparison along one of these stimulus dimensions. In this case, once a mismatching component in one dimension, e.g., colour, is found, 'Different' judgments cannot be made until subjects verify that the detected difference is in the dimension required by the task instruction. This re-check on the 'Different' pairs may result in slower 'Different' (or Fast 'Same') performance, as suggested by Krueger's theory of perceptual matching (1978). Fast 'Same' results are indeed typically reported in multidimensional S-D tasks (Nickerson, 1967; Bamber, 1969, Miller and Bauer, 1981). In the present study, the stimuli to be compared only differ in visual form. If visual comparison is based on the detection of mismatching

features, re-checking is more likely to occur on the 'Same' pairs since no mismatching can be found. Therefore, Fast-'Different' instead of Fast-'Same' performance should be expected.

The rationale for using a S-D task in the present study is as follows: if the orthographic units defined in this study do indeed function as the constituent units of Chinese words, the speed and/or accuracy of S-D decisions should be affected by the number of constituent units (unit-complexity) of the 'Same' pairs and /or the number of mismatching units in the 'Different' pairs (unit-dissimilarity). If, however, the visual comparison of Chinese words is wholistically based on global features such as the overall visual configuration of the words, the unit-complexity and unit-dissimilarity should have no significant effect on speed and accuracy of the comparisons.

Experiment 1: Visual comparison of pairs of Chinese words

Introduction

An exploratory experiment was conducted, prior to the main study, in order to examine the effect of the number of mismatching units of the 'Different' pairs (Unit-dissimilarity) on the visual comparison of Chinese words. The effect of dissimilarity has been well-documented in experiments on visual comparison for both written English words, measured in terms of mismatching letters (*Barron and Pittenger 1974*), and also for non-verbal material (*Miller and Bauer, 1978*). In this experiment, unit-dissimilarity was measured in terms of the number of mismatching orthographic units, as previously defined. Therefore, the presence of unit-dissimilarity effects in this experiment would support the claim that the

orthographic units defined in this study are indeed those in the visual processing of Chinese words.

Method

Subjects

30 native readers and speakers of Chinese (21 males and 9 females), from mainland China, aged 25-35, were paid to participate in this experiment. They were all graduate students at the University of Oxford. All of them were right-handed with normal or corrected to normal vision.

Stimulus Conditions

Single character words, in simplified form, were used as stimuli. There were 60 trials, with equal numbers of 'Same' and 'Different' pairs. The 'Different' pairs were subdivided into four categories according to their unit-dissimilarity (U.D). The unit-dissimilarity was represented by the ratio of the number of mismatching units to the total number of constituent units of the two stimuli in the pairs. In the stimuli used in this experiment the number of units ranged from 1 to 5 units, with the majority having 2 or 3 units. Thus unit-dissimilarity in this experiment was classified into the five categories: 0 ('Same'), $1/3$, $1/2$, $2/3$, 1 (No similarity). An approximately equal number of pairs was constructed in each category of unit-dissimilarity (except 0 ('Same')). Stimuli consisting of any number of units other than 2 or 3 were classified into the closest unit-dissimilarity category according to the absolute value. For instance, 5-unit stimuli with one unit different between two characters in a pair ($1/5$) were classified into $1/3$ since $1/5$ (0.20) is closer to $1/3$ (0.33) than to any of the other categories.

Apparatus

The experiment was conducted on a three channel tachistoscope controlled by a BBC computer. Total viewing distance (including the viewing hood) was 580 mm. A pair of words, one above the other, each 10 mm x 10 mm and separated 5 mm from edge to edge, were written in black at the centre of the white display card (150 mm x 100 mm). When the subjects pressed a response button, the generator in the tachistoscope supplied a noise free signal to the timer. The reaction time (RT) and error for each trial and the mean RTs and error rates across trials for each subject were recorded in the BBC computer by a predetermined programme.

Procedure

The subjects' task was to judge whether the two stimuli on the target card were physically the 'Same' or 'Different' .

Before the experiment, each subject was presented with written instructions, which explained the procedures of the experiment and the requirements for making responses. Subjects were asked to respond by pressing one of the two buttons in front of them with the index finger of either the right or the left hand. They were asked to respond as accurately and quickly as possible. Half the subjects were instructed to press the left button (with the left hand) for the 'Same' judgments and the right button (with the right hand) for the 'Different'. The other half were reversed. Fixation on a red fixation point (+), whenever it appeared, was stressed by the experimenter throughout the experiment. The fixation point appeared midway between the position of the 2 test stimuli.

Each trial started with the appearance of a white display card with a red cross "+" as the fixation point at the centre, followed by a verbal signal 'ready'. Then the experimenter pressed the bar of the computer to present the stimuli. The target card appeared for 150 msec. and then the red fixation point (+) returned to replace

it. Once a response button had been pressed, the individual trial was finished and the experimenter placed the next target card into the tachistoscope and started the next trial according to the same procedure.

Results

The mean RT and error rates of the 30 subjects by the five categories of unit-dissimilarity are presented in Table 4.1 along with the results of the simple regression analysis. The "Same" pairs were regarded as zero for the calculation of unit-dissimilarity and were included in the correlation.

1) A unit-dissimilarity effect

There was an inverse correlation for both comparison speed and accuracy with the number of mismatching units in the stimuli (unit-dissimilarity). RT and error rates increased as unit-dissimilarity decreased across the five categories (0, 1/3, 1/2, 2/3, 2/2 (3/3)). This result is illustrated in Figure 4.1. The correlation was significant on both RT (*Correlation*=-0.22, *P*<0.001) and error rates (*Correlation*=-0.46, *P*<0.001) across the five categories, including the 'Same' pairs. After removing the U.D of zero ('Same' pairs), the correlation remained significant on error rates (*Correlation*=-0.37, *P*<0.001) but not on RT (*Correlation*=-0.15, *P*>0.05). Performance on the 'Same' pairs and the 'Different' pairs (including four categories of U.D.) was also compared on RT and error rates. 'Different' judgments were performed faster and more accurately than the 'Same' judgements. The difference between 'Same' and 'Different' judgments was highly significant in both RT and error rates (*RT*: $t_{(29)}=3.49$, *P*<0.01; *error rates*: $t_{(29)}=3.01$, *P*<0.01).

Table 4.1 Mean reaction times (RT msec.) and error rates (%) of 30 subjects by unit-dissimilarity in Experiment 1, with the results of the correlation analysis

	Unit-dissimilarity†					Correlation analysis	
	0	1/3	1/2	2/3	1	Correlation	P values
RT (msec.)							
Mean	863.4	808.7	762.7	730.5	669.9	-0.22	0.00
(SD)	(147.1)	(176.3)	(163.5)	(125.3)	(170.6)		
Error rate (%)							
Mean	37.3	38.6	26.7	18.9	7.7	-0.46	0.00
(SD)	(17.0)	(20.0)	(13.0)	(21.0)	(17.0)		

† 0=Same, 1=totally different, without any matching orthographic units

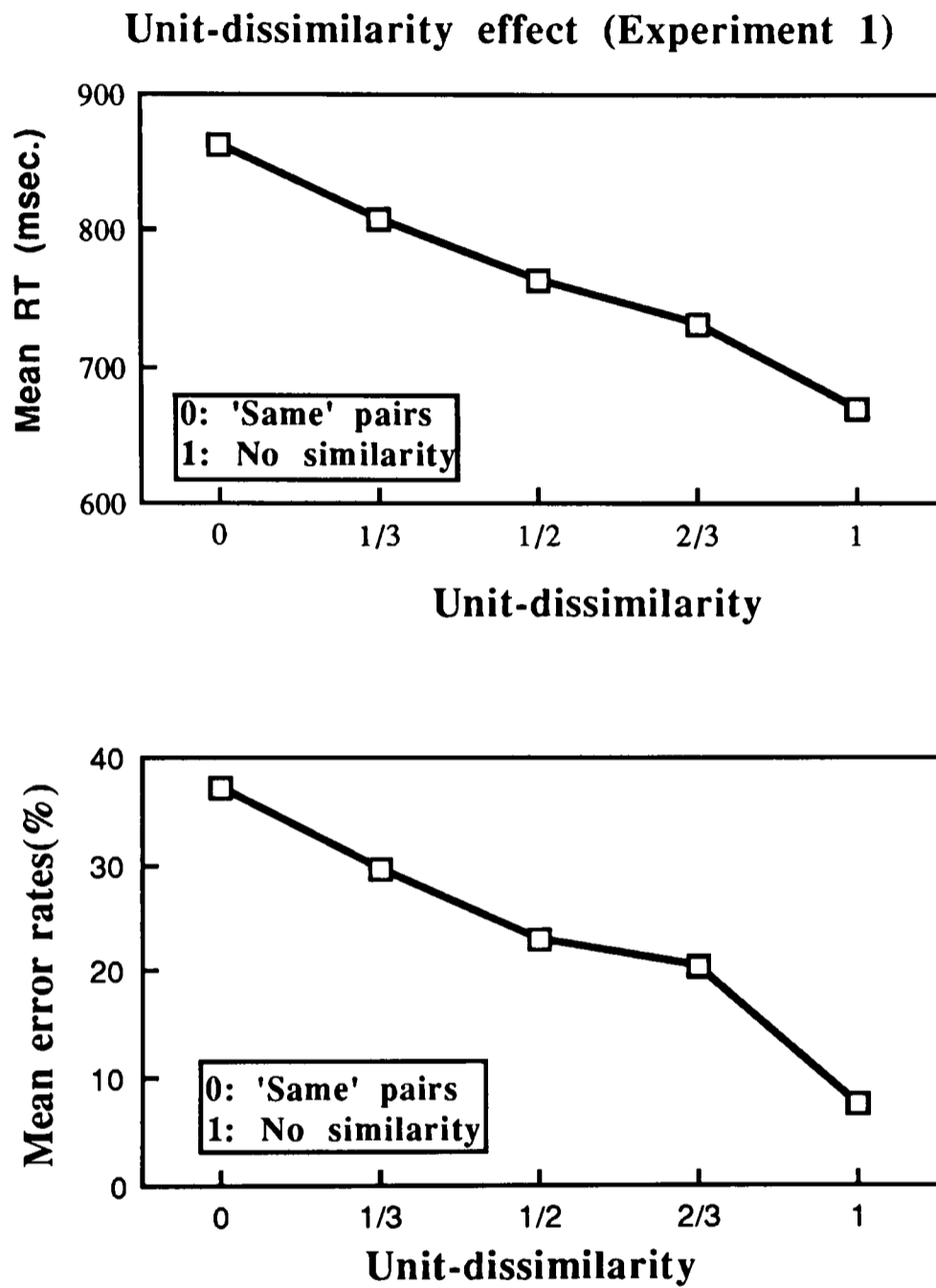


Figure.4.1 Mean RTs and error rates of 30 Chinese subjects by unit-dissimilarity of the word-stimuli

Discussion

This experiment shows that the visual comparison of Chinese words is affected by the number of mismatching units, the orthographic units as defined in this study. The demonstration of a unit-dissimilarity effect suggests that the orthographic units of Chinese can be (implicitly) recognized and used in the visual analysis of Chinese words. Fast 'Different' rather than Fast 'Same' performance was observed in this experiment, suggesting that Fast-'Same' may be a task-specific rather than irreversible phenomenon.

Experiment 2: Visual comparison of Chinese words, pseudowords and nonwords

Introduction

With better controlled stimuli and the task computerised, this experiment investigated further the nature and the functional roles of orthographic units in the visual analysis of Chinese characters. The aims of this experiment are twofold. First, the experiment is concerned to identify the orthographic units that are functional in reading Chinese words. In the previous experiment, it was demonstrated that comparisons of pairs of Chinese words seem to be based on orthographic units, as evidenced by the significant unit-dissimilarity effect. However, it is not clear whether the unit-dissimilarity effect demonstrated in the last experiment might not be confounded by the number of individual strokes. In some previous studies, the visual complexity of Chinese words was counted over the number of individual strokes (*e.g.*, Seidenberg, 1985a). Thus it seems necessary to compare the roles of individual strokes and of stroke-patterns, i.e., the orthographic units proposed in this study, in the visual processing of Chinese words. In this paradigm, the unit-dissimilarity effect has to be based on orthographic units, since pairs of words that differ by only a single stroke, *e.g.*, and are rare, especially for words with more than one unit. In order to

understand the functional roles of individual strokes and of orthographic units (stroke-patterns), both stroke-complexity and unit-complexity were examined in pairs of identical stimuli ('Same' pairs) in the current experiment. By manipulation of stimuli, stroke-complexity and unit-complexity were examined in an orthogonal design. In this experiment, stimuli containing either 2 or 3 units can have the same total number of strokes (8-11). Thus, the effect of unit-complexity, inferred from the difference between 2-unit and 3-unit stimuli with between 8 and 11 strokes, can be studied without the confounding of stroke-complexity. Similarly, stimuli with different numbers of strokes can have either 2 or 3 units. Hence the effect of stroke-complexity can be assessed in 2-unit and 3-unit stimuli separately and independent of unit-complexity.

Second, this experiment aimed to examine the word superiority effects in the visual analysis of Chinese characters. Word superiority effects are among the most ubiquitous effects, which any word recognition model has to account for. There are two types of word superiority effects. One is also sometimes called the word-letter superiority effect; that is, a letter is recognized better in a word context than in isolation, especially in conditions of backward visual masking (*Henderson, 1982*). This type of word superiority effect, first reported by Cattell (1886) more than a hundred years ago, was confirmed by Reicher (1969) and by Wheeler (1970) and has been studied by many other researchers (*e.g., Johnston & McClelland, 1973; see Henderson, 1982 for review*). Another type of word superiority effect, contrasting words and pseudowords, has also been widely obtained in various tasks, including the simultaneous "same-different" comparison task. Comparisons of word-pairs showing superior performance on RT to pseudowords in the simultaneous "same-different" comparison task, has been widely reported. The word superiority effect is evident in particular for "same" decisions but occasionally also for "different" RTs in English (*Barron & Pittenger, 1974; Baron, 1975; Chambers & Forster, 1975; Barron & Henderson, 1977; Taylor et al, 1977;*

Bruder, 1978; Carr et al, 1979). The major account of word superiority over pseudowords in English is usually considered to be the lexicality of words (*Henderson, 1982*). The lexicality of single characters in Chinese is sometimes thought to be "vague". Instead of "words", Chinese scripts are conventionally called "characters" in English, and are sometimes considered to be equated to "morphemes" in English (*Hung & Tzeng, 1981*). The demonstration of a word superiority effect with single characters in Chinese would support the view that Chinese characters should be equated to "words" rather than "morphemes", as discussed in Chapter 1 of this thesis. Moreover, study of the role of lexicality in word recognition may contribute to the recent debate about whether (or in what sense) "word" units exist in the functional framework of word recognition (*Seidenberg & McClelland, 1989; Besner, et al, 1990, Seidenberg & McClelland, 1990*). The mutual interaction between word constituents (such as letters) and words, which characterises many interactive models of word recognition, such as the parallel distributed processing model of *McClelland & Rumelhart (1981)*, provides one class of explanation for both types of word superiority effect. In this experiment, recognition of constituent orthographic units of Chinese words will be inferred from unit-complexity effects and/or unit-dissimilarity effects. If there is a word superiority effect over pseudowords and nonwords in Chinese, visual comparisons should be faster and/or more accurate on words than that on pseudowords and nonwords. Secondly, a word superiority effect affecting the recognition of orthographic units should manifest itself in the form of stronger unit-complexity and unit-dissimilarity effects in words than in pseudowords and nonwords. These two types of word superiority effects in Chinese, if obtained, may well suggest a functional parallelism between the orthographic units of Chinese and the letters in English orthography, in the context of current interactive models for word recognition.

Along with word superiority relative to pseudowords, a pseudoword advantage relative to illegal nonwords has also been widely obtained in English. The study of this pseudoword advantage has concentrated on the roles of orthographic structure in word recognition. In the visual analysis of English words, positional information is considered to be important for letter recognition (*Allport, 1977*). That is, a variable of importance in word or letter-string processing is the single letter positional frequency (SLPF) (or positional redundancy). Due to the constraints of English orthography, the single letter positional frequency is highly intercorrelated with sequential redundancy (bigram, trigram frequency) and pronounceability. Many studies attempted to find an independent role of these different factors without confounding each other (*Gibson, et al, 1970; Massaro et al, 1979, Henderson and Chard, 1980*). Although single letter positional frequency appears to be a main factor in the pseudoword advantage, it is not entirely clear whether it is uniquely causal or whether the effect can be mediated by other factors such as sequential redundancy.

The situation is rather different in Chinese. That is, in Chinese, words, pseudowords and nonwords can be formed by manipulation of the positions and the combinations of co-existing orthographic units. It is possible to construct pseudowords by substitution of one orthographic unit from a real word, leading to a meaningless (non-lexical) combination of the orthographic units, e.g., . However, the sole difference between the construction of illegal nonwords and pseudowords is that in a nonword the lexical radical is placed in an illegitimate position (such as three dots on the right side, e.g.,).

As summarized in Table 4.2, the position of the lexical radical constitutes the sole difference between pseudowords and nonwords. Hence, a pseudoword advantage over nonwords, if obtained, would suggest the independent role of the

positional regularity of lexical radicals without confounding with other orthographic attributes. Therefore, the functional role of positional information in orthographic units may be better understood in the study of Chinese orthography.

Table 4.2 Comparison of the orthographic attributes of words, pseudowords and nonwords in Chinese

	Words	Pseudowords	Nonwords
Positional regularity	+	+	-
Lexicality	+	-	-
Orthographic familiarity	+	-	-
Unit familiarity	+	+	+

Notes: Given that the pseudowords and nonwords were formed by substitution of one orthographic unit from the real words used in this experiment, the units used in words, pseudowords and nonwords were identical.

Method

Subjects

12 native Chinese (6 males and 6 females), aged 20 to 35 years old were paid to participate in the experiment. All subjects were graduates from mainland China, and were currently studying for a PhD at the University of Oxford. They were all right handed, and had normal or corrected to normal vision.

Stimulus-Conditions

The stimuli used in this experiment were 2-unit or 3-unit Chinese words, pseudowords and nonwords. The composite spatial arrangements of orthographic units of all stimuli was either top-bottom 守 or left-right 列. The two stimuli in a pair always had the same spatial design 表 or 削. The number of top-bottom and 杂 剑.

left-right stimuli was equal in different conditions. In each condition, these spatial designs were randomly ordered within a block.

There was a total of 360 trials, over 3 word-types in the experiment. There were nine stimulus-conditions for each type of word, as illustrated in Figure 4.2. They were nested, in the order of four variables, i.e., word type (word, pseudoword and nonword), unit-complexity (2-unit vs 3-unit), stroke-complexity ('Same' pairs) or unit-dissimilarity ('Different' pairs), and judgment ('Same' or 'Different') (Figure 4.2). Stroke-complexity was classified post hoc according to the number of strokes in each stimulus of a pair. This was only done in 'Same' pairs, since in 'Different' pairs the two stimuli almost always have a different number of strokes. Unit-dissimilarity, based on the number of mismatching units, of course could only be classified in 'Different' pairs. The number of 'Same' and 'Different' pairs in each condition was identical. Apart from the variable of stroke-complexity, the number of stimuli of each subtype was identical. The variables of word-type, unit-complexity and unit-dissimilarity were all presented in blocks, as indicated by a square in Figure 4.2. Each block had equal numbers of 'Same' and 'Different' pairs, and these were ordered randomly. The randomisation of stimuli within each block was performed by the computer programme, and was different for individual subjects. The presentation of the blocks was also counterbalanced among the subjects. Data from 'Same' pairs in each block were further divided for analysis into two groups, "simple" or "complex", according to stroke-complexity (Figure 4.2). Each variable is described in detail below:

Word-type

Words: Real words were selected from frequency bands that ranged from 100 to 5011 per hundred thousand (100,000) as based on the revised Frequency Count Dictionary published by Beijing Language University (1988). Words in different conditions were carefully matched for word frequency. Any two words that were selected to form a real word pair had no semantic or phonological relationship to

each other, and they cannot form a meaningful two-character CI (see Chapter 1 for two-character CI).

Pseudowords: Pseudowords were constructed by substitution of one unit from real words to form a non-existing combination of orthographic units. Lexical radicals of pseudowords were always in their legitimate positions (see the example in Figure 4.2).

Nonwords: Nonwords were constructed by placing the lexical radicals of the pseudowords in illegal positions. Therefore they were examples of illegal Chinese orthography, although the constituent orthographic units used in the nonwords were also selected from the real words used in this experiment.

Unit-complexity

All stimuli consisted of either 2-unit or 3-unit characters in the simplified version. Two words in a pair, for both 'Same' and 'Different' judgments, always have the same number of constituent units.

Stroke-complexity of 'Same' pairs

The number of constituent strokes of each word in the 'Same' pairs ranged from 5 to 11 strokes for 2-unit stimuli, and from 8 to 22 strokes for 3-unit stimuli. In the 2-unit stimuli, "simple" refers to stimuli containing 5-7 strokes, while "complex" refers to those with 8 to 11 strokes. For 3-unit stimuli, stimuli with 8-11 strokes were classified as "simple" and those with 12 to 22 strokes were classified as "complex". The numbers of stimulus-pairs with 8-11 strokes in 2-unit were 13 and in 3-unit were 22 (Figure 4.2).

Experiment 2 & 3

Task: Is the pair of scripts the 'Same' or 'Different' ?
(360 trials in 15 blocks)

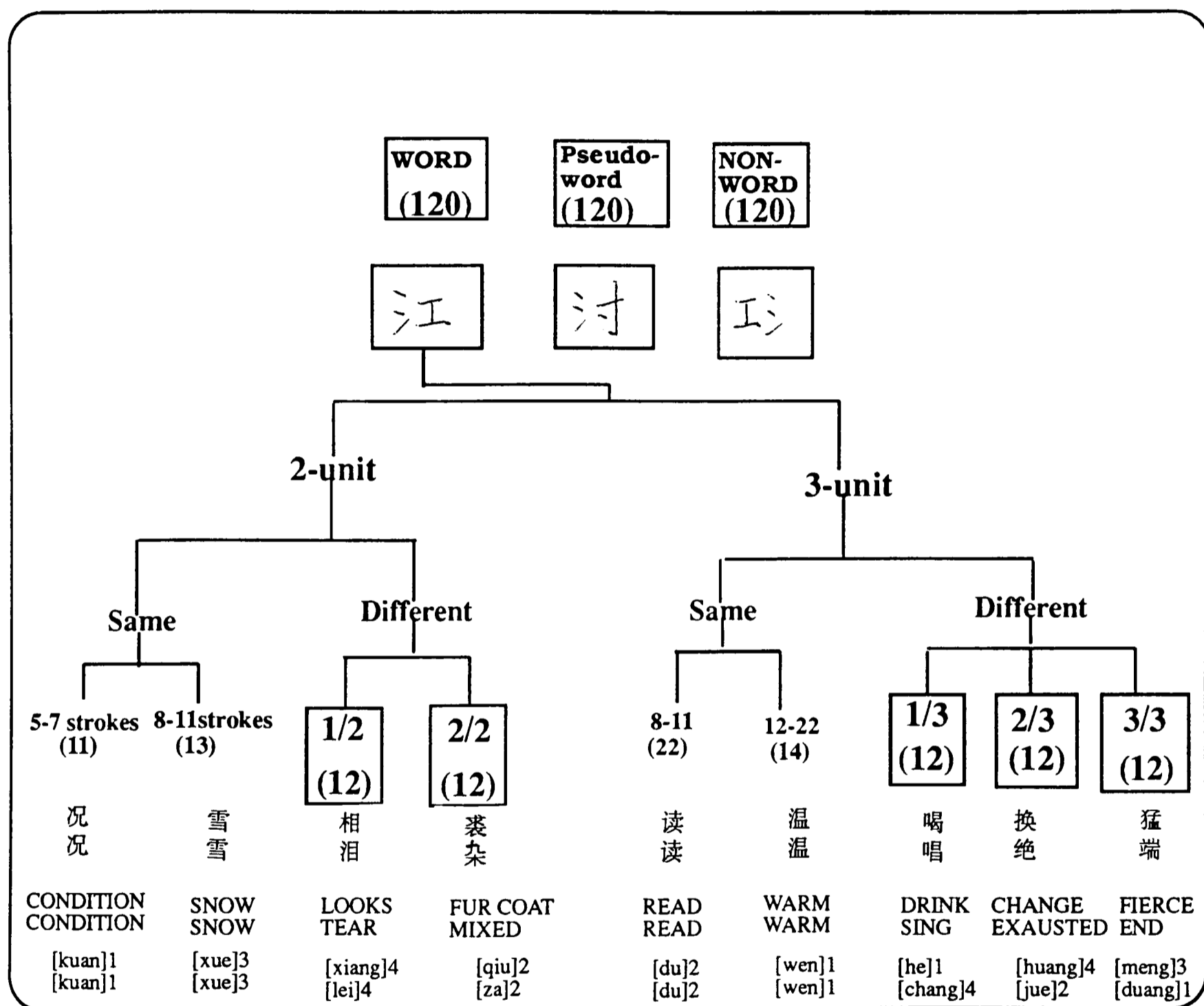


Figure 4.2 Stimulus conditions in Experiment 2 and 3

Unit-dissimilarity for 'Different' pairs

The two stimuli in a 'Different' pair always had the same number of constituent units. The unit-dissimilarity was represented by the ratio of the number of mismatching units over the constituent units of each stimulus. For instance, 1/2 represents a pair of 2-unit stimuli with one mismatching unit between them. In 2-unit pairs, unit-dissimilarity included 1/2 and 2/2, while in 3-unit pairs it included 1/3, 2/3, 3/3. The spatial position of the mismatching units in the 'Different' stimuli could be at the *upper right, right, left, or upper left*. The numbers of pairs with mismatching units at certain position (e.g., left, right) were matched in the different conditions. For each pair, the mismatching unit(s) were always at the same location in their constituent stimulus. For instance, for a pair of stimuli 喝 , 唱 the mismatching units 和 were both on the lower right position of their constituent words. The orthographic status of mismatching units, lexical radical 揭 喝 or non-radical component 削 劍 , was also balanced in different conditions. In the same category of unit-dissimilarity, stimuli with different mismatching units in terms of position and status were randomly presented within a block.

Apparatus

The experiment was carried out on a Packard Bell IBM PC-AT286 monograph computer. A programme was designed to control stimulus presentation, record response time and error rate, and calculate the mean value of both reaction times and error rates in each condition. This programme was also connected with a Chinese word processing package-ET to retrieve real words and construct the pseudowords and nonwords used in the present experiment. The pairs of stimuli were shown in printed KAI format in white on a black background. Stimuli were presented simultaneously, one above the other. Each

character was 1 cm x 1 cm, and the two characters in a pair were separated by 1.5 cm centre to centre. Each pair of stimuli covered a visual angle of circa 1° horizontally by 2° vertically from a viewing distance of 30 cm.

Procedure

The task was simply to judge whether the two members of a pair of stimuli were 'Same' or 'Different' in terms of their visual form.

Each trial began with a fixation cross, "+", 0.2 cm x 0.2 cm. It appeared in the centre of the screen. It was displayed for 0.5 seconds and then disappeared. After a further interval of 0.5 seconds, a pair of stimuli would appear in the positions immediately above and below the fixation point. The pair remained on the screen until the response key was pressed. Once the response key was pressed, the stimuli disappeared immediately from the screen. There was then an interval of 0.5 seconds before the fixation point, "+", re-appeared to start the next trial.

Before the experiment, subjects were asked to read carefully the instructions in Chinese on the screen. The written instructions explained task content and the procedures. The experiment was also further described where necessary. The subjects were told prior to the experiment that there may be some "unknown" or even nonsense characters among the stimuli. They were simply required to make a judgement concerning the physical identity of the stimuli, and nothing else. The "Z" and "/" keys, which are located at the left and right corner of the keyboard respectively, were used as the two response keys. Subjects were asked to respond as accurately and quickly as possible, using the index finger of left or right hand to press the left or right key. All subjects were asked to press the "/" key (right) for the 'Same' judgements and the 'Z' key (left) for the 'Different' judgements. The two response keys were also labelled correspondingly with 同 SAME and 异 DIFFERENT, in Chinese.

Prior to the main experiment, there were 30 pairs of words for a warm-up exercise. After this practice session, the mean reaction times and error rates for this session were shown on the screen. If the error rate was equal to or greater than 20%, more practice was given using different stimuli until an accuracy of at least 85% was reached.

During the formal experiment, there was a break after every 120 trials, accompanied by a piece of quiet electronic music. Subjects could rest as long as they needed. After the experiment was completed, subjects were asked to report the strategies they used in the experiment.

Results

1) Word-type effect on both judgments

Table 4.3a shows the mean correct RTs by judgments and word-type, along with the results of statistical analyses. The corresponding data for error rates are given in Table 4.3b. As indicated in Table 4.3a, RT increased markedly from words to pseudowords to nonwords, on both 'Same' and 'Different' judgements. On 'Same' judgements, the mean RTs were 756.8 ms, 846.2 ms and 941.2 ms respectively for words, pseudowords and nonwords; on the 'Different' judgments, they were 778.7 ms, 836.0 ms and 916.5 ms respectively. Note that the standard deviation also increased markedly from words to pseudowords to nonwords, particularly on 'Same' judgements (Table 4.3a). This word-type effect in both judgments is illustrated in Figure 4.3, for RT and error rates separately. The overall analysis by ANOVA (2 X 3) on RT, with main effects of judgment and word-type, revealed a highly significant word-type effect ($F_{(2,11)}=20.99$, $P<0.001$) but no judgment effect ($F_{(1,11)}=0.06$, $P>0.05$). Overall mean RTs for 'Same' and 'Different' judgments were 848 msec. and 844 msec. respectively. Accordingly, the word- and pseudoword superiority effects cannot be attributed to a response bias, as in some matching

Table 4.3a. Mean reaction times in msec. of 12 native Chinese subjects by judgments and word types in Experiment 2 with the results of the statistical analyses

Word-type	No. of stimuli	Same judgment		Different judgment		Overall mean
		Mean	SD	Mean	SD	
Word (W)	120	756.8	74.2	778.7	71.4	767.7
Pseudoword(PW)	120	846.2	115.6	836.0	72.6	841.1
Nonword (NW)	120	941.2	147.5	916.5	93.2	928.8
All types	360	848.1	112.4	843.7	79.1	845.9

<i>ANOVA (2 x 3)</i>	F value	P value
Main effects		
Judgment (2)	0.06	0.81
Word-type (3)	20.09	<0.001
Interaction		
Judgment x Word-type	1.91	0.17

<i>Newman-Keuls (Q value)</i>	Same	Different
W vs PW	6.1**	3.4*
W vs NW	7.0**	10.7**
PW vs NW	3.9*	3.7*

Table 4.3b. Mean error rates of 12 subjects by judgments and word-type in Experiment 2, with statistical analyses

Word-type	No. of stimuli	Same judgement		Different judgement		Overall Mean
		Mean	Mean	Mean	Mean	
Word (W)	120	7.6	9.8	8.7		
Pseudoword(PW)	120	11.2	14.0	12.6		
Nonword (NW)	120	10.4	21.5	15.9		
All types	360	9.7	15.1	12.4		

Statistical analysis		
<i>Newman-Keuls (Q values)</i>		
W vs PW	1.82	4.08*
W vs NW	1.70	6.43**
PW vs NW	0.61	3.97*

*, ** P<0.05 and P<0.01 respectively

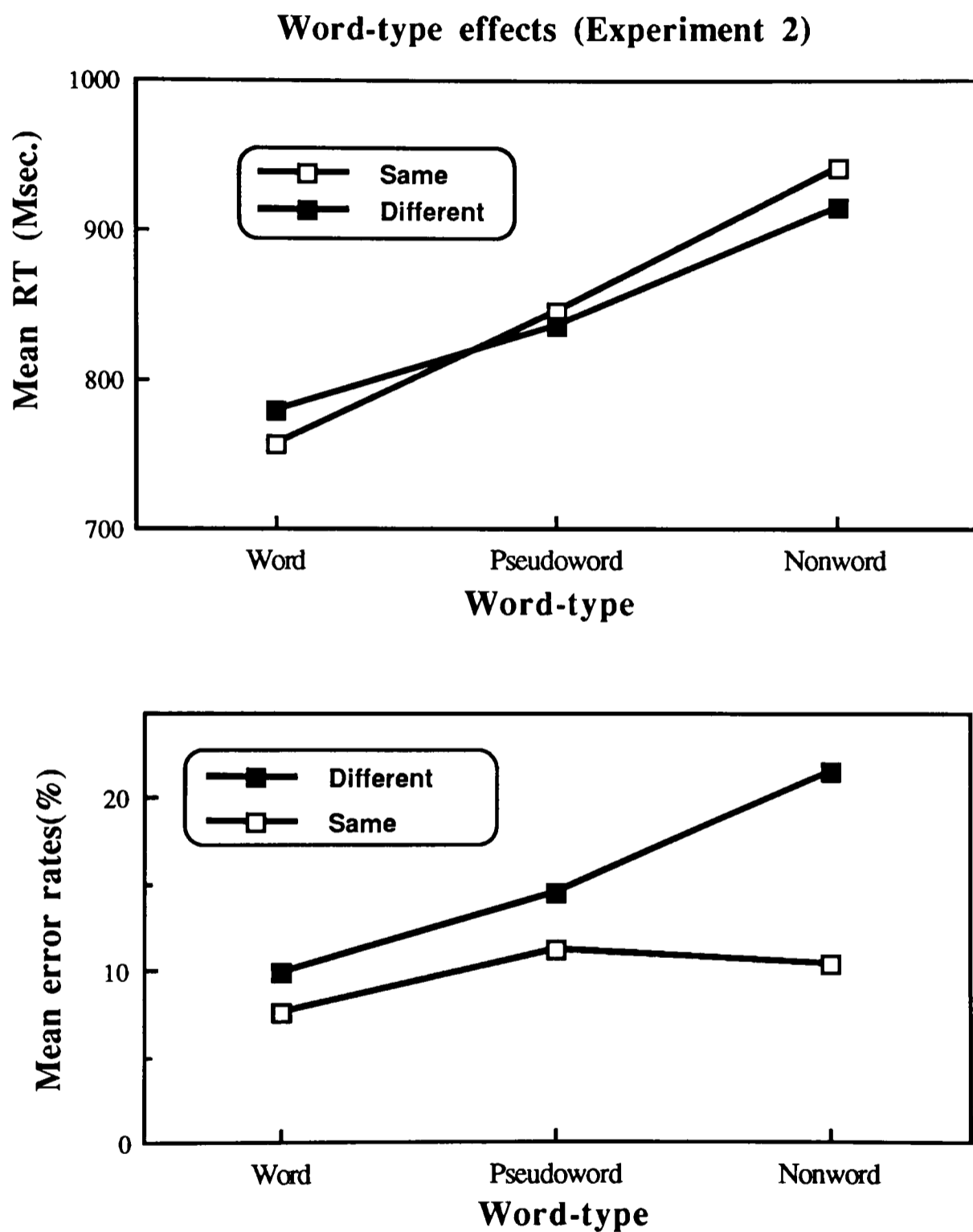


Figure 4.3 Mean RT or error rates (%) for word-type on both judgments

tasks with alphabetic stimuli (*cf Henderson, 1982*). The interaction of the two main effects was not significant ($F_{(3,22)}=1.91, P>0.05$).

The overall pattern on error rates (Figure 4.3) was consistent with that on RT, except that, on 'Same' judgements, the error rate for nonwords (10.4%) was not higher than that for pseudowords (11.2%). Further analyses of both RT and error rates were performed separately for the 'Same' and 'Different' judgments to examine further the word-type effect. In RTs, the differences between words and pseudowords (word superiority) and between pseudowords and nonwords (pseudoword advantage) were significant for both judgments, by Newman-Keuls Q test (Table 4.3a). In error rates, the word superiority and pseudoword advantage were also both significant, but only on the 'Different' judgments (Table 4.3b).

2) 'Same' judgments: unit-complexity and stroke-complexity effects

Tables 4.4a and 4.4b present the mean values of RT and error rates of the 12 native Chinese subjects respectively by stimulus-conditions, along with the statistical analyses. The 'Same' judgments included 7 stimulus-conditions with three variables, i.e, word-type (3), unit-complexity (2) and stroke-complexity (2). Data for 'Same' pairs were obtained, in separate blocks, for words, pseudowords and nonwords and for 2-unit stimuli and 3-unit stimuli of each word-type. From each block, by-subjects means for stimuli with different numbers of strokes (stroke-complexity) were further obtained. As shown in Table 4.4a, in each condition words were judged faster than pseudowords (word superiority effect) and pseudowords, in turn, were faster than nonwords (pseudoword advantage). In each word-type, performance on the 3-unit stimuli was slower than on 2-unit stimuli. This unit-complexity effect remained evident with control of the number of strokes (8-11) in 2-unit and 3-unit stimuli. The unit-complexity effect on RT was greatest for words, less so for pseudowords, and became much less evident for

Table 4.4a Mean reaction times (msec.) on 'Same' judgments by word-type, unit-complexity and stroke-complexity in Experiment 2, with the results of the statistical analyses

No. of Units	No. of Strokes	Word		Pseudoword		Nonword		All types	
		Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
2-unit	5-7	711.7	(66.5)	790.7	(109.0)	941.9	(220.6)	814.7	(132.0)
	8-11	720.7	(76.5)	838.5	(154.8)	939.3	(173.5)	832.8	(131.9)
	All	716.2	(71.5)	814.6	(131.9)	940.6	(197.0)	823.7	(131.9)
3-unit	8-11	808.2	(85.8)	883.4	(113.6)	953.8	(190.5)	881.8	(130.0)
	12-22	801.4	(85.8)	865.1	(118.4)	943.3	(171.6)	869.9	(125.3)
	All	804.8	(85.8)	874.2	(116.0)	948.5	(181.0)	875.8	(127.6)

ANOVA(3 x 2 x 2)

Main effects

Word-type (W)

F values

16.08

P values

<0.001

Stroke-complexity (S)

0.08

0.78

Unit-complexity (U)

11.99

<0.001

Interactions

W x U

2.18

0.13

S x U

2.09

0.18

W x S

0.50

0.61

W x S x U

0.79

0.46

Newman-keuls (Q values)

Word

Pseudoword

Nonword

Unit-complexity

All stimuli

7.21**

4.95*

0.28

8-11 strokes

3.93*

1.63

0.31

Stroke-complexity

2-unit

0.92

3.37

0.08

3-unit

0.77

1.45

0.63

*, ** P<0.05 and P<0.01 respectively

Table 4.4b Mean error rates (%) on 'Same' judgments by word type, unit-complexity and stroke-complexity in Experiment 2.

No. of units	No. of strokes	<u>Word</u> Mean	<u>Pseudoword</u> Mean	<u>Nonword</u> Mean	<u>All stimuli</u> Mean
2-unit	5-7	6.3	3.4	7.9	5.9
	8-11	6.4	4.5	3.7	4.9
	All	6.3	3.9	5.8	5.4
3-unit	8-11	6.7	4.0	4.0	4.9
	12-22	7.3	5.0	4.0	5.4
	All	7.0	4.5	4.0	5.1

'Same' judgments: Unit-complexity and stroke-complexity effects (Experiment 2)

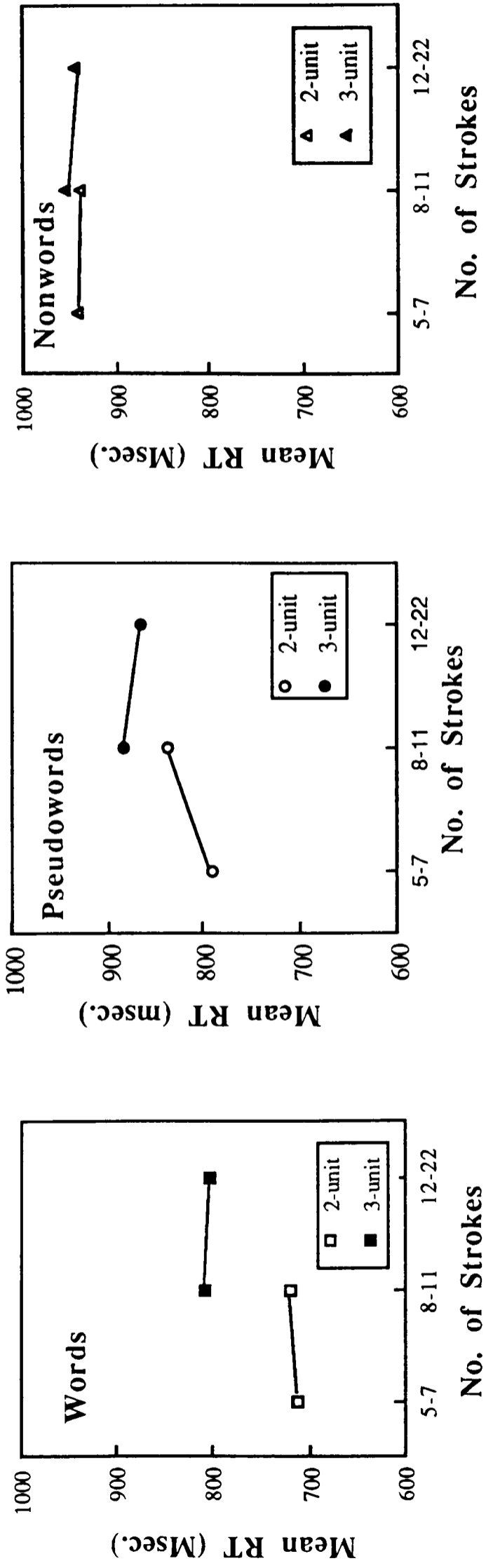


Figure. 4.4. Mean 'Same' RTs (msec.) by stroke-complexity and unit-complexity

nonwords. In contrast to the effect of unit-complexity on RT, there was no RT difference between different levels of stroke-complexity. An ANOVA ($3 \times 2 \times 2$) was performed on the RTs of the 'Same' pairs to examine the main effects of word-type, unit-complexity, stroke-complexity and their interactions. Both word-type ($F_{(2,11)}=16.08, P<0.001$) and unit-complexity effects were highly significant ($F_{(1,11)}=11.99, P<0.001$). By contrast, the stroke-complexity effect was not significant. No interactions among the three main effects, or between any two of them were significant (Table 4.4a).

Further analyses on both RT and error rates examined the unit-complexity and stroke-complexity effects separately in words, pseudowords and nonwords. The unit-complexity effect and stroke-complexity effect are illustrated in Figure 4.4 separately for words, pseudowords and nonwords. Overall, the difference in RT between the 2-unit and 3-unit stimuli (unit-complexity effect) was significant for both words and pseudowords, but not for nonwords (Table 4.4a). After controlling the number of strokes (8-11), the difference in RT between 2-unit and 3-unit (unit-complexity) was significant only for words, although the difference was also appreciable for pseudowords (Table 4.4a). The stroke-complexity effect was also examined separately in 2-unit and 3-unit stimuli, to avoid any possible confounding effect from unit-complexity. As confirmed by the results from the ANOVA, there was no significant stroke-complexity effect on RT in any of the three types of stimuli, words, pseudowords or nonwords, in either the 2-unit or 3-unit stimuli (Table 4.4a).

3) 'Different' judgments: Unit-dissimilarity effect

The variables examined in the 'Different' judgments were the effects of word-type, unit-complexity and unit-dissimilarity with 3, 2, and 2 levels respectively. The overall mean values of RT and error rate for the subjects are shown respectively in Tables 4.5a and 4.5b. The data are also illustrated in Figure 4.5.

Table 4.5a Mean reaction time (msec.) on **Different** judgments by word-type, unit-complexity and unit-dissimilarity in Experiment 2, with statistical analyses

No. of Unit	Degree of Dissimilarity	Word		Pseudowords		Nonwords		All types	
		Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
2-unit	1/2	795	(79)	886	(156)	930	(125)	870	(120)
	2/2	693	(82)	744	(109)	843	(93)	760	(95)
	All	744	(80)	815	(132)	886	(109)	815	(107)
3-unit	1/3	864	(90)	883	(123)	927	(73)	891	(95)
	2/3	809	(59)	860	(89)	930	(90)	866	(79)
	3/3	765	(81)	826	(125)	979	(84)	857	(97)
	All	813	(77)	856	(112)	945	(81)	871	(90)

<i>ANOVA(3 x 2 x 2)</i>		F value	P value
Main effects			
Word-type (W)		11.51	<0.001
Unit-dissimilarity (UD)		39.40	<0.001
Unit-complexity (UC)		11.39	<0.001
Interactions			
W x UD		4.20	<0.001
W x UC		0.26	0.78
UC x UD		4.96	0.05
W x UD x UC		1.14	0.34

<i>Newman-Keuls (Q value)</i>	<u>Words</u>	<u>Pseudowords</u>	<u>Nonwords</u>
Unit-complexity 2/2 vs 3/3	6.07*	6.26*	4.22
Unit-dissimilarity 1/2 vs 2/2	11.07**	3.30	2.74
1/3 vs 3/3	4.96*	3.55	1.83
1/3 vs 2/3	3.43	1.33	1.10
2/3 vs 3/3	3.48	2.44	1.10

*,** P<0.05 and P<0.01 respectively

Table 4.5b. Mean error rates (%) on **Different** judgments by word-type, unit-complexity and unit-dissimilarity in Experiment 2.

Unit	U.D	<u>Word</u> Mean	<u>Pseudoword</u> Mean	<u>Nonword</u> Mean	<u>All types</u> Mean
2-unit	1/2	4.1	7.6	10.4	7.4
	2/2	4.8	9.7	10.4	8.3
	All	4.4	8.6	10.4	7.8
3-unit	1/3	7.6	12.5	8.3	9.5
	2/3	9.0	14.6	3.8	9.1
	3/3	12.5	11.8	9.0	11.1
	All	9.7	13.0	7.0	9.9

'Different' judgments: Unit-dissimilarity effect (Experiment 2)

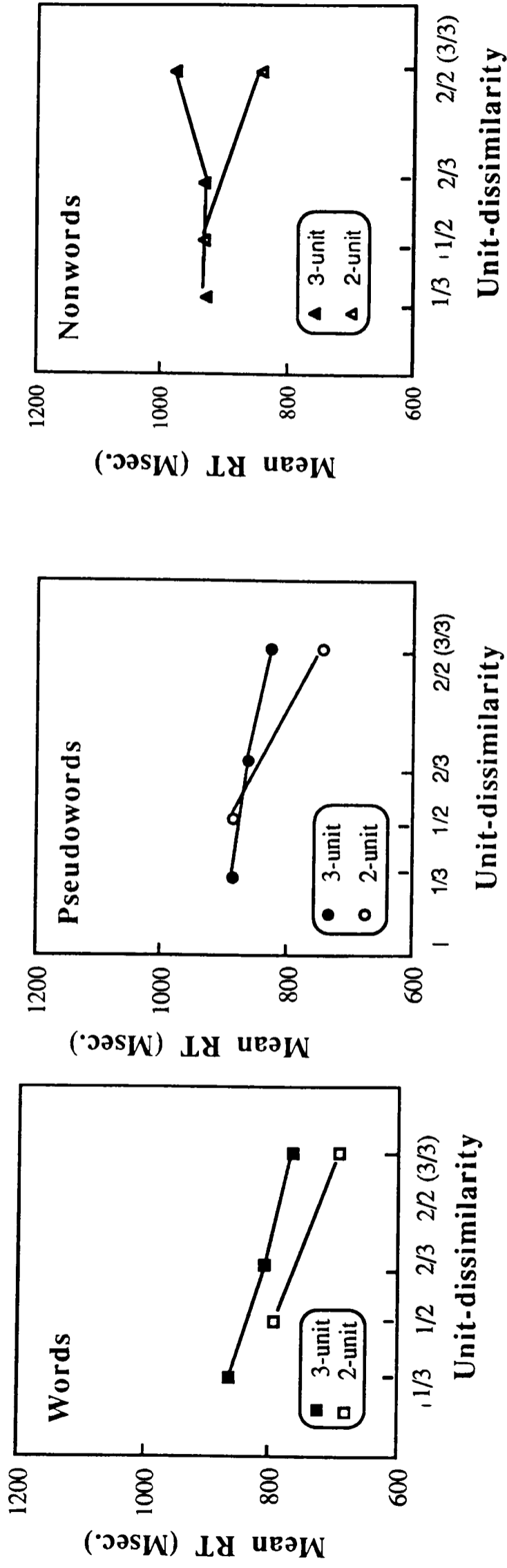


Figure. 4.5. Mean 'Different' RTs (msec.) by unit-dissimilarity and unit-complexity

As in the 'Same' judgements, in each condition there was a clear word superiority and pseudoword advantage in terms of comparison speed. Regarding RT differences between the different degrees of unit-dissimilarity, with 2-unit stimuli performance was consistently faster on pairs of stimuli with a greater number of mismatching units (1/2 vs 2/2). The effect was equally apparent for words, pseudowords and nonwords. For the 3-unit stimuli, the effects of unit-dissimilarity were evident for words, less so for pseudowords, and not for nonwords. In this respect, the effect of unit-dissimilarity parallels the effects of unit-complexity seen in the 'Same' judgments. Moreover, performance was faster on the 2-unit stimuli than that on the 3-unit stimuli. The main effects of word type, unit-complexity, unit-dissimilarity and their interactions were examined by ANOVA on RTs. For the 3-unit stimuli, the category of 2/3 unit-dissimilarity was not included, so that the number of categories of unit-dissimilarity in the 3-units can be matched with the 2-units, as required by the nested ANOVA. The three main effects were all highly significant (*Unit-dissimilarity*: $F(1,11)=39.40$, $P<0.001$, *Word-type*: $F(2,11)=11.51$, $P<0.001$, *Unit-complexity*: $F(1,11)=11.39$, $P<0.001$). A significant interaction was found between unit-dissimilarity and word-type ($F(2,22)=4.20$, $P<0.001$). There was also a marginally significant interaction between unit-dissimilarity and unit-complexity ($F(1,22)=4.96$, $P=0.05$). No other interactions were found among the three main effects.

Further analyses were performed to examine the differences in RT and error rates between the different levels of unit-dissimilarity in words, pseudowords and nonwords separately. The data are illustrated in Figure 4.5 separately for words, pseudowords and nonwords. The RT differences between different levels of unit-dissimilarity were only significant between 1/2 and 2/2, and between 1/3 and 3/3 and for words only, but not for pseudowords and nonwords (Table 4.5a). This accounts for the significant interaction between word type and unit-dissimilarity shown in the ANOVA.

The overall pattern of error rates was consistent with the main effects on RT. However, none of the differences in error rate between the different conditions were statistically significant.

Discussion

This experiment further confirms that the orthographic units, proposed in this study on the basis of purely orthographic-linguistic criteria, have psychological reality as functional word constituents in the visual analysis of Chinese characters. The evidence is provided both by the unit-complexity effect and by the unit-dissimilarity effect. Stroke-complexity did not appear to have any significant influence upon the speed or accuracy of visual comparisons. Furthermore, the effects of both unit-complexity and of unit-dissimilarity were most evident for word-comparisons, somewhat less in pseudoword-comparisons, and in nonwords the unit-complexity effect, at least, was no longer apparent. This suggests that recognition of individual orthographic units is better in a word context (that is, a word superiority effect for recognition of orthographic units). Moreover, comparisons of word-pairs were significantly faster than on pseudowords, and pseudowords in turn were faster than nonwords. Hence both superiority of words over pseudowords and pseudowords over nonwords has been demonstrated in this experiment. The effect of superiority of words over pseudowords indicates the functional role of lexicality (and perhaps visual familiarity) in the visual analysis of Chinese words. The advantage of pseudowords over nonwords indicates that positional regularity of the lexical radicals may play an important role in the visual analysis of Chinese words. These two types of word-superiority effect suggest that there is mutual interactive processing between orthographic units and words in the visual processing of Chinese. In general, these results suggest some functional parallelism between orthographic units in Chinese and

letters in English. Such orthographic units as the smallest stroke-pattern can be regarded as *functional* orthographic units in Chinese.

Experiment 3: Visual comparisons of Chinese script by non-Chinese readers

Introduction

This experiment was conducted to investigate whether the components (i.e., orthographic units) of Chinese can be recognized through visual familiarity even without the lexical information of these components. Using exactly the same stimuli and design as in the last experiment, non-Chinese readers, who encountered Chinese orthography for the very first time, were used as subjects. The issue addressed in this experiment is: Can the orthographic units of Chinese words be learned solely on the basis of visual familiarity with their graphic patterns, independent of lexical information and without any explicit instruction? As introduced earlier, orthographic units as the smallest stroke-pattern are not conventionally taught in the acquisition of literacy in Chinese. The sound and meaning of some orthographic units are no longer available even for highly educated Chinese readers, e.g., . The learning by native readers of Chinese of some orthographic units, at least, must rely mainly on visual familiarity with the graphic patterns through extensive reading experience.

In fact, acquisition of the orthographic units in Chinese without explicit linguistic knowledge has already been demonstrated in an early study (*Hull, 1920*) of associationism theory. In that study, English subjects, who did not know Chinese characters, were required to make feature-sound associations with Chinese characters. An artificial sound was taught as an association to each character. In fact, characters with the same components (radicals) were always assigned the same sounds. However, these components (radicals) of Chinese characters were not explicitly taught in the experiment. Subjects were required to

report the sound after each character was presented. Ninety six stimuli were divided into five sets in which each set had 12 stimuli with the same radical. After some practice, English subjects could even report the sounds of new characters before they were taught. This indicated that non-Chinese subjects can recognize the components (radicals) of Chinese characters when they recur frequently and consistently, even if they are not explicitly taught.

In this experiment, the encoding or recognition of orthographic units by non-Chinese is examined by means of the unit-complexity and unit-dissimilarity effects. To address the question whether the recognition of orthographic units of Chinese characters can be independent of any other linguistic knowledge of Chinese, in this experiment the 360 visual comparison trials previously used in Experiment 2 were repeated with English subjects. The stimuli were presented in three sequential blocks, in the order of words, pseudowords and nonwords, for all subjects. In this case visual familiarity increases as "lexicality" (for Chinese readers) diminishes across the stimuli, from the first block (words) to the third block (nonwords). If the orthographic units of Chinese characters can be recognized from visual familiarity alone, independently of any other linguistic information, then unit-complexity and unit-dissimilarity should be more likely to emerge by the last block (nonwords). If so, this would be the opposite effect to that shown in the visual processing of the same set of stimuli by native Chinese subjects: there both unit-effects were most evident in words and almost disappeared in nonwords. The word-types variable was therefore deliberately confounded with order (degree of practice) in this experiment. This confounding should have no significant effect if the Chinese characters were merely regarded as visual patterns by non-Chinese subjects. As graphic patterns, the words, pseudowords and nonwords in this experiment were all 1 cm x 1 cm squares in the same printed format. Neither the configuration as a whole nor the types of components were obviously different among the three types of words. Therefore

it is implausible to expect a word-type effect, as such, in the performance of non-Chinese subjects, who merely treated these compact squares as graphic patterns.

Method

Subjects

12 native English graduates (6 male and 6 female, aged 20-35 years old), at the University of Oxford, were paid to participate in the experiment. They were all right handed and had normal or corrected to normal vision. Prior to the present experiment, no subject had any significant exposure to Chinese characters.

Stimulus-Conditions

The stimuli and conditions used in this experiment were identical to those of Experiment 2. The presentation of the stimuli was also the same, except that the three types of characters were presented in three sequential blocks in the order of words, pseudowords and nonwords, for all subjects. The 2-unit and 3-unit stimuli in each word-type were counterbalanced among subjects.

Procedure

Apart from the experimental instructions, which were shown in English, the apparatus, experimental task and procedures were identical to the last experiment. The stimuli were called graphic patterns, with no knowledge of Chinese characters being involved.

Results

1) Visual familiarity and judgments effects

The visual familiarity effect was inferred from comparison over successive blocks. Mean values of RTs and error rates of the 12 non-Chinese subjects are

Table 4.6a. Mean reaction time (msec.) of 12 non-Chinese subjects by blocks and judgments in Experiment 3, along with statistical analyses

Blocks	No. of stimuli	Same		Different		Both		Statistical analysis t values † (Same vs Diff.)
		Mean	(SD)	Mean	(SD)	Mean	(SD)	
First (Word)	120	1295	(220)	1173	(198)	1234	(209)	2.52*
Second (Pseudoword)	120	1320	(249)	1157	(199)	1238	(224)	1.88
Last (Nonword)	120	1317	(248)	1212	(187)	1264	(250)	1.86
All blocks		1311	(239)	1181	(195)	1245	(228)	
<i>ANOVA (2 x 3)</i>		F value		P value				
Main effects								
Judgment (2)		5.56		0.03				
Blocks (3)		0.28		0.76				
Interaction								
Judgment x Blocks		0.47		0.63				

† t values were estimated by a paired t test with a degree of freedom of 11.

* P<0.05

Table 4.6b. Mean error rates (%) of 12 non-Chinese subjects by blocks and judgments in Experiment 3, with statistical analyses

Blocks	No. of stimuli	Same		Different		Both		Statistical analysis z value † (Same vs Diff.)
		Mean		Mean		Mean		
First	120	8.0		3.3		5.6		-2.94*
Second	120	7.1		4.3		5.7		-2.00*
Last	120	8.8		5.0		6.9		-1.89
All blocks		8.0		4.2		6.1		

† z values were estimated from Wilcoxon signed rank test.

* P<0.05

Word-type(block) effects in Experiment 2 & 3

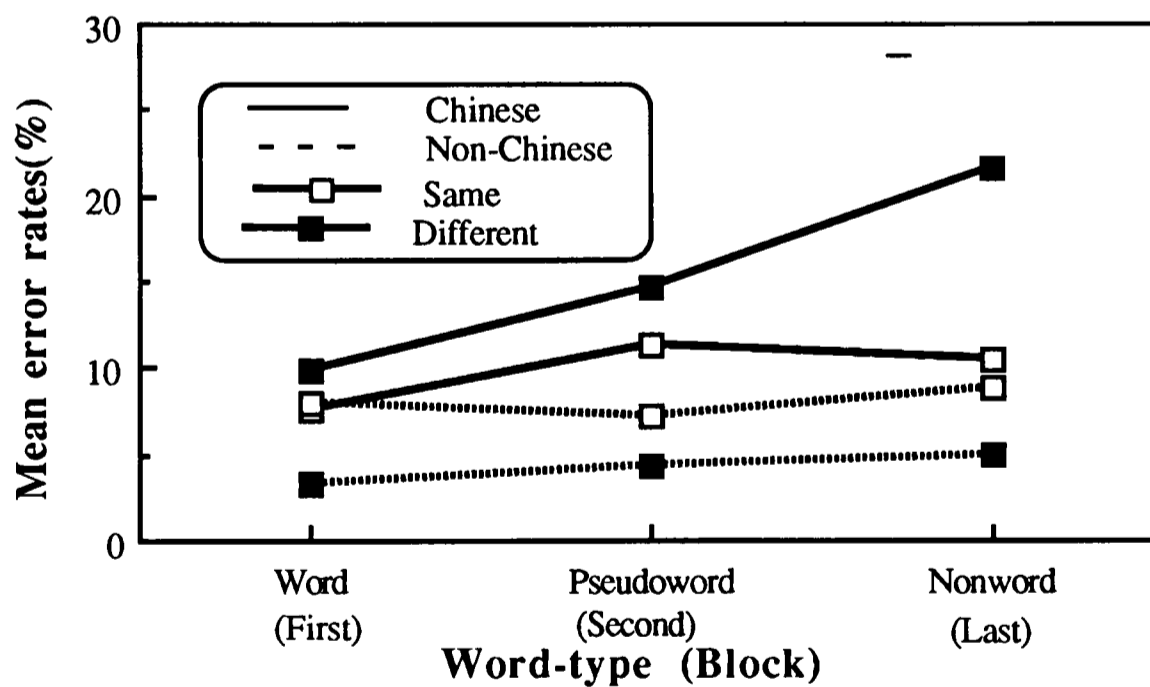
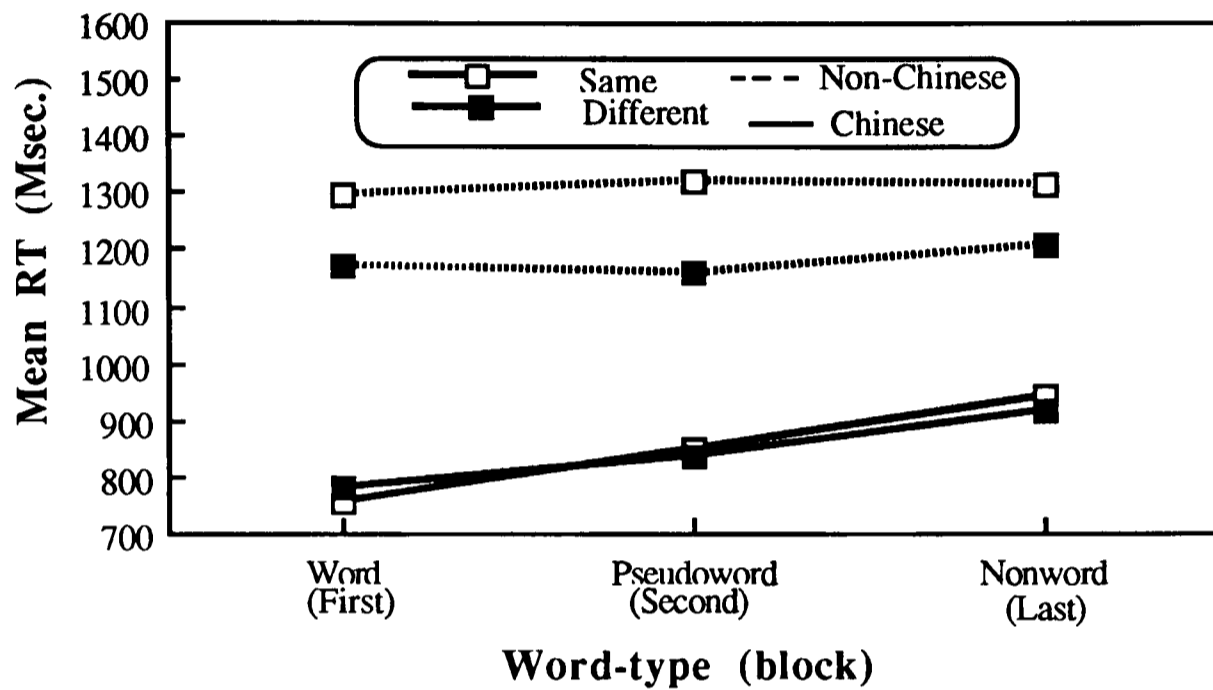


Figure 4.6 Mean RTs or error rates of both Chinese and non-Chinese subjects

shown in Tables 4.6a and 4.6b. Overall there were no apparent differences across the three sequential blocks (or word-types), in either RT or error rate analysis, on either 'Same' or 'Different' judgments. Performance on 'Different' judgements was consistently faster and more accurate than that of 'Same' judgments, in all three blocks. An ANOVA, with two factors of judgment and block, was performed on the RT data. A statistically significant effect was found only for judgment ($F_{(1,11)}=5.56$, $P<0.05$), but not for block ($F_{(2,11)}=0.28$, $P>0.05$). No interaction was evident on RTs between the two main effects ($F_{(2,22)}=0.47$, $P>0.05$). Therefore visual comparisons of Chinese scripts as graphic patterns by non-Chinese subjects did not improve significantly during 360 trials of practice.

For each block separately, performance on 'Different' judgements was faster than 'Same' judgements; the difference was significant in the first block but not in the second or the last block by paired t tests (Table 4.6a). Likewise, for each block, the 'Different' judgements were consistently more accurate than the 'Same' judgements, but the difference was significant only for the first and second blocks by Wilcoxon signed rank tests (Table 4.6b).

2) 'Same' judgments: unit-complexity and stroke-complexity effects

Tables 4.7a and 4.7b present the mean values of 'Same' RTs and error rates, subdivided respectively by block, unit-complexity and stroke-complexity. Overall, no difference in RT was evident among the three blocks. In each block, however, mean RT was consistently shorter on 2-unit than on 3-unit characters, even for stimuli with the same number of strokes (8-11 strokes). In contrast to Experiment 2 (Chinese subjects), this unit-complexity effect was most evident in nonwords, that is, in the last block, after 240 trials practice. For the first two blocks, the mean RT was shorter on characters with fewer strokes both in 2-unit and 3-unit stimuli, whereas in the last block no such effect was found; if anything, the

Table 4.7a. Mean reaction time (msec.) on 'Same' judgments of 12 non-Chinese subjects, by block (word-type), unit-complexity and stroke-complexity in Experiment 3, with statistical results

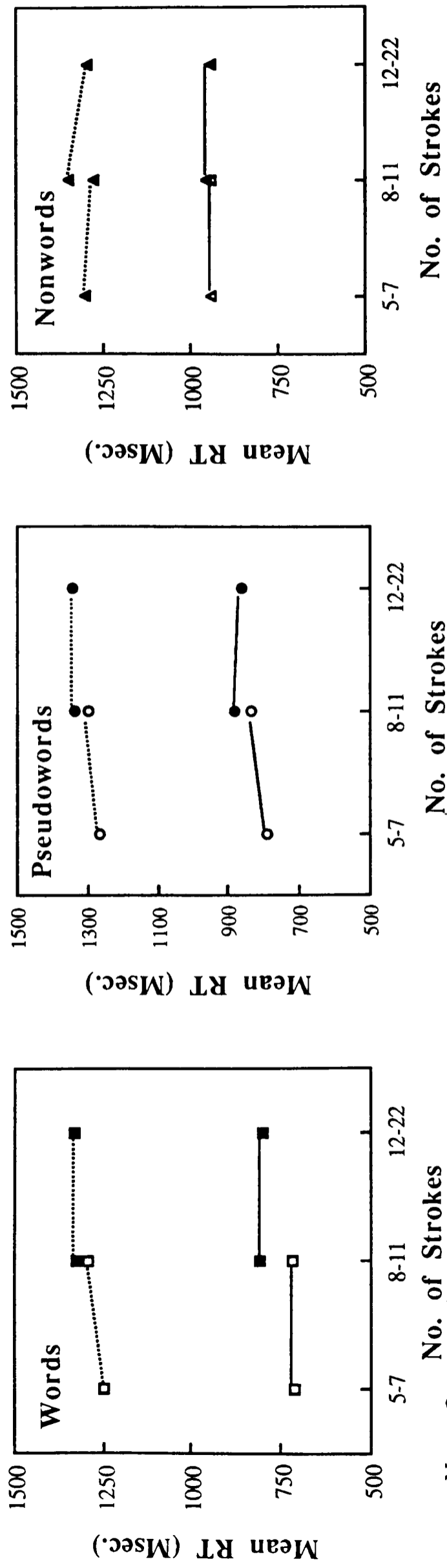
No. of Unit	No. of Strokes	First		Second		Last		All blocks	
		Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
2-unit	5-7	1251	(258)	1269	(256)	1301	(313)	1273	(275)
	8-11	1292	(308)	1300	(240)	1278	(339)	1290	(296)
	All	1271	(283)	1284	(248)	1289	(326)	1281	(285)
3-unit	8-11	1326	(256)	1340	(393)	1346	(106)	1337	(252)
	12-22	1333	(240)	1343	(331)	1295	(97)	1324	(193)
	All	1329	(248)	1341	(362)	1320	(101)	1330	(222)

<u>ANOVA(3 x 2 x 2)</u>	F value	P value
Main effects		
Blocks(B)	0.02	0.97
Stroke-complexity (S)	0.00	0.94
Unit-complexity (U)	0.94	0.35
Interactions		
B x U	0.10	0.95
S x U	0.38	0.55
B x S	1.41	0.27
B x S x U	0.00	0.99

Table 4.7b. Mean error rates (%) on 'Same' judgments of 12 non-Chinese subjects, by block (word-type), unit-complexity and stroke-complexity in Experiment 3

No. of unit	No. of stroke	First	Second	Last	All blocks
		Mean	Mean	Mean	Mean
2-unit	5-7	1.1	0.7	3.7	5.1
	8-11	1.4	3.7	1.9	2.3
	All	1.2	2.2	2.8	3.7
3-unit	8-11	2.2	1.3	2.9	2.1
	12-22	2.2	3.2	1.3	2.2
	All	2.2	2.2	2.1	2.1

'Same' judgment: Unit-complexity and stroke-complexity effects (Experiment 2 & 3)



Note: Open symbols: 2-unit stimuli; Filled symbols: 3-unit stimuli

Figure. 4.7. Mean 'Same' RTs by stroke-complexity and unit-complexity for of Chinese (solid line) and non-Chinese (dot line) subjects

stroke-complexity effects were reversed. The three main effects of block, stroke-complexity and unit-complexity were examined by an ANOVA on the RT data for 'Same' judgments. However, neither main effects nor interactions were significant.

On error rates (Table 4.7b), the patterns of block, unit-complexity and stroke-complexity effect were consistent with those on RT. However, no significant differences were found on error rates.

3) 'Different' judgments: unit-dissimilarity

The mean 'Different' RTs of the 12 non-Chinese subjects are presented in Table 4.8a, subdivided by block, unit-complexity and unit-dissimilarity, along with the statistical analyses. The corresponding data on error rates are given in Table 4.8b. As seen in Table 4.8a, the unit-dissimilarity effect became evident only in later blocks. By the second and the last blocks, it was present as a consistent trend on both 2-unit and 3-unit stimuli. Although the three main effects of block, unit-complexity and unit-dissimilarity remained non-significant by the ANOVA on 'Different' RTs, there was a highly significant interaction of unit-dissimilarity and block revealed by this ANOVA ($F_{(1,22)}=6.73$, $P<0.001$). Further analyses on both RT and error rates were applied separately in each block. There was a significant difference in RT between 1/3 and 3/3 in the last block ($t_{(11)}=2.56$, $P<0.05$), but not in the first two blocks. Consistent with RT, the error rates on 3/3 unit-dissimilarity in the last block appeared to be smaller than on stimuli with 1/3 unit-dissimilarity. However, this difference in error rate was not statistically significant.

4) Comparison of performance between Chinese and non-Chinese readers

Experiments 2 and 3 differed only in that they were performed by native Chinese and non-Chinese subjects respectively. Therefore it is possible to compare the visual processing of Chinese words as words and as graphic patterns by these Chinese and non-Chinese subjects respectively.

Table 4.8a. Mean reaction time (msec.) on 'Different' judgments of 12 non-Chinese subjects, by block (word-type), unit-complexity and unit-dissimilarity in Experiment 3, with statistical analyses

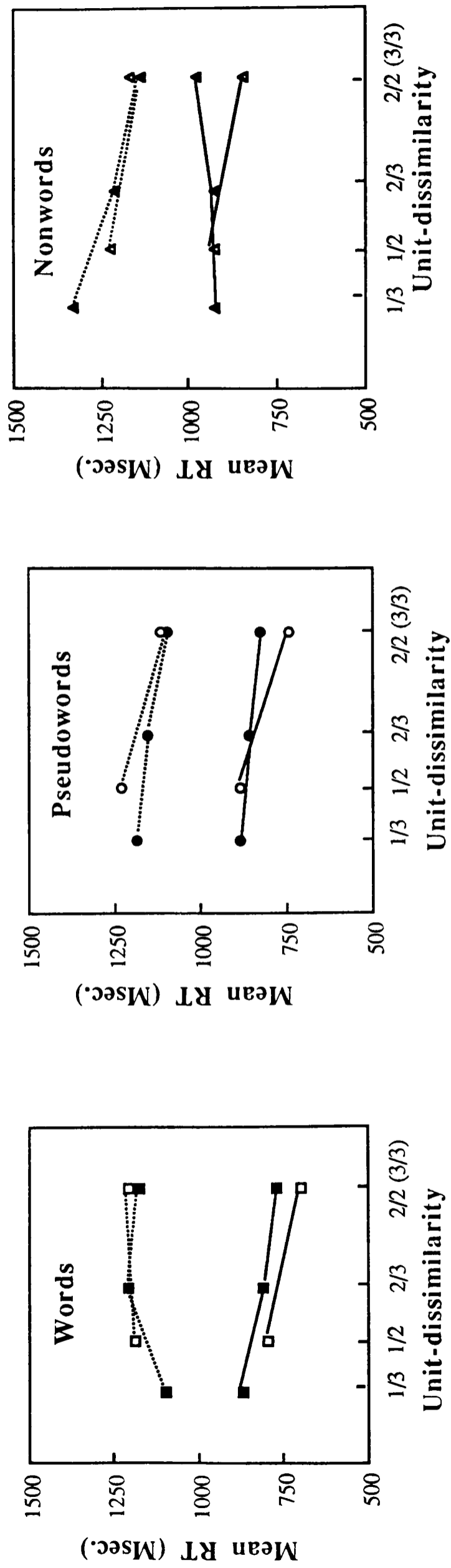
No. of Unit	U.D.	First		Second		Last		All blocks	
		Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
2-unit	1/2	1186	(237)	1230	(268)	1223	(224)	1213	(243)
	2/2	1209	(234)	1113	(237)	1165	(255)	1162	(242)
	All	1197	(239)	1171	(252)	1194	(239)	1212	(242)
3-unit	1/3	1095	(196)	1186	(243)	1330	(280)	1203	(283)
	2/3	1203	(243)	1157	(231)	1208	(217)	1189	(230)
	3/3	1171	(235)	1096	(185)	1136	(268)	1134	(229)
	All	1156	(225)	1146	(219)	1224	(255)	1175	(247)

<i>ANOVA(3 x 2 x 2)</i>		F value	P value
Main effects			
Blocks (B)		0.77	0.47
Unit-dissimilarity (UD)		1.16	0.30
Unit-complexity (UC)		0.36	0.56
Interactions			
B x UD		6.73	0.00
B x UC		1.41	0.26
UC x UD		0.18	0.68
B x UD x UC		1.70	0.20

Table 4.8b. Mean error rates (%) on 'Different' judgments of 12 non-Chinese subjects by block (word-type), unit-complexity and unit-dissimilarity in Experiment 3

No. of unit	No. of stroke	First	Second	Third	All blocks
		Mean	Mean	Mean	Mean
2-unit	1/2	4.2	4.1	3.4	3.9
	2/2	2.7	4.8	6.9	4.8
	All	3.4	4.4	5.1	4.3
3-unit	1/3	1.4	4.8	5.5	3.9
	2/3	3.5	3.5	5.6	4.2
	3/3	4.8	3.4	4.2	4.1
	All	3.2	3.9	5.1	4.1

'Different' judgment: Unit-dissimilarity effect (Experiment 2 & 3)



Note: Open symbols: 2-unit stimuli; Filled symbols: 3-unit stimuli

Figure. 4.8. Mean 'Different' RTs by unit-dissimilarity and unit-complexity for Chinese (solid line) and non-Chinese (dot line) subjects

a) Speed of performance

The performance of the Chinese and non-Chinese subjects across the three types (or blocks) of stimuli, on both 'Same' and 'Different' judgments, is illustrated in Figure 4.6. Overall, the set of Chinese characters was matched much faster by the Chinese subjects than when they treated merely as graphic patterns, by the non-Chinese subjects. However, accuracy of performance between Chinese and non-Chinese subjects was not greatly different at least for the 'Same' judgments. Chinese subjects were slightly less accurate than non-Chinese subjects. However, there was no significant speed-accuracy trade-off effect in either group, as assessed by correlations between RT and error rates in each condition.

b) Fast 'Different' performance

In contrast to the widely reported fast-'Same' effect found in the early 1970s studies, a faster 'Different' effect was observed in the non-Chinese subjects performance. However, this judgement effect was significant only in the first 120 trials, when the subjects were less familiar with the stimuli as graphic patterns (Table 4.6a, b). Moreover, there was no judgment effect for the Chinese subjects. This further supports the view that Fast-'Same' is by no means irreversible.

c) 'Same' judgments: Unit-complexity and stroke-complexity

In Figure 4.7, the mean 'Same' RTs of both Chinese and non-Chinese subjects are illustrated separately for words (or first block for non-Chinese), pseudowords (or second block) and nonwords (or last block), by different levels of unit-complexity and stroke-complexity.

In general, performance by the non-Chinese subjects was consistently slower in each condition, compared with the Chinese subjects. For Chinese subjects, the

difference in RT between the 2-unit and 3-unit characters, i.e., the unit-complexity effect, was significant for words, and tended to decrease from words to pseudowords to nonwords. For non-Chinese, there appeared to be the opposite tendency, although the effect was not significant. There was no stroke-complexity effect in the 'Same' judgments either in Chinese subjects or in non-Chinese subjects.

d) 'Different' judgments: Unit-dissimilarity

In Figure 4.8, the RTs for both Chinese and non-Chinese subjects by unit-dissimilarity are illustrated separately for words (first block for non-Chinese), pseudowords (second block) and nonwords (last block). For Chinese subjects, the effect of unit-dissimilarity decreased across words, pseudowords and nonwords. For non-Chinese subjects, the opposite effect was found. The effect of unit-dissimilarity was not shown in the first block (words) but began to emerge in the second block (pseudowords). It became significant in the last block (nonwords) for at least 3-unit stimuli.

Discussion







In the present experiment, non-Chinese subjects were required to judge the visual identity of Chinese characters as graphic patterns in three sequential blocks. Unit-complexity and unit-dissimilarity effects, which significantly influenced the performance of Chinese subjects, did not exert any significant influence upon non-Chinese subjects' performance on the first and second blocks with 240 trials. However, the unit-dissimilarity effect became significant in the last block (after 240 previous trials). Consistent with unit-dissimilarity, the unit-complexity effect also appeared, although as a non-significant trend, in the last block. Together these results suggest that the pattern of orthographic units in Chinese can be recognized by visual familiarity alone, independently of linguistic knowledge. However, the amount of trials (360 pairs) used in this experiment was

very small compared to real reading. It is noteworthy that the RT variance for non-Chinese subjects was uniformly large (about 250 ms) over words, pseudowords and nonwords, which was similar to (or greater than) Chinese subjects in nonword comparisons. The fewer overall significant effects found in non-Chinese subjects may be attributed, to some extent, to the large variance over all three blocks (contrasted with the much smaller SDs of Chinese subjects, especially in word-comparisons). Thus, an experiment with a larger number of trials is needed in the future.

General Discussion

1) The functional orthographic unit of Chinese words: *the stroke-pattern but not the individual strokes*

In the present study, for skilled Chinese readers, the visual comparison of pairs of Chinese characters was affected by the number of constituent units (unit-complexity) for pairs of 'Same' stimuli and by the number of mismatching units (unit-dissimilarity) for pairs of 'Different' stimuli. Since both unit-complexity and unit-dissimilarity were measured in terms of the orthographic units defined in this study (stroke-pattern), the demonstration of both unit-complexity and unit-dissimilarity effects suggests that these orthographic units of Chinese words can be (implicitly) recognized and are in fact used by skilled native Chinese readers in the visual analysis of Chinese script. The unit-complexity effect was not confounded by a possible stroke-complexity effect since it was significant after controlling for the number of strokes. By contrast, individual strokes had no influence upon visual comparisons of Chinese characters by native readers, even with control of the unit-complexity effect. This finding is inconsistent with the conclusion put forward in a recent study of the visual analysis of Chinese characters that used an illusory conjunction paradigm (Fang & Wu, 1989). In that

study, both individual strokes and stroke-patterns were found to produce migration errors, and on these grounds both were therefore considered as independent perceptual units. The authors claimed that "... *local components of single strokes and stems are equally likely to be independent perceptual units of character perception*". However, at least from the stimuli they provided in the paper, e.g., , migration errors of individual strokes appeared to occur only in characters containing a single unit. The present study is primarily concerned with words that have more than one unit (ZI) (i.e., over 90% of Chinese words). The claim that individual strokes are unlikely to have a significant influence upon the visual processing of Chinese words can at least hold for over 90% of Chinese words. Stroke-complexity did not even affect the visual analysis of Chinese characters by the non-Chinese subjects, who could only treat the Chinese scripts as graphic patterns. This suggests that even as graphic patterns the individual strokes have little significance. As introduced in Chapter 2, the separation of strokes in Chinese words is a matter of writing convention, rather than perception. In Experiment 3, any tendency toward better performance on stimuli with fewer strokes, i.e., a stroke-complexity effect, had disappeared in the second block and was even reversed in the last block. This indicates that non-Chinese subjects may attempt to identify the individual strokes at first. However, they seem to give up the attempt when they find that such a stroke as , e.g., a dot or a line is not an independently recurring component and that the number of strokes can be as many as 22. This was confirmed by subjects' report of their strategies after the experiment. Some of them could even report some of the stroke-patterns (i.e., units) such as (in their own words), "B-like pattern", "  ", "dots in a horizontal line or a vertical line, "..." or "  ", "a square", "  ", etc.. Note that they described the four dots at the bottom as a horizontal line rather than merely four dots. In fact, the individual strokes in the stroke-pattern, e.g.,  in Chinese may not be qualitatively different from a single stroke in an alphabetic letter, e.g., – in 'f'. The functional orthographic unit in Chinese, corresponding to

individual letters in English, thus appears to be the stroke-pattern rather than the individual strokes (at least for over 90% of Chinese words that contain more than one unit).

2) The pseudoword advantage in Chinese: *The importance of the lexical radical and its positional regularity*

In Experiment 2, pseudowords were matched faster and more accurately than nonwords. Given that the sole difference between pseudowords and nonwords was the positional regularity of the lexical radicals, this suggests that the lexical radical and its positional regularity play important roles in the visual analysis of Chinese characters. The importance of the lexical radicals has been suggested also in a recent study of the phonological processing of Chinese characters (*Perfetti and Zhang, 1991*). A radical-priming effect was found, i.e., characters primed by other characters containing the same radical were named faster. This effect, unexpected by the experimenters, was considered as a radical-based perceptual effect in that study (*Perfetti and Zhang, 1991*). The functional role of the lexical radicals as semantic components is further examined in the study of semantic processing of Chinese words in Chapter 7.

3) Word superiority over pseudowords—*"word" rather than "character"*

The visual comparison of words was found to be both faster and more accurate than that of the pseudowords. The standard account for word superiority over pseudowords, as already suggested in the English studies, appeals to lexicality, although orthographic familiarity (word-likeness) can hardly be entirely ruled out (*Henderson, 1982*). The importance of Experiment 2 is that both types of superiority effect, of words to pseudowords and of pseudowords to nonwords have been demonstrated with single-character words in Chinese. The demonstration of the superiority effects of words and pseudowords with individual characters in Chinese suggest that the so-called "characters", the basic

units in Chinese, have psychological reality equivalent to "words" rather than "morphemes" in English. The conventional term Chinese "characters" should be more appropriately changed to Chinese "words"

4) Two types of word superiority effects-*Mutual interaction between units and words*

In Experiment 2, both unit-complexity and unit-dissimilarity effects were found strongly in words, less so in pseudowords and absent in nonwords. This indicates that the recognition of orthographic units is better in word contexts. Thus a word superiority effect, in terms of the recognition of word constituents, has now been demonstrated in Chinese (with individual characters) as well as in alphabetic scripts. The absence of these unit-effects in nonword comparisons suggests that recognition of the lexical radicals may rely particularly on their positional regularity. (In the comparison of nonwords, in which the lexical radical was in an illegitimate position, neither of the unit-effects were significant.) Although the same set of lexical radicals was used in the words, pseudowords and nonwords in this experiment, the nonwords appeared so unfamiliar to the native readers of Chinese that they regarded them, in their own words as, "Japanese Kana", simply because the lexical radicals of the nonwords were in a position they have never met in real words. This further supports the claim that the identity of orthographic units is unlikely to be recognized independent of their positional information.

As discussed before, orthographic units (especially the lexical radical) were recognized better in words (Experiment 2). In the same experiment, words that had better-recognized units also showed a superiority effect over pseudowords in the visual comparison task. These two aspects of word-superiority, demonstrated in this study, suggest a mutually interactive processing between the recognition of word constituents and the recognition of words in Chinese. Hence, the interactive property of such visual word recognition models as McClelland & Rumelhart (1981) may have wider applicability than just being restricted to alphabetic

orthographies. However, an orthography-specific feature of Chinese may be that such mutual interaction is particularly important between words and their constituent lexical radicals, since the lexical radicals possess not only specific structure but also the positional regularity that determines the orthographic regularity of Chinese words (individual characters).

Conclusions

In conclusion, the present study has demonstrated that the visual analysis of Chinese characters can be based upon relatively well-defined orthographic units, particularly the lexical radicals. The functional orthographic units used by native readers are the stroke-patterns, not the individual strokes as previously thought. There is to some extent a functional parallelism between the orthographic units of Chinese and the letters of English. As in English, recognition of orthographic units in Chinese involves positional information and can benefit from word context. The mutual interaction between recognition of the units and recognition of the words has thus been demonstrated also in Chinese (with individual characters), as shown by the two types of word superiority effects here.

In addition, the visual comparisons undertaken by the non-Chinese subjects over limited amounts of visual exposures indicated that the graphic patterns of the orthographic units can be recognized by visual familiarity alone, without necessarily involving any underlying linguistic knowledge. This implies that even though the orthographic units (i.e., stroke-patterns) are not explicitly taught in learning to read Chinese words, the graphic patterns of the units can be (implicitly) acquired through extensive visual exposure, independently of any underlying linguistic features. This awaits more evidence from future studies (especially developmental studies) in reading Chinese. The word superiority over

pseudowords suggests that the individual character in Chinese may have functional parallelism with the "word" rather than the "morpheme" in English.

Chapter 5

Phonological and Semantic Processing

Introduction

The previous chapter described three experiments on visual processing of Chinese script. In the following three chapters, the thesis will focus on phonological and semantic recoding of written Chinese words, with special emphasis on the role of lexical decomposition. The studies of lexical decomposition in Chinese focus on the functional roles respectively of the lexical radicals in semantic recoding and of the non-radical components in phonological recoding. Lexical radicals and non-radical components can be distinguished, in principle, by recognizing the specific structure and position of the lexical radical. The structure and positional regularity of lexical radicals were shown to be implicitly recognized by skilled native readers of Chinese in the visual analysis of Chinese words (Experiment 2). According to the orthographic rules of Part-to-Whole Phonological Correspondence and Part-to-Whole Semantic Correspondence proposed in this study (Chapter 2), the *regular* functional role of lexical radicals and non-radical components was defined as semantic and phonetic respectively, independent of their etymological origin. The core issues to be addressed in this chapter are whether the rule-defined phonetic role of the non-radical component and the rule-defined semantic role of the lexical radical are in fact exploited by skilled readers of Chinese, in phonological and semantic processing tasks respectively.

To ensure maximum internal coherence and comparability across these investigations, both the phonological and semantic processing tasks made use of the simultaneous comparison paradigm. The critical manipulation of the stimuli concerns the occurrence of a graphic match between members of a stimulus pair,

with respect to one or other orthographic component, either on the lexical radical or on the non-radical component.

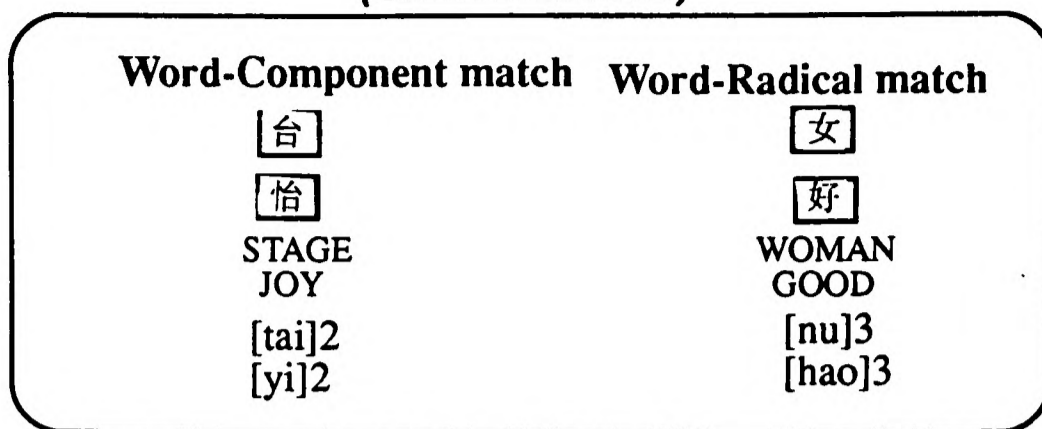
Background & Rationale

In both comparison tasks, pairs of words consisted of one simple word, and one compound word in which the simple word was embedded as a component. Which of the two orthographic components in the compound word matched the simple word was the critical manipulation, thereby forming one of two main conditions: Word-Radical match or Word-Component match. The Word-Radical match condition refers to pairs of words in which the simple word is embedded in its paired compound word as *the lexical radical*, e.g., 日 SUN [ri]4 BRIGHT [ming]2 (日 is the lexical radical of 明). The Word-Component match condition refers to pairs of words in which the simple word is embedded in its paired compound word as the non-radical component, e.g., 子 -CHILD [zi]3, 好 GOOD [hao]3 (子 is the non-radical component of 好). Both Word-Radical match pairs and Word-Component match pairs consisted of two words that were neither synonyms nor homophones of one another, nor did the pair form a meaningful two-character CI. In other words, in both the Word-Radical match and Word-Component match conditions, the component that matched graphically with the simple word in the same pair was both phonologically and semantically incongruent with its constituent word (the compound word) in the same pair. Such Word-Radical match and Word-Component match pairs are thus phonologically and semantically mismatching pairs, and can therefore be used as 'Different' pairs in both semantic and phonological comparisons.

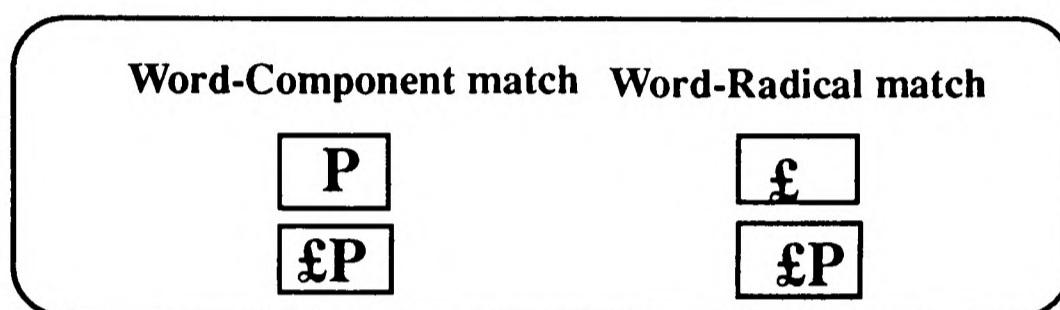
Figures 5.1 and 5.2 illustrate these two critical conditions for the phonological and semantic comparisons respectively. In both Figures, stimuli are first illustrated with real Chinese words, and also by means of symbols and geometric figures that may be more familiar to non-Chinese readers. In the symbolic

Task: Is the pair of words 'Same' or 'Different' in pronunciation? 90

*Real stimuli
(Chinese words)*



Symbolic analogy



Graphic illustration

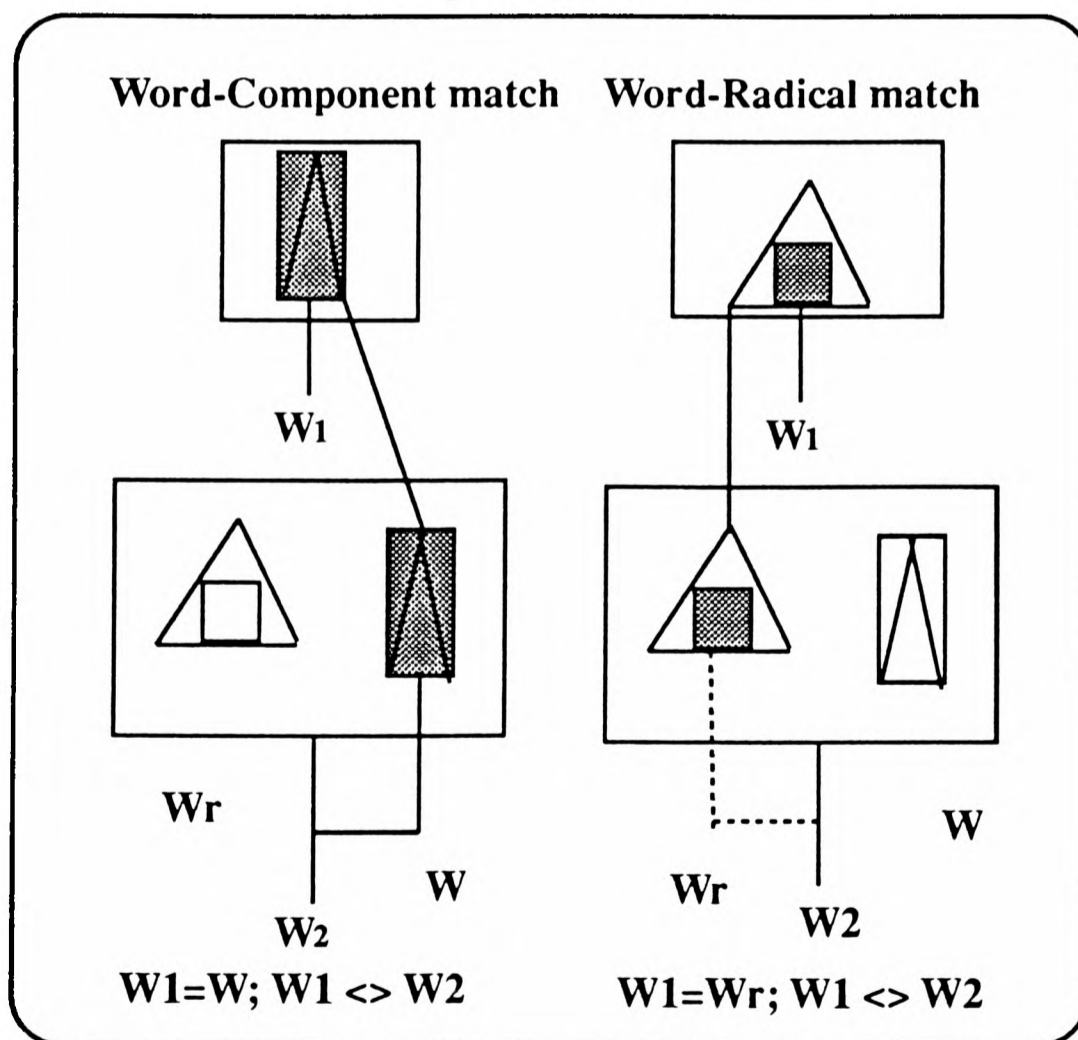
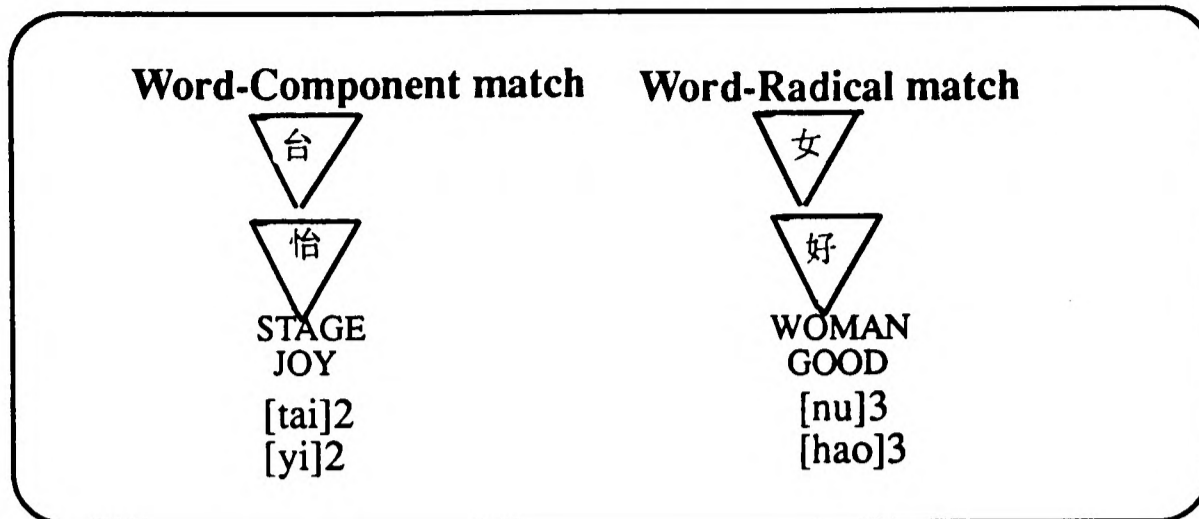
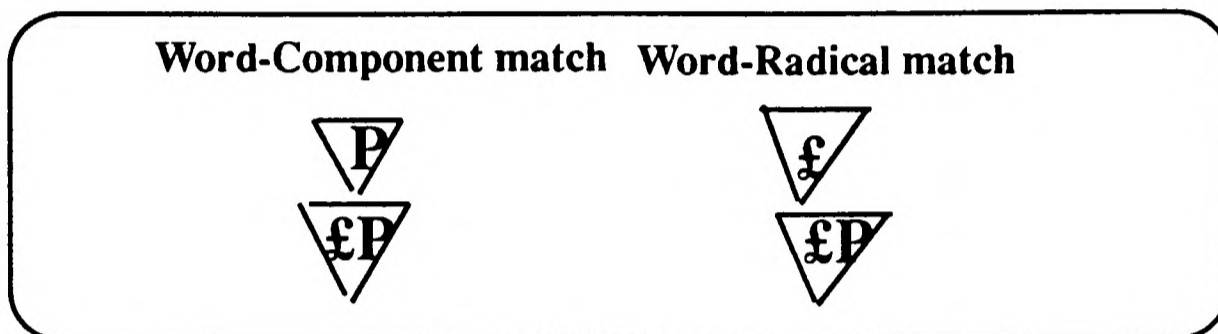


Figure 5.1 The illustrations of pair of words (W1 & W2) in Word-Component match (W) and Word-Radical (Wr) match conditions for phonological comparison task

*Real stimuli
(Chinese words)*



Symbolic analogy



Graphic illustration

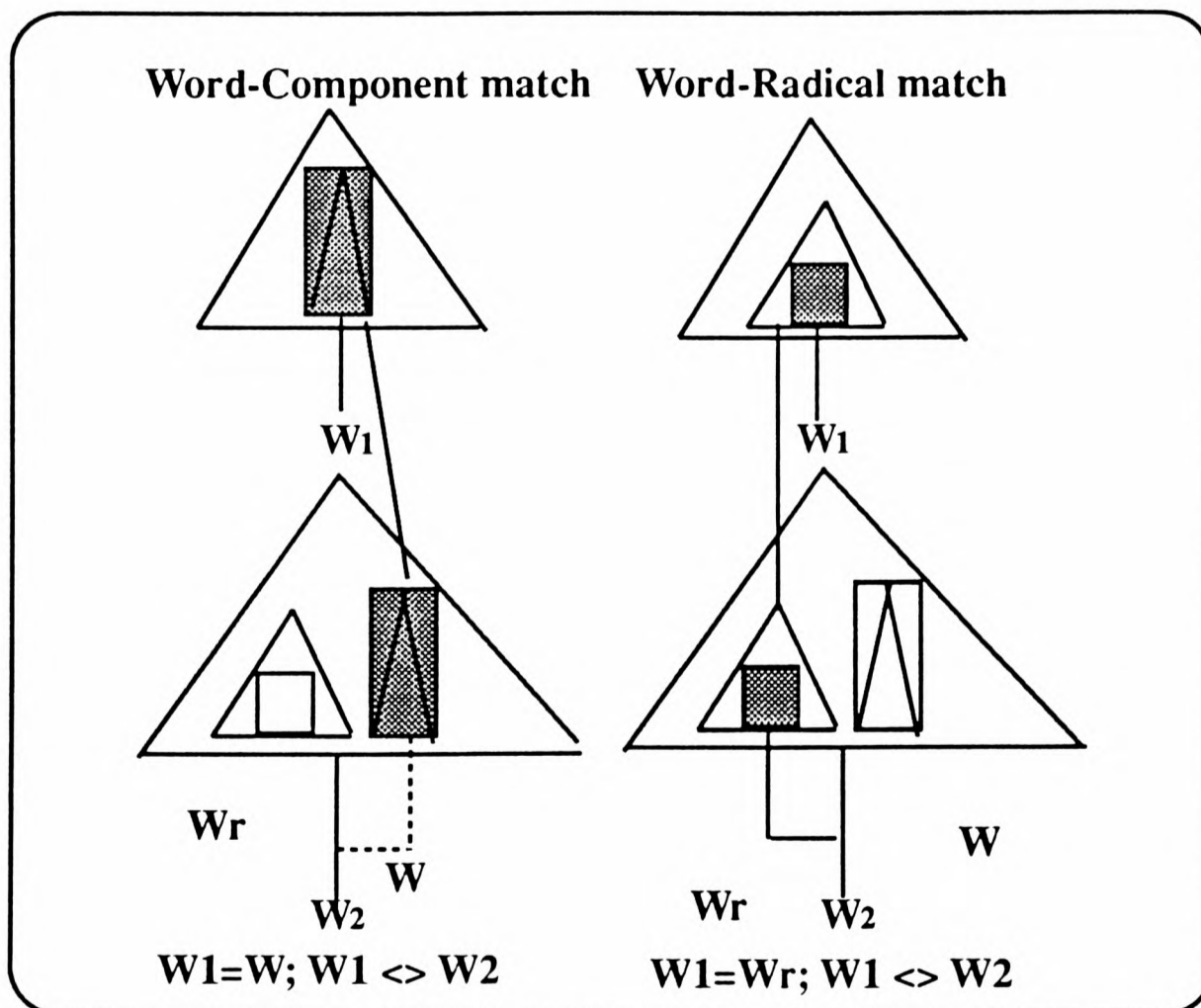


Figure 5.1 The illustrations of pair of words (W1 & W2) in Word-Component match (W) and Word-Radical (Wr) match conditions for semantic comparison task

analogy, the ideographic symbol "£" represents the lexical radical (i.e., the rule-defined semantic component), while the alphabetic letter "P" refers to the non-radical component (i.e., the rule-defined phonetic component of Chinese words). (Note that the term 'match' in the names of the stimulus conditions merely describes the manipulation of physical identity between components of a stimulus pair.) The idea is that the physical matching on one component of a pair of words may function as a 'distractor', which will bias subjects toward a false 'Same' response. Suppose that, in semantic judgments, subjects are more influenced by a matching lexical radical, and in phonological judgments they are more influenced by a matching non-radical component. Differential distractor effects of this kind would indicate a different attentional bias towards one or other orthographic component, depending on the strategic requirement of the task. In an analogy, if subjects treat £ and P as semantic and phonemic symbols, they would be more likely to judge the £ & £P pairs as semantically 'Same' symbols in the semantic comparisons. By contrast, in the phonological comparison task, P & £P pairs might be more likely to be judged as phonologically 'Same' pairs.

As introduced in Chapter 2, the lexical radical and non-radical component were defined as the *regular* semantic and phonetic components respectively. Hence a physical match on the lexical radical or on the non-radical component may act as 'semantic' and 'phonetic' distractors respectively. An attempt is made to portray this in the graphic illustration of Figure 5.1 and 5.2. There, a *triangle* represents the *semantic* component and a *square* represents the *phonological* component of Chinese words, respectively. Given that both Word-Radical match and Word-Component match pairs are non-homophone and non-synonym pairs (i.e., 'Different' pairs for both tasks), the distractor (shaded-in in Figures 5.1 and 5.2) in both conditions should bias subjects toward false 'Same' responses. The effect of the distractor, if any, would be to cause slower 'Different' responses and/or higher false 'Same' responses than pairs without any matching component.

The latter thus provide a necessary *baseline* condition for both semantic and phonological judgment tasks. Suppose that subjects allocate equal attention to all orthographic components, regardless of the recoding task involved, there should then be no significant difference, either on 'Different' RTs or on false 'Same' rates, between the two different distractor conditions. If, however, the radical and non-radical components are treated as 'semantic' and 'phonetic' component, respectively, the Word-Radical match condition (as semantic distractor) should have a stronger influence upon semantic comparisons; by contrast, the Word-Component match condition (as 'phonetic distractor') should be more important in the phonological comparison task. These opposite effects of the radical and non-radical components as differential distractors in semantic and phonological comparisons can be tested, even though the subjects and stimuli in these two main conditions are in fact identical.

Note that the pronunciation of the compound words in the Word-Component match stimuli is always incongruent with their non-radical component. Thus the compound words used in the Word-Component condition were in fact ideocompound words, in which the non-radical component as well as the lexical radical is defined as a semantic component according to the etymological word-formation rule (ideocompound rule). If this ideocompound rule was recognized perceptually, the interference of the Word-Component as a distractor should not be significantly different from that of the Word-Radical, in the phonological judgment task. If, on the contrary, the functional role of the matching component is implicitly assigned according to orthographic rule (Part-to-Whole correspondence) as defined in Chapter 2, independent of its etymology, the opposite effects, as described above, would still appear. This hypothesis was first explored in a pilot study, reported below.

Experiment 4: A pilot study of phonological and semantic comparisons of Chinese words

Introduction

This experiment, including both semantic and phonological comparison tasks, aimed to examine whether native readers of Chinese treat the lexical radical and non-radical component differently in semantic and phonological comparison tasks. The same set of stimuli in the critical conditions of Word-Radical match and Word-Component match, described already in the rationale, were presented to the same subjects in these two different tasks. In the phonological comparison task, subjects were required to judge whether pairs of words are the 'Same' or 'Different' in their *pronunciations* (in Mandarin), including the tones. Four weeks later, the same subjects were asked to judge whether the pairs of words are 'Same' (similar) or 'Different' in their *meanings*.

Method

Subjects

30 native Chinese readers and speakers from the mainland (21 males and 9 females), aged 25 to 35, were paid to participate in this experiment. Nine of them use Mandarin as their native dialect while 21 use Mandarin very often (at least in formal occasions). They were all reading for doctorates at the University of Oxford.

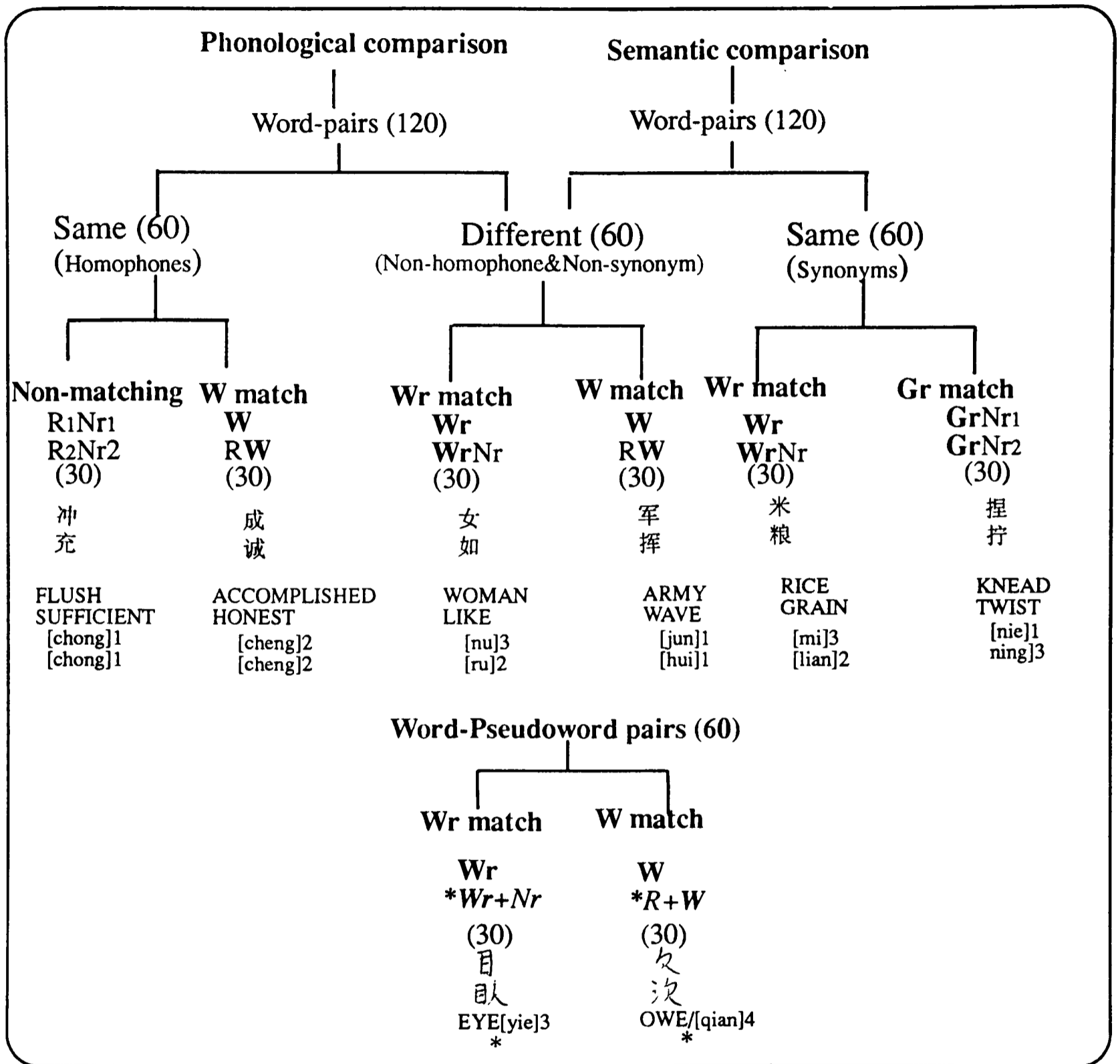
Stimulus Conditions

In each task, there were 60 word-pairs for the 'Same' judgments and 60 word-pairs for the 'Different' judgments. For each judgment, there were two conditions with 30 stimuli each. The two critical conditions for the 'Different' judgments of both tasks were Word-Radical match and Word-Component match. In addition, the experiment also used 60 word-pseudoword pairs in the Word-Radical match

Experiment 4

Task A: Does the pair of words have 'Same' or Different pronunciation ?

Task B: Does the pair of words have 'Same' or Different meaning ?



*Note: R: and Nr refer to lexical radical and non-radical component.
 W and Wr refer to Word-component and Word-radical
 Gr and G refer to Graphic-radical and Graphic-component
 *: Pseudoword
 Bold typeface indicates the matching component in each stimulus pair*

Figure 5.3. Stimulus conditions in Experiment 4

and Word-Component match conditions. The distribution of the stimuli in each condition (with corresponding examples) is shown in Figure 5.3. The stimuli for each condition are described below:

'Different' judgments for both tasks (Real-word pairs)

The pairs of words used for 'Different' judgments in both tasks were not synonyms, homophones, or meaningful two-character CI in Chinese.

Word-Radical match: In these word-pairs, the lexical radical of the compound word is identical to the adjacent whole word (simple word) in the same pair.

Word-Component match: In the word-pairs, the non-radical component of the compound word is identical to the adjacent whole word (simple word) in the same pair.

'Different' judgments for both tasks (Word-pseudoword pairs)

The two main conditions of Word-Radical match and Word-Component match were constructed in the same way as those in the real-word pairs. In these word-pseudoword pairs, the only difference was that the compound word in each pair is a pseudocompound, which consists of one lexical radical and one non-radical component in meaningless combination.

'Same' judgments for phonological comparisons (Homophones)

Graphically non-matching: In this condition, pairs of words have the same pronunciations, including tones (homophones) but have no graphic component in common.

Word-Component match: Corresponding to the Word-Component match condition for the 'Different' judgments, the pairs of words matched on the non-radical component also have the same pronunciations. That is, the non-radical

component has a pronunciation congruent with the compound word, in which it is embedded.

'Same' judgments for semantic comparisons (Synonyms)

Word-Radical match: Corresponding to the Word-Radical match condition for the 'Different' judgments, the pairs of words with a matching Word-Radical match are similar in their meanings (approximate synonyms).

Graphic-Radical match: In this condition, the pairs of words have similar meanings (synonyms) and also have the same graphic radical. The graphic radical is not a word in isolation and is therefore unpronounceable.

For all word-pairs, the two words selected for a pair do not form meaningful phrases or two-character CI in Chinese. The position of the simple word could be either at the top or at the bottom in different stimulus-pairs. The position was randomized in each condition. Pairs of homophones were always also identical in tone. Subjects were informed about this so that they would not concern themselves with the tone when they found two words with the same pronunciation. Given that Chinese pseudocompounds have virtually no pronunciations or meanings, word-pseudoword pairs were all classified as 'Different' pairs.

Apparatus

The apparatus used was a three channel tachistoscope, connected to a BBC computer. Two response buttons were placed in front of the subjects' left and right hands. Once a response button was pressed, the generator in the tachistoscope gave a noise free signal to the timer. Reaction times and errors for each trial were then recorded. The mean values of the reaction times and error rates of the stimuli in each condition were calculated for each subject, and recorded by a BBC computer.

Pairs of stimuli were presented simultaneously in the tachistoscope from a viewing distance of 508 mm. The paired stimuli, one above the other, were each 10 mm x 10 mm and separated 15 mm from centre to centre. They were written in black at the centre of a white display card (150 mm x 100 mm). A red cross ('+'), 5 mm x 5 mm, at the centre of the white display card was used as a fixation point.

Procedure

In the phonological comparison task, subjects were asked to judge whether a pair of words was the 'Same' or 'Different' in their pronunciation in Mandarin. In the semantic comparison task, the same subjects were required to judge whether pairs of Chinese words were the 'Same' (synonyms) or 'Different' in their meanings.

Each trial started with the appearance of the fixation point followed by a verbal 'ready' signal and then the stimuli. Each stimulus-pair was displayed for 150 msec. and immediately replaced by the fixation point, which remained until the next trial. Subjects were required to press the button as quickly as possible once a judgment was made. Half the subjects used the index finger of the left hand for the 'Same' judgments and the index finger of the right hand for the 'Different' judgments. The other half of the subjects responded with the hand/judgement pairing reversed. The full 180 stimuli in each experiment were presented randomly. The randomisation of the stimuli in each task was predetermined and was the same for all subjects.

Before each task, written instructions in Chinese, which explained the procedure, were given to each subject in addition to further oral instruction and explanation. Pseudowords were introduced as "unknown" words. Subjects were encouraged to guess the sound and/or meanings of the "unknown" words as if they had encountered them in reading without a dictionary at hand.

In each task, 30 practice trials were given. These trials were discarded from the stimulus set used in the main experiment. After every 60 trials (or at the subject's request), subjects were given a break of about 2 minutes. Each task lasted between 40 and 60 minutes for each subject.

Results

Performance in both tasks showed very high error rates, and in a number of subjects approximated 50%, as least in some of the conditions. In order to avoid comparisons in accuracy between different conditions being truncated by chance or near-chance error-rates in one (or more) condition(s), data from the subjects who had error rates approaching 50% overall in *either* the phonological or the semantic comparison task were discarded. Data from remaining 15 subjects were used for the analyses. These subjects had a mean error rate of 29% and 39% for the phonological and semantic comparisons, respectively. Given the high error rates even of these remaining 15 subjects, interpretation of RTs is clearly problematic. Accordingly, the principal analyses were based on mean error rates, although analyses of RT were also performed. Table 5.1a presents the mean error rates of the 15 subjects, by judgments and conditions, for both the phonological and the semantic comparisons, along with the results of the statistical analyses. The corresponding data for RTs are presented in 5.1b, along with the statistical results. Analyses were first performed on the data from each task and then on the data of the critical Word-Radical match and Word-Component match conditions from both tasks. Before turning to error rates, analyses on RTs are first reported briefly.

Reaction time (RT) analysis

Overall, 'Same' (homophone or synonym) pairs were judged faster than 'Different' (non-homophone and non-synonym) pairs. The judgment effect was

significant in both the phonological comparisons ($F_{(1,14)}=22.8$ $P<0.001$) and the semantic comparisons ($F_{(1,14)}=7.9$, $P=0.01$). The condition effect, however, was not significant on either the semantic or the phonological comparison tasks (Table 5.1b). There was no interaction between conditions and judgments (Table 5.1b) for either task.

Error rate analyses

1) Opposite effects of radical and non-radical matching on 'Different' judgments of the semantic and the phonological comparisons

The difference in error rates between Word-Radical match and Word-Component match (the condition effect) was significant on 'Different' judgments in both the phonological ($F_{(1,14)}=6.8$, $P<0.05$) and the semantic comparison ($F_{(1,14)}=20.3$, $P<0.001$) tasks. Performance on the 'Same' pairs was more accurate than that on the 'Different' pairs in the phonological comparison task. This judgment effect was marginally significant in the phonological comparison task. However, there was no such effect in the semantic comparison task (Table 5.1a). There was no significant interaction between the two main effects in either the phonological or semantic comparison tasks (Table 5.1a).

Further analyses were performed to examine the critical condition effect on the 'Different' judgements alone. The most important finding is that, as predicted, mean false 'Same' rates in the two critical conditions of Word-Radical match and Word-Component match showed a cross-over pattern between the phonological and semantic comparisons. With word-word pairs, the mean false 'Same' rates of Word-Radical match and Word-Component match were respectively 28% and 40% in the phonological comparison task, while they were respectively 45% and 31% in the semantic comparison task. The same cross-over pattern of false 'Same' rates in the two critical conditions was also found in the word-pseudoword pairs. Here the mean 'Same' response rates for Word-Radical match and Word-Component match

Table 5.1a Mean error rates (%) of 15 subjects by judgments and conditions in Experiment 4, with the statistical analyses.

Judgement	Condition (No. of stimuli)	Pronunciation		Meaning			
		Mean	(SD)	Mean	(SD)		
Same judgement		<i>Phonological comparisons</i>					
	Non-matching (30)	25	(12)	—	—		
	Word-component (30)	22	(11)	—	—		
	All homophones (60)	23	(12)	—	—		
		<i>Semantic comparisons</i>					
	Word-radical (30)	—	—	47	(12)		
	Graphic radical (30)	—	—	24	(15)		
	All synonyms (60)	—	—	35	(13)		
Different judgment		<i>Word-pairs</i>					
	Word-component (30)	40	(15)	31	(12)		
	Word-radical (30)	28	(8)	45	(14)		
	All pairs (60)	34	(11)	38	(13)		
		<i>Word-pseudoword pairs</i>					
	Word-component(30)	39	(12)	27	(15)		
	Word-radical (30)	22	(8)	45	(11)		
	All word-pseudo (60)	30	(10)	36	(13)		
Statistical analysis							
		Pronunciation		Meaning		Pronunciation & meaning	
ANOVA		F value	P value	F value	P value	F value	P value
Main effects							
	Judgments (J)	4.1	0.06	0.00	0.96	—	—
	Conditions (C)	6.8	0.02	20.30	0.00	—	—
	Task (T)	—	—	—	—	1.7	0.22
	Word type (W)	—	—	—	—	2.8	0.12
Interactions							
	J x C	3.0	0.11	0.60	0.44	—	—
	T x W	—	—	—	—	0.6	0.43
	T x C	—	—	—	—	46.5	0.00
	C x W	—	—	—	—	0.2	0.63
	T x W x C	—	—	—	—	3.4	0.09

Table 5.1b. Mean reaction times (msec.) of 15 subjects by judgments and conditions in Experiment 4, with the statistical analyses.

Judgement	Condition (No. of stimuli)	Pronunciation		Meaning			
		Mean	(SD)	Mean	(SD)		
Same judgement		<i>Phonological comparisons</i>					
	Non-matching (30)	672	(107)	—	—		
	Word-component (30)	614	(73)	—	—		
	All homophones (60)	643	(90)	—	—		
		<i>Semantic comparisons</i>					
	Word-radical (30)	—	—	632	(47)		
	Graphic radical (30)	—	—	594	(70)		
	All synonyms (60)	—	—	613	(58)		
Different judgment		<i>Word-pairs</i>					
	Word-component (30)	725	(89)	650	(56)		
	Word-radical (30)	725	(92)	667	(74)		
	All pairs (60)	725	(90)	658	(65)		
		<i>Word-pseudoword pairs</i>					
	Word-component(30)	693	(115)	607	(54)		
	Word-radical (30)	680	(102)	616	(36)		
	All word-pseudo (60)	686	(108)	611	(45)		
Statistical analysis							
		Pronunciation		Meaning		Pronunciation & meaning	
ANOVA		F value	P value	F value	P value	F value	P value
Main effects							
	Judgments (J)	22.8	0.00	7.9	0.01	—	—
	Conditions (C)	3.9	0.07	2.6	0.13	0.1	0.74
	Task (T)	—	—	—	—	10.1	0.00
	Word type (W)	—	—	—	—	19.7	0.00
Interactions							
	J x C	3.9	0.07	0.6	0.44	—	—
	T x W	—	—	—	—	0.1	0.73
	T x C	—	—	—	—	1.6	0.23
	C x W	—	—	—	—	0.2	0.65
	T x W x C	—	—	—	—	0.0	0.93

were 22% and 39% , respectively, in the phonological comparison task, while they were 45% and 27%, respectively, in the semantic comparison task. In both word-word pairs and word-pseudoword pairs, the difference in false 'Same' rates between Word-Component match and Word-Radical match was highly significant for both phonological (*real word*: $t_{(14)}=3.27$, $P=0.01$; *pseudoword*: $t_{(14)}=8.40$, $P<0.001$) and semantic comparisons (*Real word*: $t_{(14)}=7.43$, $P<0.001$; *Pseudoword*: $t_{(14)}=5.25$, $P<0.001$). Figure 5.4 illustrates the cross-over pattern of false 'Same' rates for Word-Radical match and Word-Component match as between the semantic and phonological comparison tasks, in both real word-pairs and word-pseudoword pairs.

An overall ANOVA was performed on the 'Different' judgments alone, (i.e., for the Word-Radical match and Word-Component match) in both the real word and the pseudoword conditions, and for both phonological and the semantic comparison tasks. The analysis examined the interaction of the critical condition effect with the effects of task (Phonological vs Semantic) and word-type (Real word vs Pseudoword). The ANOVA was performed on both error rates and RT.

On error rates, there was a highly significant interaction between task and condition ($F_{(2,26)}=46.5$, $P<0.001$). The main effects of condition, task and word-type were not significant (Table 5.1). The interaction between word type and condition was not significant (Table 5.1).

On RT, the main factors of task and word-type were both highly significant (Task: $F_{(1,14)}=10.1$, $P<0.001$; Word-type: $F_{(1,14)}=19.7$, $P<0.001$). Neither the condition effect nor the interaction between main effects (Table 5.1b) was significant. Word-pseudoword pairs were judged faster than real word pairs. The semantic comparison task was significantly faster than the phonological comparison task with no evident difference in error rate between the two tasks (Table 5.1b). The task effect was significant both in real word pairs ($t_{(14)}=4.53$,

**Component-matching effects in semantic and phonological judgments
(Experiment 4)**

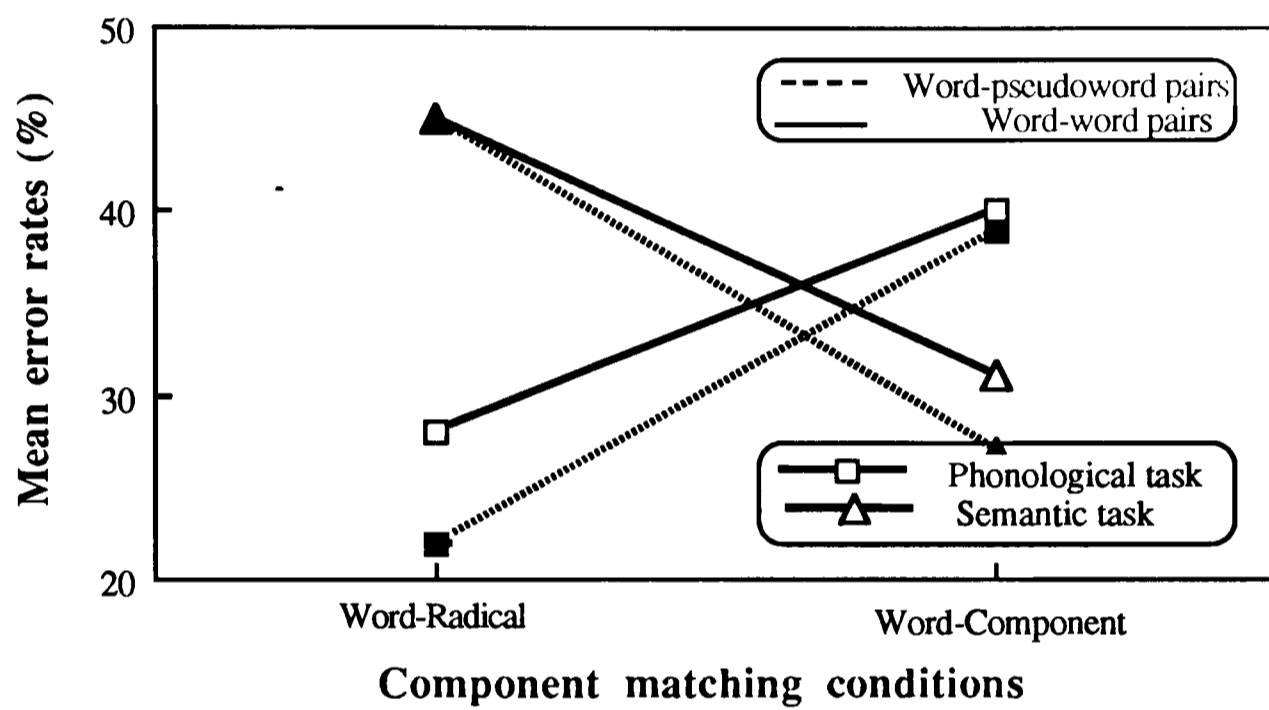


Figure 5.4 Mean error rates in component-matching conditions

$P < 0.001$) and word-pseudoword pairs for Word-Radical match ($t_{(14)} = 7.05$, $P < 0.001$) and in both real word-pairs ($t_{(14)} = 2.14$, $P = 0.05$) and word-pseudoword pairs for Word-Component match ($t_{(14)} = 2.86$, $P = 0.01$) in further analyses by paired t tests.

Discussion

The most important finding is that (1) in the pronunciation task subjects made significantly higher false 'Same' responses for pairs of words matching on the non-radical component (i.e., on the rule-defined phonetic component) than those matching on the lexical radicals. On the contrary, (2) when the task required judgments of meaning, the same subjects made significantly higher false 'Same' responses on word-pairs matching on the lexical radical (i.e., on the rule-defined semantic component) than on word-pairs matching on the non-radical components. The same cross-over pattern of false 'Same' rates was confirmed in both word-word and word-pseudoword pairs. That is, on semantic judgments the lexical radical acted as the more powerful distractor; on phonological judgments it was the non-radical component that had the greater distractor effect. This suggests that skilled readers of Chinese adopt a different selective attentional bias in semantic and phonological judgment tasks, in accordance with the Part-to-Whole correspondence rules, defined in this study.

However, this rule-like behaviour was only shown by error rates. The result is thus not fully convincing, especially since the mean accuracy of performance was not very high. Therefore, the finding needs to be re-examined with longer stimulus exposure time to permit meaningful RT analysis. This was done in the main studies of phonological and semantic comparisons described in the next two chapters.

Given that the subjects, apparatus, and procedures were exactly the same in all six experiments in the main studies, they are described in the following General Methods section and will not be repeated for each individual experiment.

General Methods for Experiments 5-10

1) Subjects

12 native readers and speakers of Chinese from the mainland (6 males and 6 females), aged 25 to 35, were paid to participate in the following six experiments on semantic and phonological comparison. An interval of at least two months elapsed between different experiments. All subjects graduated from university in mainland China and were D.Phil students at the University of Oxford. They all speak Mandarin. For nine of them, Mandarin is their native dialect; the other three use Mandarin at least for communication on formal occasions.

2) Apparatus

All experiments were conducted on a Packard Bell IBM PC-AT286 computer with monograph screen. Computer programmes were designed to randomize trials, control stimulus presentation, time the responses (RTs), record the errors and compute the mean value of RT and error rate, both by subjects and by items. The ET Chinese word processing package was connected with the computer programme both for word retrieval and for the construction of the pseudowords used in the experiments. Pairs of stimuli, one above the other, each 1 cm x 1 cm and separated 1.5 cm from centre to centre, appeared in a standard print format in white on a black background on the computer screen. Each pair covered a visual angle of 1° horizontally by 1.5° vertically from a viewing distance of 30 cm.

3) Paradigm & Procedures

For the phonological comparisons, subjects were asked to judge whether the words in a pair were the 'Same' or 'Different' in the Mandarin pronunciation (including tone). For the semantic comparison, they were required to judge whether the paired words were the 'Same' (synonym) or 'Different' in their meanings.

Before each experiment, subjects first read written instructions on the screen. Pseudowords were introduced as "unknown" words, and subjects were encouraged to guess the pronunciation of the "unknown" words in the phonological comparison task and to guess their meaning in the semantic comparison task.

Each experiment task was preceded by a practice session, with 30 stimulus-pairs. After the practice session, the mean reaction time and error rate of the 30 trials were shown on the screen. If the error rate was equal to or higher than 20%, further practice with different stimulus-pairs was given until the total accuracy was equal to or better than 85%.

The sequence of each trial started with a fixation cross, "+", 0.2 cm x 0.2 cm, at the centre of the screen. The cross indicated the position in which the stimuli would appear. It lasted 0.5 seconds. After a 0.5 second gap, a pair of stimuli was presented simultaneously in the positions immediately above and below the fixation point. The stimuli remained on the screen until the subject pressed a response key. Once a response key was pressed, the stimulus immediately disappeared and, after another 0.5 second, the fixation point returned to start the next trial. Fixation on the cross was always required and emphasized through each experiment.

In all experiments, subjects were asked to respond as quickly and accurately as they could by pressing one of two buttons. The response key for 'Same' was the left corner key 'Z' and for 'Different' was the right corner key slash '/'. Z and / were labelled 同 'Same' and 异 'Different' in Chinese respectively. Subjects were asked to press the left key (Z) for 'Same' and the right key (/) for 'Different' judgements, using their index finger of the left hand and the right hand, respectively.

After every 60 trials, the subject had a break for 2 minutes, accompanied by a piece of quiet electronic music. At the end of each experiment, subjects were required to report their strategies.

Chapter 6

From script to sound in Chinese

Theoretical issues

This chapter reports and discusses three experiments on phonological comparison of Chinese words, with special emphasis on lexical decomposition.

Three major theoretical issues are addressed:

1) What is the role of semantic recoding in the phonological processing of Chinese words ?

2) Is pronunciation of Chinese words using lexical decomposition biased towards the non-radical component, that is, the phonetic component, as defined by the lexical decomposition rule proposed in this study?

3) If so, how is this process modulated by type-frequency indices of concordance with other items in the lexicon?

Experiment 5: Phonological comparisons on pairs of Chinese words

Introduction

The main objective of this experiment is to investigate the role of semantic recoding in the phonological processing of Chinese words. In English, there is little doubt that the pronunciation of words can be achieved by direct mapping from orthography, without semantic mediation. This is supported, for example, by the observation of "word-meaning blindness" or "semantic dyslexia: patients may name a written English word correctly without knowing its meaning, even when the word has an irregular pronunciation (*Schwartz, Marin and Saffran, 1979*). To address the issue of whether such direct mappings from orthography to phonology exist also in the phonological recoding of Chinese words, Experiment 5 examines the synonym effect in a phonological comparison task. Suppose that the phonological recoding of a Chinese word is not possible without semantic

recoding. In this case, pairs of words, which are different in sounds but closely similar in meanings (near synonyms), e.g., 盘 DISH [die]²碗 BOWL [wan]³, would be judged more slowly or provoke more false 'Same' responses than pairs of words which are different in both meaning and pronunciation, e.g., 安 SAFE [an]¹ and 草 GRASS [cao]³. If, however, no such synonym effect is found, this would support the view that the phonological processing of Chinese words can be independent of semantic mediation. The direct mapping from orthography to phonology, described in current word recognition models, may not be unique to alphabetic scripts.

Stimulus Conditions

A total of 120 pairs of words was used in this experiment. Figure 6.1 illustrates the distribution of the 120 stimuli in the four different conditions, with corresponding examples. There were equal numbers of 'Same' (homophone) and 'Different' (non-homophone) pairs. For the 'Different' judgments, there were two different conditions, i.e., Control and Synonym, each with 30 pairs of stimuli. For comparability with the two different conditions for the 'Different' judgments, the 60 pairs for the 'Same' judgments were also divided into two equal groups (Homophones 1 and 2). Stimuli in each condition are further described as follows:

'Same' judgments (homophone pairs)

Homophone: The 60 pairs of homophones consisted of two different words which each have the same pronunciation, including tone, but no semantic relationship.

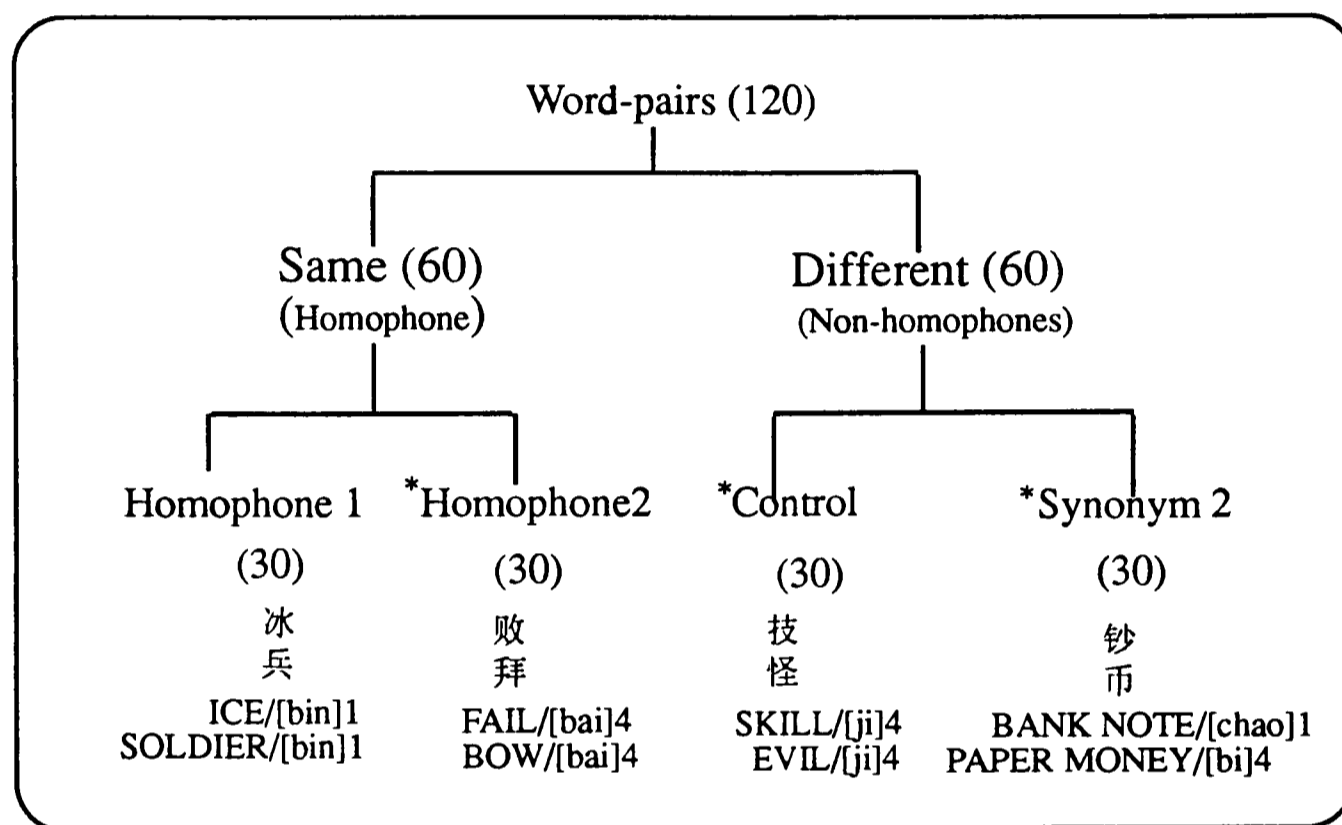
'Different' judgments (non-homophone pairs)

Synonym: In this condition, two different words in a pair have different pronunciations but closely similar meanings.

Control: This involves pairs of different words which are neither homophones nor synonyms.

Experiment 5

**Task: Does the pair of words have the Same or Different pronunciations?
(Phonological comparison)**



** the same set of stimuli was also used in Experiment 8*

Figure 6.1. Stimulus conditions in Experiment 5

Stimuli in these different conditions were randomised in this experiment. The randomisation was controlled by a computer programme and was different for each individual subject. The apparatus, method and procedures used in this experiment were the same as those described in the General Methods section of chapter 5.

Results

Table 6.1a presents the mean RTs (both by-subjects and by-items) by judgement and condition, along with the main statistical analyses. The corresponding data for error rates are given in Table 6.1b. An ANOVA was performed on RTs by-subjects while an ACOVA with the word frequencies of the two words in each pair as two co-variates was conducted on RTs by-items. In both the by-subjects and by-items analyses, the mean RTs were faster for 'Same' (homophone) than for 'Different' judgments (non-homophone). However, there was no apparent difference between the different stimulus-pair conditions. In the ANOVA by-subjects, the main effect of judgment was highly significant ($F_{(1,11)}=57.5, P<0.001$) but not the condition effect (Table 6.1a). The interaction between conditions and judgments was also not significant. After taking the word frequencies into account, this pattern was confirmed by the ACOVA by-items analysis; there was a significant judgment effect ($F_{(1,29)}=57.9, P<0.001$) but no significant condition effect and no interaction between them (Figure 6.1a). The two co-variates of word frequencies (individually or together) had no significant association with the two main effects nor their interaction in the ACOVA (Table 6.1a).

1) Synonym effect on 'Different' judgments

The difference between the synonym condition and the control condition was examined on 'Different' judgments in terms of both RTs and error rates. Consistent with the general analysis in the ANOVA, there was no significant difference in RT

Table 6.1a Synonym effects on phonological comparison: Mean reaction time (msec.) in Experiment 5, with statistical analyses

Judgments	Condition	No.of stimuli	By Subjects		By items		By Subjects	By items	
			Mean	(SD)	Mean	(SD)	(1) vs (2)	(1) vs (2)	
Same									
	Group 1	30	1022	(211)	1034	(121)			
	Group 2	30	1038	(194)	1042	(117)			
	All	60	1030	(202)	1044	(119)			
Different									
	(1)Control	30	1201	(253)	1217	(122)	$t_{(11)}=0.64$ P=0.55	$t_{(11)}=0.89$ P=0.38	
	(2)Synonym	30	1215	(242)	1242	(98)			
	All	60	1208	(247)	1229	(110)			
Statistical analysis									
			ANOVA		ACOVA		Co-1†	Co-2†	Co-1&2
			F value	P value	F value	P value	F value	F value	F value
Main effects									
	Judgments (J)		57.5	0.00	55.0	0.00	0.02	0.90	0.57
	Conditions (C)		0.1	0.35	0.6	0.46	0.01	0.27	0.17
Interactions									
	J x C		0.0	0.90	0.8	0.37	5.03*	0.81	2.59

† Co-1 and Co-2 refer to co-variate of the word frequency of the top and the bottom word respectively

*: P<0.05

Table 6.1b Synonym effects on phonological comparison: Mean error rates (%) in Experiment 5

Judgment	Condition	No. of stimuli	By Subjects		By-items		By Subjects	By items
			Mean	SD	Mean	SD	(1) vs (2)	(1) vs (2)
Same								
	Group 1	30	2.4	(2.3)	4.3	(7.2)		
	Group 2	30	4.7	(3.7)	10.0	(10.6)		
	All	60	3.5	(3.0)	7.1	(8.9)		
Different								
	(1) Control	30	2.3	(2.3)	5.6	(5.5)	Z=2.70*	Z=1.54
	(2) Synonym	30	4.7	(3.2)	9.9	(9.7)		
	All	60	3.5	(2.7)	7.7	(7.6)		

Z values were estimated by the Wilcoxon signed rank test.

* P<0.05

between the synonym and the control condition, by paired t test ($t_{(11)}=0.64$, $P=0.55$). In the by-subjects analysis, synonyms showed higher error rates (4.7%) than the control condition (2.3%) (Wilcoxon signed rank test $Z=2.70$, $P<0.05$). However, after adjustment for word frequency, the synonyms effect became non-significant on error rates ($Z=1.54$, $P>0.05$). Moreover, the effect was also not significant on RT ($t_{(11)}=0.89$, $P=0.38$).

Discussion

This experiment demonstrated that synonym-pairs were not judged significantly more slowly than pairs of words that were different in both pronunciation and meaning, in phonological comparisons. This suggests that semantic mediation or recoding is not necessary for phonological recoding of Chinese words. Nonetheless, some semantic recoding may take place. In parallel processing models of word recognition, semantic and phonological processing are automatically activated in word recognition. Subjects may use the semantic information even if it is not necessary to perform the task. In this experiment, subjects did have numerically higher error rates on synonyms than on the control pairs, even though the difference was not reliable in the by-items analysis. Hence, as in English, semantic recoding may be unnecessary but still occur in tasks involving the phonological processing of Chinese words. The implication of this finding will be discussed in detail in the general discussion of this chapter.

Experiment 6: Phonological comparisons of Chinese words with matching components

Introduction

The experiment was designed to address two issues concerning lexical decomposition in the phonological recoding of Chinese words, as follows:

1) Rule-defined phonetic component

In Chapter 2, the non-radical component of compound Chinese words was defined as the phonetic component, according to the rule of Part-to-Whole Phonological correspondence. The functional salience of the non-radical component in phonological recoding tasks received some support in the pilot study (Experiment 4), with stimulus-pairs matching either on the radical or on the non-radical component. In that pilot study (Experiment 4), subjects were biased by the matching non-radical component towards a false 'Same' response. This suggested that skilled readers of Chinese may attend selectively to the non-radical component in phonological judgments, in accordance with the Part-to-Whole Phonological Correspondence rule proposed in this study. However, in Experiment 4 this attentional bias was observed on error rates but not on RTs. The present experiment aimed to re-examine these component matching effects again, in a better controlled design.

In the present experiment on phonological comparison, instead of tachistoscopic displays of 150 msec., stimuli were displayed until a response was made. Moreover, the matching components (both lexical radical and non-radical components) were extended to include also Graphic-Radical and Graphic-Component, constituents which cannot be used as words in isolation in modern Chinese. These new conditions were intended to address the question of whether the independent lexical status of a component may affect the process of lexical decomposition. Is the pronunciation of Chinese words, based on the lexical

decomposition of the phonetic component, influenced by the congruency and consistency of that orthographic component in the lexicon? In Brown's model (1987), the frequency of an orthographic component and of its phoneme correspondences, summed over all appropriate words in the lexicon, plays an important role in lexical decomposition in English. In the current experiment, three different indices of an orthographic component were evaluated, summed over all words in the dictionary (Xin Hua, 1979). These were the Radical frequency index, the Congruency index and the Consistency Index. These three indices were evaluated for each of the matching components in the critical Word-Radical match and Word-Component match pairs (N=60). All three types of frequency indices were based on the type-frequencies of the components in the dictionary. The type-frequency of any orthographic component is defined as the total number of words in the dictionary in which the component appears, regardless of its word frequency. For instance, the component (also a word in isolation) SUN 日 (the matching component of the word-pair 日 & 明) appears in 151 words in the dictionary. The type frequency of 日 SUN is therefore 151.

2) Three types of frequency indices for orthographic components

a) *Radical Frequency Index (RFI)*

The "radical frequency" of an orthographic component was defined as the ratio of the number of words in the dictionary which contain that component as their *lexical radical*, over the type frequency of the same component. Given that the matching components of Word-Component match pairs were selected from simple words other than the list of 189 radicals in the dictionary, they cannot be regarded as the lexical radical in any of their constituent words. Therefore the radical frequency for the matching component in these Word-Component match pairs would be zero, according to the criterion given above. To avoid zeros (as required by the regression analysis), the actual Radical Frequency Index of a

component in this experiment is expressed by the radical frequency, defined above, plus 1, over type frequency plus 1. For instance, the matching component in the pair of stimuli 兵 & 宾, appears in 7 words in the lexicon, i.e., its type frequency is 7. None of these words has 兵 as their lexical radical. The Radical frequency of 兵 is $0+1/7+1$, i.e., $1/8$. The maximum value of RFI is 1.0. However, the Radical Frequency Index of a matching component in the Word-Radical match pairs can be less than 1, since although the matching component must be a lexical radical in the stimulus items selected for this condition, it may serve as the non-radical component in other words in the dictionary. For instance, in the Word-Radical match pair 马 & 驴, the matching component 马 HORSE [ma]3 is the lexical radical in 驴 DONKEY [lu]2. However, it is the non-radical component in 妈 MOTHER [ma]1. (See definition of lexical radicals in Chapter 2.) Thus the type-frequency of a given component can be higher than the radical frequency index of the same component. The distribution of the Radical Frequency Index of the lexical radicals and of the non-radical components is bimodal, with little or no overlap between them (Figure 6.4).

b) Congruency Index

The Congruency Index denotes the part-to-whole phonological congruence between a component and all other words containing that component in the lexicon. It is defined as the ratio of the total number of words in the dictionary that contain this component, and in which the pronunciation of the whole word is congruent with the pronunciation of the component *in isolation*, as a simple word, over the type frequency of the component. For example, 甲 [jia]3 FIRST, the matching component of the stimulus-pair 甲 & 押, appears in 12 words in the lexicon. Therefore the type frequency of 甲 is 12. Seven of these words are pronounced [jia]3, which is congruent with the pronunciation of in isolation. The Congruency Index of 甲 is therefore $7/12$.

c) Consistency Index

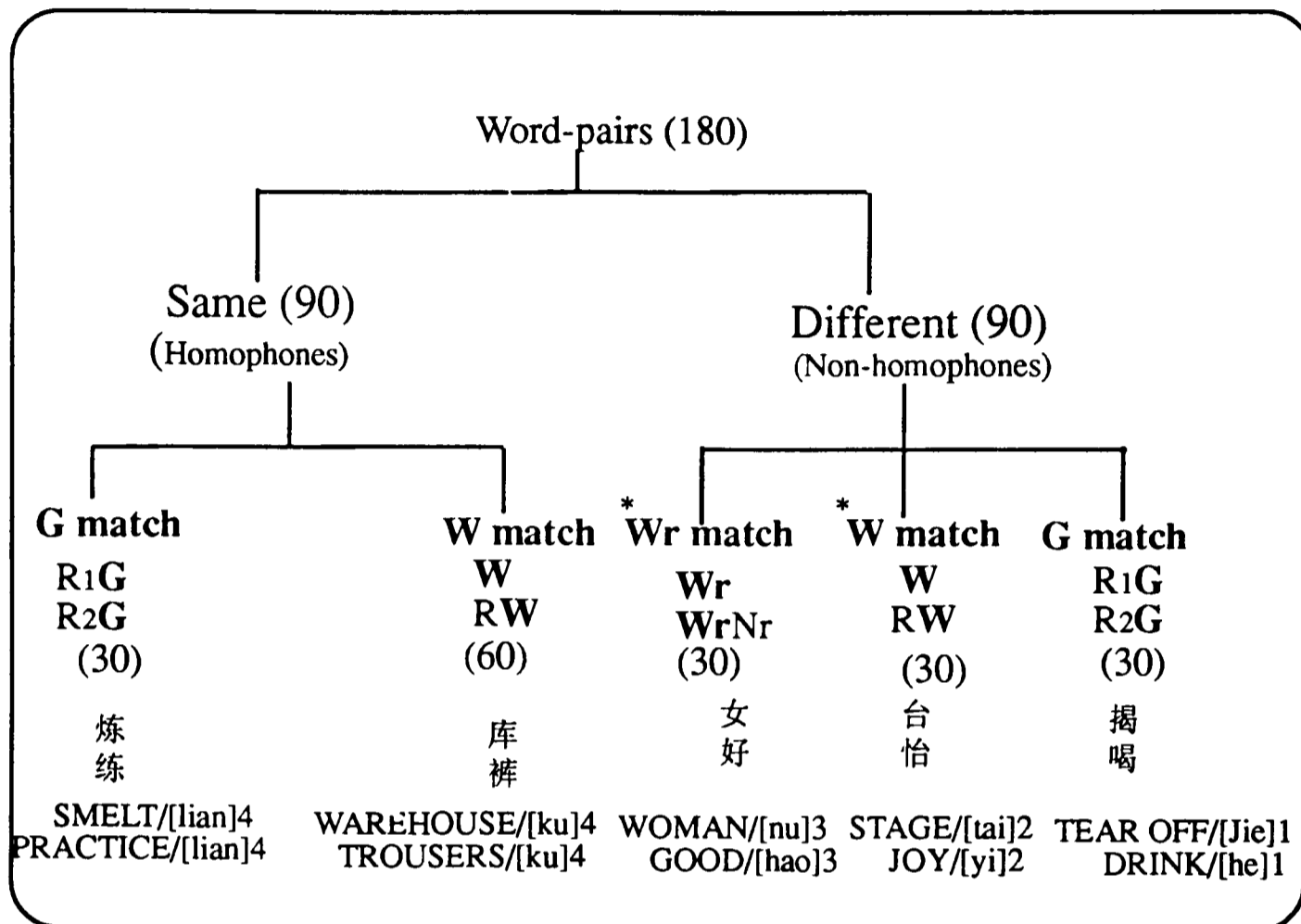
The Consistency Index denotes the "phonological incongruent consistency" between a component in a particular compound word and all other words that contain it in the lexicon. It is defined as the ratio of the total number of words in the dictionary, containing that component, whose whole-word pronunciation is consistent with the pronunciation of the particular compound word in question, divided by the type-frequency of the matching component. Note that the pronunciation of the compound words used in this experiment are always incongruent with the matching component in isolation. For example, 台 in isolation is the simple word [tai]2/ STAGE. It is the matching component of the stimulus-pair 台 怡 ([tai]2/ STAGE & [yi]2/ JOY). It appears in 17 words. The type frequency of 台 is 17. Five out of 17 of these words in the lexicon are pronounced [yi]2, (e.g., 贻 LEAVE BEHIND), which is not congruent with its non-radical component 台 [tai]2 in isolation, but consistent with the compound word 怡 [yi]2 in the same pair. The Consistency Index of 台 is thus 5/17.

Stimulus Conditions

There were 180 pairs of words used as stimuli in this experiment. Figure 6.2 illustrates the distribution of the 180 items in the three conditions for each judgment, with examples of each. The various stimulus conditions depend on the nature of the matching component in the stimulus-pair: i.e., which component were identical to the whole adjacent simple word in the same pair. There were three types of matching component, namely, lexical radical, non-radical and Graphic-Component match. Accordingly, the three conditions for 'Different' judgements were Word-Radical match, Word-Component match and Graphic-Component match. For 'Same' judgements, there were only two conditions, the Word-Component match and Graphic-Component match conditions. Word-Radical match pairs cannot be constructed for the 'Same' judgements, since lexical

Experiment 6

**Task: Does the pair of words have the Same or Different pronunciations ?
(Phonological comparison)**



* the same set of stimuli was also used in Experiment 9

Note: R: and Nr refer to lexical radical and non-radical component.

W and Wr refer to Word-component and Word-radical

Gr and G refer to Graphic-radical and Graphic-component

Bold typeface indicates the matching component in each stimulus pair

Figure 6.2. Stimulus conditions in Experiment 6

radicals (simple words) cannot have pronunciations congruent with the compound words in which they occur as lexical radicals. In addition, pairs of words that had no orthographic component in common (stimulus-pairs used in the last experiment) were treated as the baseline condition in contrast to these component matching conditions. Baseline data for the 'Same' judgments were taken from the Homophone 2 condition and for the 'Different' judgments were taken from the Control condition in the last experiment .

All stimuli in this experiment were presented randomly. The randomisation was controlled by a computer programme and was different for individual subjects. All stimulus-pairs used in this experiment consisted of two words which were matched on one component. The same matching component did not occur twice in the experiment. The two words in a pair cannot form a meaningful two-character CI in Chinese. The stimuli in each condition are described in detail as follows:

'Same' judgments (Homophone pairs)

Word-Component match: In homophone pairs, the non-radical component of the compound word was identical to the whole adjacent simple word in the same pair.

Graphic-Component match: In homophone pairs, two words share a common component, which is the non-radical component in both words. Unlike the Word-Component match, the matching graphic-component in isolation is unpronounceable in modern Chinese.

'Different' judgments (Non-homophone pairs)

Word-Radical match: In non-homophone pairs, the lexical radical of the compound word was identical to the whole adjacent simple word in the same pair.

Word-Component match: Constructed in the same way as Word-Component match for the 'Same' judgments, except that the two words have different pronunciations.

Note: In the Word-Component match condition, the simple word necessarily has a pronunciation incongruent with its paired compound word. Therefore the compound words used in this experiment were all ideocompound words, i.e., phonologically irregular words in Chinese. (See Chapter 2.) The matching components in this condition were the non-radical components of their compounds, and therefore *phonetic components* according to the Part-to-Whole phonological correspondence rule defined in this study. (See chapter 2.) However, they are semantic components by etymological rule (i.e., by the ideocompound rule)

Graphic-Component Match: Constructed in the same way as the Graphic-Component match for the 'Same' judgments, except that the two words had different pronunciations.

Baseline data: Baseline data for the 'Same' judgments were taken from those of homophone 2; and for 'Different' judgments were taken from those of the control items (non-homophone and non-synonym) in Experiment 5.

The apparatus, paradigm and procedures, and the statistical methods used in the present experiment were the same as described in the General Methods section.

Results

Table 6.2a presents the mean RTs averaged both by subjects and by items, with the statistical analyses. The corresponding data for mean error rates are shown in Table 6.2b.

1) 'Same' and 'Different' judgments

The RTs for the 'Same' judgments were generally faster than for the 'Different' judgments in both by-subject and by-item data. There was also a difference in RT between the different conditions. An ANOVA (2 x 3) and ACOVA (2 x 3) were performed respectively on by-subjects and by-item RT, with the two main factors

of judgements and conditions. The two co-variates examined in the ACOVA were the word frequencies of the two words in each pair. In the ANOVA, the two main effects of judgement and condition were both highly significant (*Judgement*: $F_{(1,11)}=16.4$, $P<0.001$, *Condition*: $F_{(1,11)}=17.8$, $P<0.001$). There was no interaction between judgment and condition (Table 6.2a). For the RT by-items analysis (ACOVA), adjusted for word frequency, significant effects of judgement and condition persisted (*Judgment*: $F_{(1,29)}=19.2$, $P<0.001$; *Condition*: $F_{(1,29)}=10.7$, $P<0.001$). Again there was no interaction between the main effects (Table 6.2a). Word frequency was not (either individually or combined) a significant co-variate in either main factors or in their interactions (Table 6.2a).

Further analyses were performed on 'Different' and 'Same' judgments separately. Comparisons between the different component matching conditions were examined by Newman-Keuls Q test. Comparisons between data from each component condition and the baseline data were assessed by Dunnett test respectively for both RT and error rates.

2) 'Different' judgments

The primary concern of this experiment, i.e., the contrast between Word-Radical match and Word-Component match conditions, was examined by further analyses on the 'Different' judgments. Figure 6.3 shows performance in the three component matching conditions on the 'Different' judgments; the figure also includes the baseline data taken from Experiment 5. The 'Different' RTs and error rate are shown separately as by-subjects and by-items data. As can be seen in Figure 6.3, there was an evident difference in RT between the two non-radical matching conditions (Word-Component match and Graphic-Component match) and the Word-Radical match condition. In the by-subjects analysis, the mean RT of 1247 msec for Word-Component match ($Q_{(2,3)}=4.30$, $P<0.01$), and 1325 msec for Graphic-Component match ($Q_{(3,3)}=4.76$, $P<0.01$) were both significantly

slower than that of 1143 msec for Word-Radical match. After adjustment for word frequency in the by-items analysis, the effect persisted. In the by-items analysis, mean RTs were 1303 msec. and 1368 msec. respectively for Word-Component match ($Q_{(2,3)}=3.74, P<0.01$) and Graphic-Component match ($Q_{(3,3)}=4.68, P<0.01$), compared with 1166 msec in the Word-Radical match condition. However, the differences in RT between the two non-radical matching conditions, i.e., Word-Component match and Graphic-Component match, were not significant in either by-subject or by-items analyses (Table 6.2a). The contrast between Word-Radical match and the baseline RTs (Figure.6.3) was not significant. Nor was the difference between Word-Component match and the baseline. However, the difference in RT between the baseline and the Graphic-Component match was significant (Dunnett test ($t_d=2.70, <0.05$)). The difference remained after adjustment for word frequency in by-items analysis ($t_d=2.66, P<0.05$).

On error rates, the pattern was broadly consistent with that found on RT (Figure 6.3). The false 'Same' response rates in the by-subjects data were 13.0% and 12.1% respectively for Word-Component match and Graphic-Component match, compared with only 3.5% for the Word-Radical match condition. For the by-items data, the corresponding figures were 26.1%, 25.3% and 7.4% respectively for Word-Component match, Graphic-Component match and Word-Radical match condition. These differences in the false 'Same' response rates between Word-Component match and Word-Radical match conditions were highly significant for both the by-subjects data ($Q_{(2,3)}=8.83, P<0.01$) and the by item data ($Q_{(2,3)}=4.62, P<0.01$). So were the differences between the Graphic-Component and the Word-Radical match conditions, both by-subjects ($Q_{(3,3)}=8.07, P<0.01$) and by items ($Q_{(3,3)}=5.21, P<0.01$).

Table 6.2a Mean reaction times (msec.) in Experiment 6 analysed by subjects (Ss) and by items, with the results of the statistical analyses

Judgment	Condition	No. of stimuli	By Subjects		By items		Comparison group	By Subjects	By items
			Mean	(SD)	Mean	(SD)		t or Q value§	t or Q value§
Same									
	(1)Baseline data	30	1038	(194)	1042	(117)	(1) vs (2)	0.06	0.42
	(2)Word-Component	60	1027	(195)	1054	(109)	(1) vs (3)	4.04	5.62**
	(3)Graphic-Component	30	1202	(313)	1257	(238)	(2) vs (3)	3.90**	5.74**
	All	120	1089	(234)	1117	(155)			
Different									
	(1)Baseline data	30	1201	(253)	1217	(122)	(1) vs (2)	1.00	1.53
	(2)Word-Component	30	1247	(349)	1303	(264)	(1) vs (3)	1.26	0.80
	(3)Word-Radical	30	1143	(249)	1166	(93)	(1) vs (4)	2.70*	2.66*
	(4)Graphic-Component	30	1325	(314)	1368	(313)	(2) vs (3)	4.30**	3.74**
							(2) vs (4)	1.61	1.16
							(3) vs (4)	4.76**	4.68**
	All	120	1229	(291)	1263	(198)			
Statistical analysis									
			ANOVA		ACOVA		Co-1†	Co-2†	Co-1&2
			F value	P value	F value	P value	F value	F value	F value
Main effects									
	Judgments (J)		16.4	0.00	19.2	0.00	1.45	1.34	0.85
	Conditions (C)		17.8	0.00	7.8	0.00	1.96	0.31	1.28
Interactions									
	J x C		1.5	0.24	1.2	0.30	1.93	0.15	1.05

§ Dunette t tests were used for comparisons against baseline data while Newman-Keuls Q tests were used for all other comparisons

† Co-1 and 2 refer to word frequency of the top and the bottom word respectively

*, ** P<0.05 and <0.01 respectively

Table 6.2b Mean error rates (%) in Experiment 6 analysed by subjects (Ss) and by items

Judgment	Condition	No. of stimuli	By Subjects		By items		Comparison	By Subjects	By items
			Mean	(SD)	Mean	(SD)		t or Q value§	t or Q value§
Same									
	(1)Baseline data	30	4.7	(3.7)	10.0	(10.6)	(1) vs (2)	0.99	1.60
	(2)Word-Component	60	4.2	(3.6)	8.4	(10.1)	(1) vs (3)	4.86*	3.52**
	(3)Graphic-Component	30	6.8	(3.9)	13.7	(14.9)	(2) vs (3)	4.34**	1.69
	All	120	5.2	(3.7)	10.7	(11.8)			
Different									
	(1)Baseline data	30	2.3	(2.3)	5.6	(5.5)	(1) vs (2)	7.37**	4.34**
	(2)Word-component	30	13.0	(6.5)	26.1	(26.3)	(1) vs (3)	0.95	0.41
	(3)Word-radical	30	3.5	(2.4)	7.4	(11.0)	(1) vs (4)	6.71**	4.16**
	(4)Graphic component	30	12.1	(6.4)	25.3	(20.6)	(2) vs (3)	8.83**	4.62**
							(2) vs (4)	1.44	0.21
							(3) vs (4)	8.07**	5.21**
	All	120	10.3	(4.4)	16.1	(15.8)			

§ Dunette t tests were used for comparisons against baseline data while Newman-Keuls Q tests were used for all other comparisons

*, ** P<0.05 and <0.01 respectively

'Different' Judgements: Component matching effects in phonological comparison (Experiment 6)

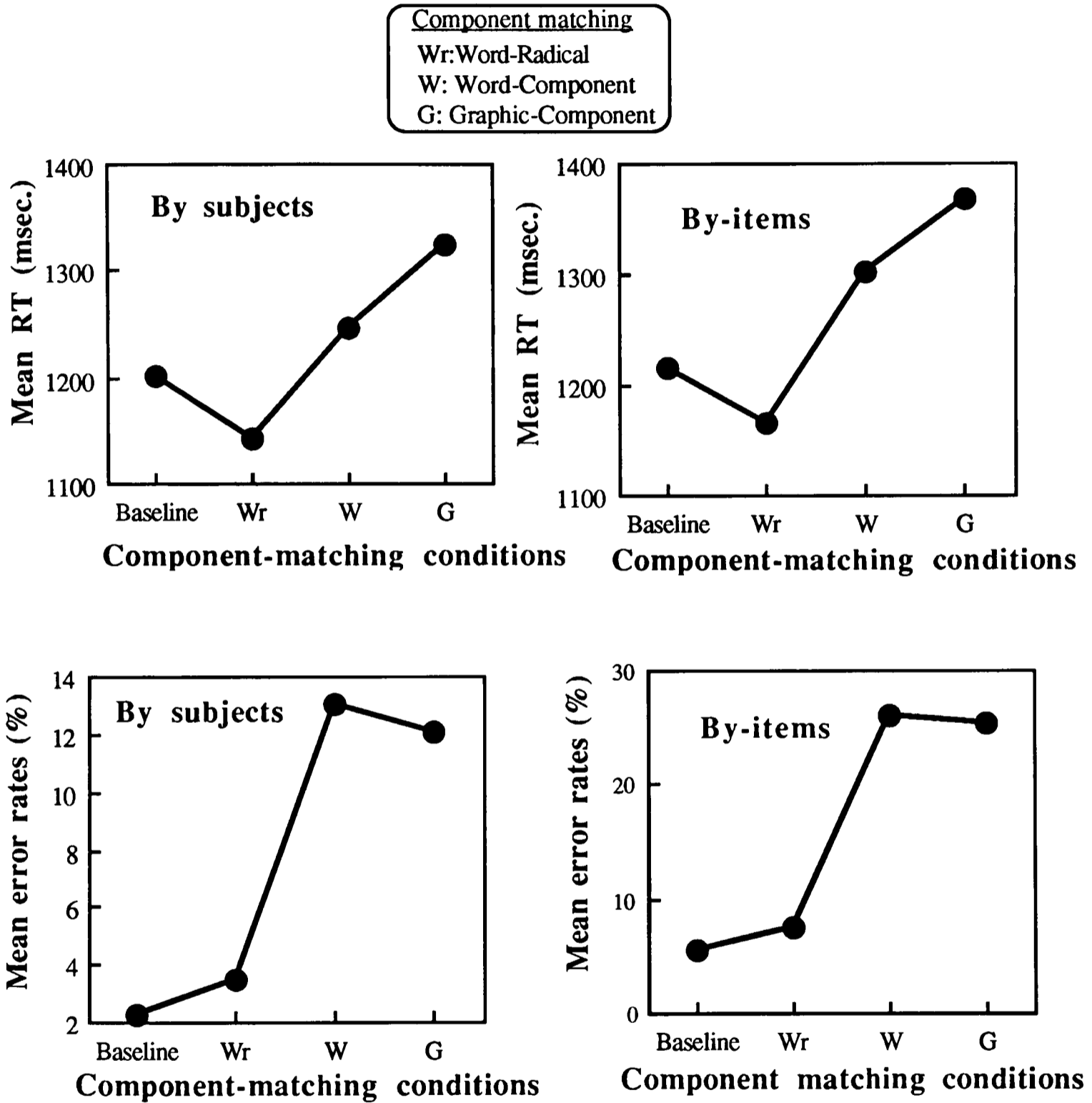


Figure 6.3 Mean RTs and error rates (by subject and by items)

In comparison with the baseline condition (2.3% by-subjects and 5.6% by-items), both of the non-radical matching conditions resulted in significantly higher error rates, both by-subject (*Word-component*: $t_d=7.37$, $P<0.01$, *Graphic-component*: $t_d=6.71$, $P<0.01$) and by-items (*Word-component*: $t_d=4.34$, $P<0.01$; *Graphic-component*: $t_d=4.16$, $P<0.01$). However, There was no significant difference between the baseline and Word-Radical match conditions.

3) 'Same' judgments: effects of different non-radical components

Further analyses were performed on the 'Same' judgments. The two conditions were Word-Component match and Graphic-Component match. As indicated in Table 6.2a, Word-Component match pairs were judged faster and more accurately than Graphic-Component match pairs, and the differences were significant both for RT ($t_{(11)}=3.90$, $P<0.01$) and error rates ($t_{(11)}=4.34$, $P<0.01$) by paired t tests. These results remained significant in the by-items analysis on RT ($t_{(29)}=5.74$, $P<0.01$) but not on error rates after taking word frequency into consideration (Table 6.2b).

Compared to the baseline condition, Word-Component match pairs did not differ significantly in either RT or error rate. In contrast, Graphic component pairs were consistently slower than the baseline (by subjects: $t_d=4.04$, $P<0.01$ and by-items: $t_d=5.62$ $P<0.01$), and had higher error rates than the baseline condition (by subjects: $t_d=4.86$, $P<0.01$ and by items analysis: $t_d=3.52$ $P<0.01$).

4) The indices of Radical frequency, Congruency and Consistency of component : effects on phonological processing

Table 6.3 lists the stimuli against the three types of frequency indices (Radical Frequency Index, Congruency Index and Consistency Index) as well as the word frequencies of the whole words (Simple word frequency (SWfreq.) and Compound word frequency (CWfreq.)) for the 30 Word-Radical match and 30 Word-Component match pairs. (The complete set of stimuli can be seen in

Table 6.3 Radical frequency index (RF) and congruency and consistency indices (Cong.I and Cons. I) and mean RTs and error rates in phonological and semantic comparison.

Stimuli	RF	Cong. I	Consis.I	SWfrq.	CWfrq.	Phonological		Semantic	
						RT	E%	RT	E%
1	.077	.417	.167	951	254	1.034	.5	1.056	0
2	.8	.071	.071	3639	1781	1.021	0	.942	0
3	.056	.647	.294	2063	564	1.258	.167	1.024	0
4	.077	.417	.417	2461	238	1.364	.5	1.067	.143
5	.037	.846	.154	2035	868	1.167	.667	1.048	.071
6	.1	.444	.333	3788	122	1.342	0	1.059	0
7	.125	.429	.286	32767	8838	1.142	.167	1.037	.071
8	.091	.7	.1	10702	1665	1.261	.282	1.119	.571
9	.1	.444	.111	5630	2439	1.375	.333	1.024	0
10	.2	.1	.25	8102	548	1.482	.167	1.058	0
11	.077	.1	.083	8932	337	1.142	0	.992	0
12	.083	.545	.182	9225	1615	1.099	.25	.992	0
13	.091	.9	.1	32767	741	1.148	.25	1.033	0
14	.1	.889	.222	8512	4375	1.434	.25	1.193	0
15	.167	.4	.8	11017	2174	1.276	.25	1.021	0
16	.2	.25	.1	105	277	1.323	.083	1.143	.143
17	.1	.333	.333	343	536	1.336	.667	1.022	0
18	.1	.111	.556	1759	2509	1.548	.167	1.010	.125
19	.125	.714	.286	155	1377	1.324	.167	1.241	0
20	.167	.1	.2	658	1980	1.243	.167	1.081	.143
21	.2	.1	.25	841	3960	2.418	.917	.954	0
22	.2	.5	.75	72	4795	1.216	.083	1.312	0
23	.143	.667	.333	1847	3291	1.488	.083	1.114	.143
24	.25	.1	.333	2378	5641	1.380	.25	1.077	0
25	.7	.111	.111	442	10857	1.151	0	1.026	0
26	.8	.071	.071	2063	6697	1.077	.083	1.001	0
27	.091	.5	.6	3573	3733	1.133	.083	1.043	.071
28	.056	.647	.059	3639	6675	1.200	.167	.950	.071
29	.111	.75	.25	8561	5127	1.049	0	1.052	.214
30	.1	.556	.111	39	2384	1.671	.417	1.118	.071
31	.7	.111	.111	442	376	1.181	0	1.099	.214
32	.1	.05	.05	525	304	1.123	.083	1.186	.143
33	.875	.161	.032	431	492	1.271	.167	1.061	0
34	.92	.125	.042	1278	498	1.174	0	1.008	.643
35	.929	.091	.018	2470	564	.988	0	.820	.929
36	.1	.091	.091	3097	66	1.152	0	1.113	0
37	.104	.106	.021	2245	1515	1.118	0	1.026	.071
38	.976	.006	.002	5840	714	1.231	0	.977	.286
39	.1	.091	.091	3595	2074	1.308	0	1.186	.429
40	.962	.08	.04	4618	1272	1.177	0	1.093	.214
41	.795	.026	.026	3484	2854	1.259	0	1.198	.071
42	.963	.038	.009	5127	3938	1.069	0	1.435	.429
43	.873	.114	.014	4712	4574	1.193	0	1.006	.071
44	.722	.114	.029	9098	619	1.091	.167	1.146	0
45	.1	.007	.007	11061	9629	1.111	0	1.368	.429
46	.947	.081	.027	133	570	1.144	.417	1.183	.714
47	.75	.429	.143	227	907	1.085	0	.942	0
48	.958	.087	.043	476	774	1.234	.333	1.321	.214
49	.917	.106	.021	2245	4336	1.039	0	1.010	0
50	.927	.1	.025	122	3180	1.243	.25	1.087	.286
51	.1	.014	.014	227	3815	1.054	.083	.971	0
52	.875	.161	.032	431	9828	1.181	.083	.970	0
53	.1	.026	.026	3484	4176	1.457	.083	1.094	0
54	.895	.167	.056	2727	10857	1.172	.083	.974	0
55	.968	.043	.011	310	8307	1.119	.083	1.008	0
56	.1	.014	.014	310	2273	1.181	.083	1.260	.071
57	.1	.026	.026	44	2124	1.280	.25	1.014	0
58	.963	.038	.009	3031	8561	1.091	.083	1.153	0
59	.98	.031	.01	111	21326	1.124	0	1.084	.643
60	.994	.011	.006	7035	32767	1.119	0	1.225	.071
Y	Regression	RF	Cong.I	Cons.I	SWF	CWF			
	RT	.116	.169	.077	.014	.015			
	R ²								
Phon	F(59)	7.613	11.766	4.873	.850	.867			
	P	.007	.001	0.031	.360	.355			
	E%	R ²	.190	.034	.001	.051			
	F(59)	14.860	13.646	2.06	.054	3.138			
	P	.000	.000	.157	.818	.082			
Sem.	RT	R ²	.029	.001	.000	.015			
	F(59)	1.759	1.287	.049	.017	.893			
	P	.190	.261	.826	.890	.348			
	E%	R ²	.100	.053	.003	.001			
	F(59)	6.542	3.825	3.252	.150	.038			
	P	.0138	0.060	.076	.700	.846			

Appendix I of this thesis.) Mean RT and error rates for the 60 items over the 12 subjects are also shown. Word frequencies of simple or compound words (SWfreq. and CWfreq.) were not significantly associated with the phonological comparisons, by simple regression analyses, either for RTs or for error rate (Table 6.3). By contrast, the three different types of frequency indices of the matching components (RFI, Congruency Index, and Consistency Index) individually were all significantly associated with RTs and error rates, by simple regression tests. The results of the regression analysis are presented in Table 6.3. As illustrated in Figure 6.4, there was an inverse association between the Radical Frequency Index of a component and both RT and error rate in the phonological comparison task. The lower the radical frequency, (i.e., the less frequently this component occurs as a lexical radical), the longer are the reaction times and the higher are the error rates for comparisons on pairs of stimuli containing that component as the matching element. For Congruency Index (Figure 6.5) and Consistency Index (Figure 6.6), the associations were both positive. The higher the index of congruency or consistency of a matching component, the longer are the reaction times and the higher are the error rates for the comparisons on the pairs of stimuli containing the component, indicating the influence of that component in the phonological task.

The three frequency indices of a component can be correlated with each other. The three indices of a component used in this experiment had significant correlations (*RFI & Congruency Index: $r^2=0.403$, $F_{(59)}=44.52$, $P<0.001$; RFI & Consistency: $r^2=0.292$, $F_{(59)}=27.27$, $P<0.001$; Consistency Index & Congruency Index: $r^2=0.104$, $F_{(59)}=7.66$, $P<0.01$).*

In order to examine the independent effect of each of these three indices of the matching components on phonological processing, a multiple regression test was performed with RT or error rates as the dependent variable, and the three frequency indices and the word frequencies of the two words as 5 independent variables. The most important independent variable was found to be the

**Radical frequency index: effect in a phonological judgment task
(Experiment 6)**

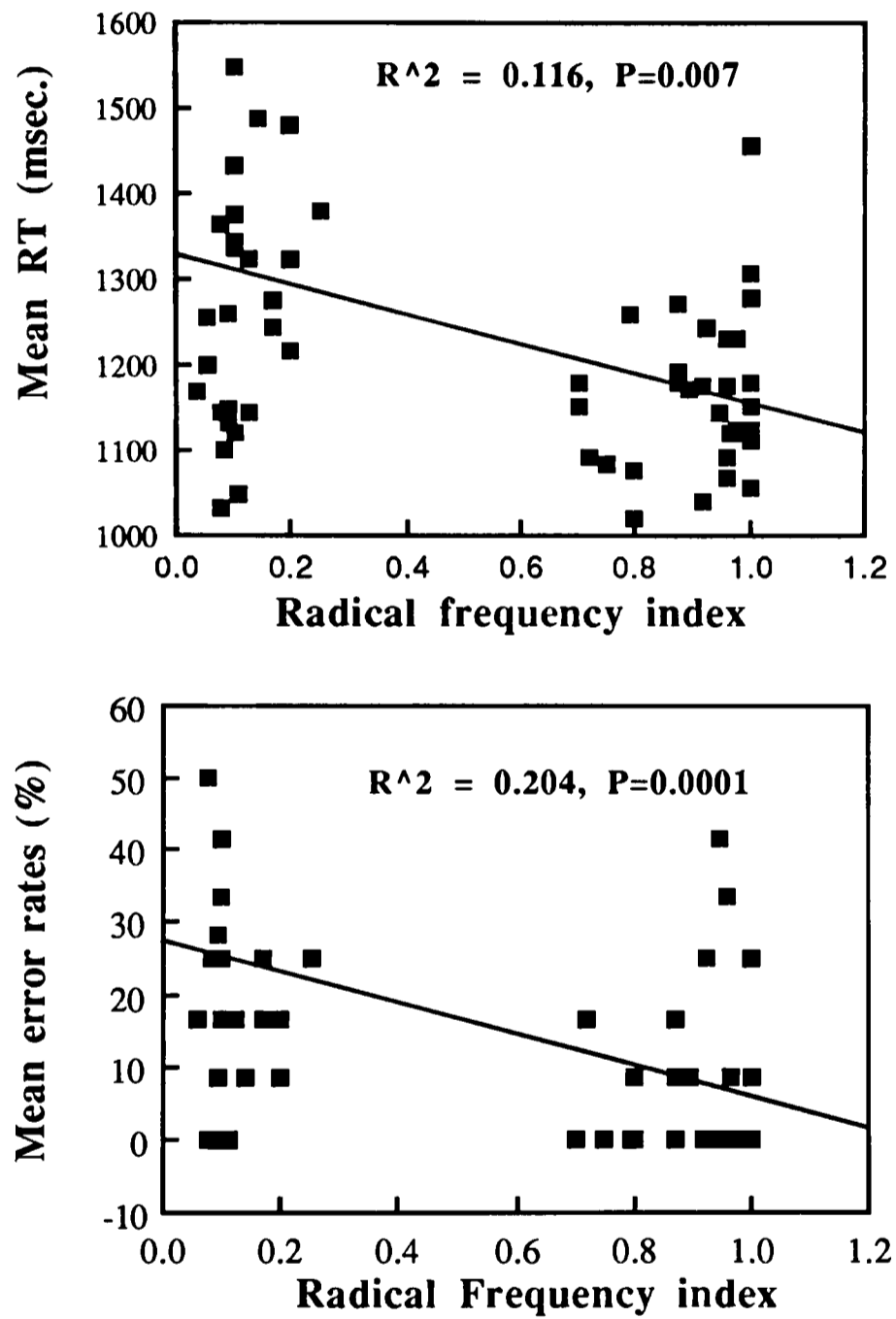


Figure 6.4 Correlation between mean RTs and error rates and radical frequency index

**Congruency index: effect in a phonological judgment task
(Experiment 6)**

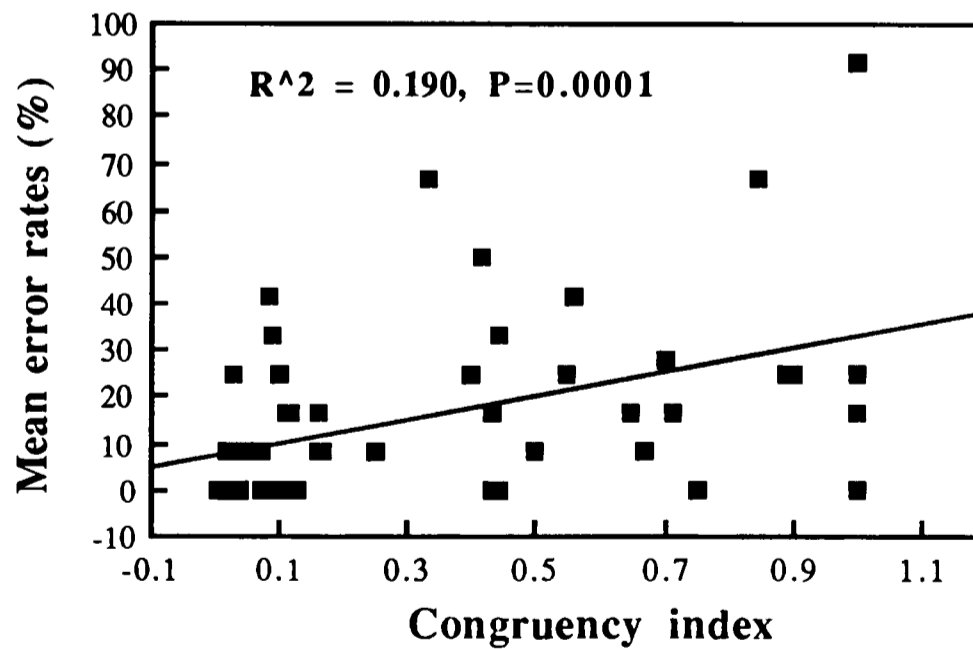
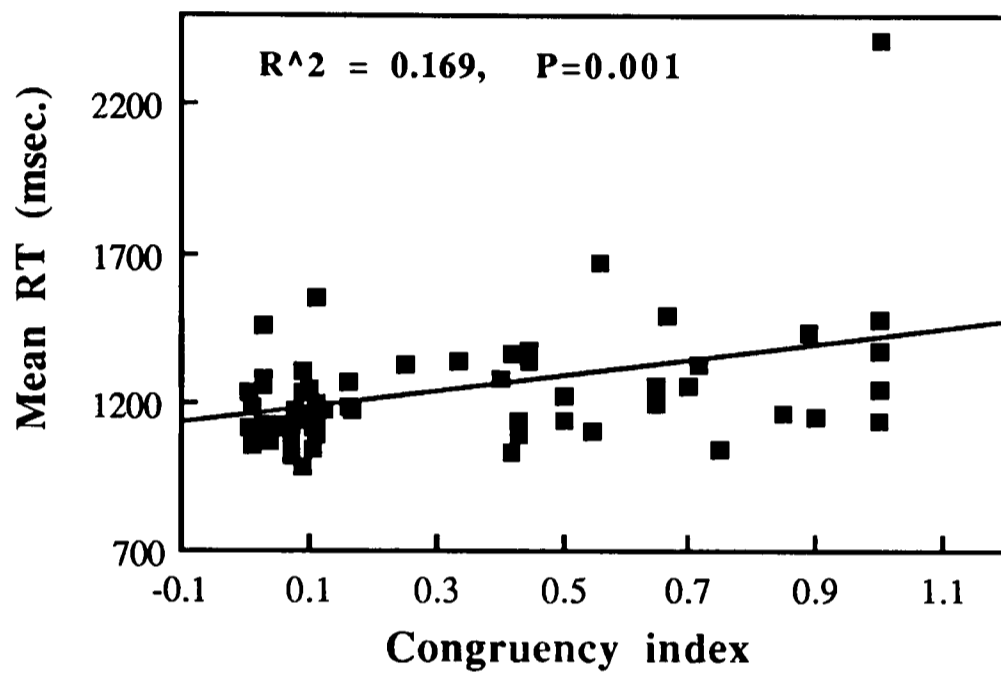


Figure 6.5 The associations between mean RTs and error rates and congruency index

Consistency index: effect in a phonological judgment task
(Experiment 6)

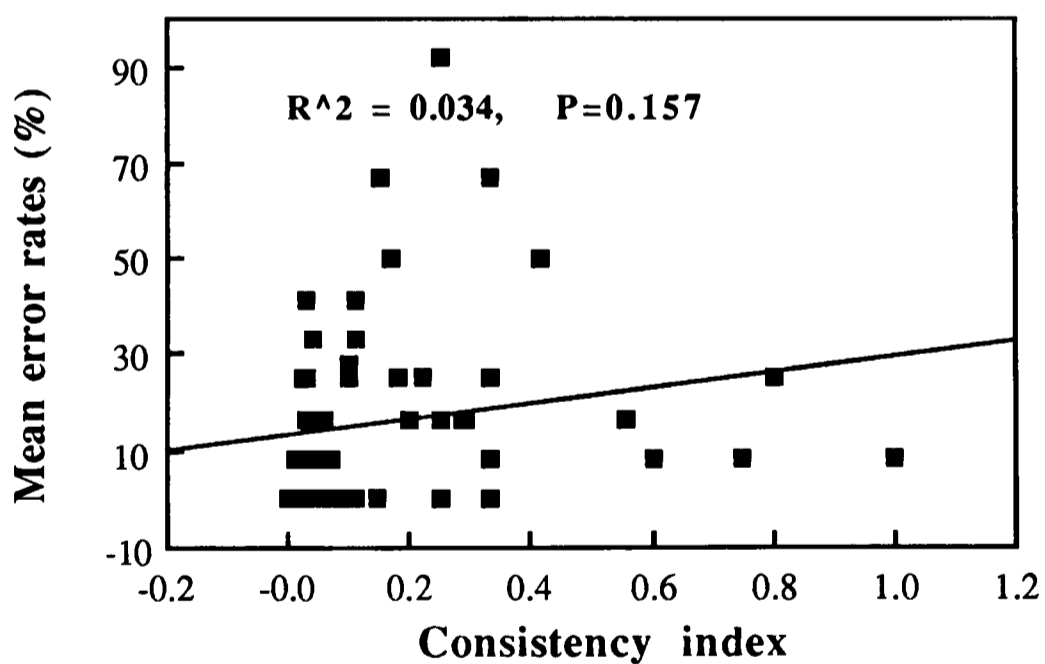
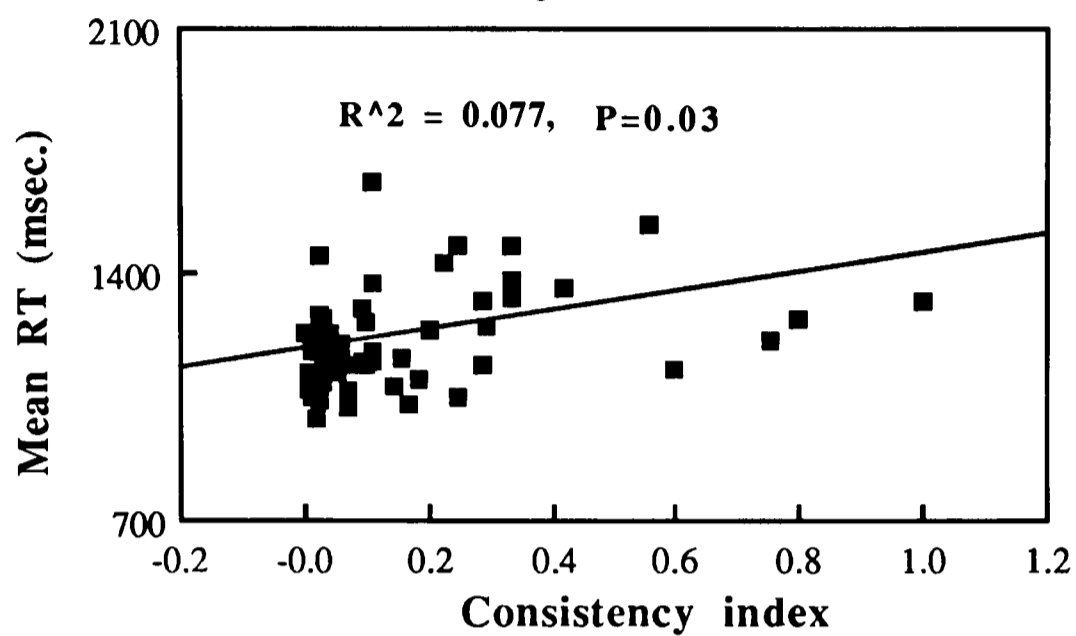


Figure 6.6 The correlation between mean RTs and error rates and consistency index

Congruency Index, i.e., phonological regularity in Chinese. After adjusting for RFI, congruency and consistency indices and for the word frequencies of both words in a pair, only the Congruency Index remained significantly correlated with RT in the phonological comparisons. None of these three factors had a significant effect on error rates, after adjusting for each other and for word frequencies.

Discussion

This experiment has demonstrated that there is selective attentional bias toward the non-radical component of compound words in a phonological judgment task. This was shown by the fact that word-pairs with a matching non-radical component had significant effect in biasing the subjects toward false 'Same' responses in the phonological comparison task. By contrast, a matching radical component as a distractor had no significant effect in the phonological comparison task. As discussed in the introduction of Chapter 5 and in this chapter, the non-radical component may function as the regular phonetic component, in accordance with the Part-to-Whole Phonological Correspondence rule introduced in this study. The benefits and costs of lexical decomposition indicated by the Word-Component match conditions in 'Same' and 'Different' judgments were inevitably confounded with the judgment effect. Nonetheless, the highly significant, fast 'Same' responses toward a matching component having a pronunciation congruent with the compound word in which it occurs may be at least partly attributed to the beneficial effect of phonological regularity (i.e., congruency in pronunciation between the non-radical component and the word that contains it).

As already noted (See Stimulus Condition), the compound words in the Word-Component match condition always have pronunciations incongruent with their own non-radical component. These compound words were all (necessarily) ideocompound words, in which the non-radical components are *semantic*

components according to ideocompound word-formation rule. However, the impact of these non-radical components as distractors in a *phonological* task suggests that the non-radical component is treated initially as the phonetic component, according to the *orthographic* rule, regardless of etymology rule. Skilled readers of Chinese appear to treat the non-radical components of Chinese words (ZI) as the phonetic token no matter whether in etymology they are semantic or phonetic components. This further supports the view that it is the regular orthographic structure, rather than the etymology rules, that is relevant to reading.

Furthermore, the results of this experiment may be taken as favouring the connectionist view of rules as connections. In current connectionist accounts, the implementation of any rule is through modifying connection weights, and the influence of each of the three component-frequency indices (RFI, Congruency and Consistency) may perhaps be thought as reflecting statistical "rules" implemented in this way. All three component frequency indices individually appeared to show a significant influence on the phonological comparison task. After adjustment for each other, however, only the Congruency Index remained significant. This result suggests that phonological recoding in Chinese (after lexical decomposition) is dominated by "derived" phonology, rather than by an "analogy" procedure.

Experiment 7: Phonological comparisons of word-pseudoword pairs

Introduction

This experiment was designed to examine the critical Word-Radical match and Word-Component match conditions with word-pseudoword pairs. This has been already investigated in Experiment 4. In addition to the pseudocompound condition (as used in Experiment 4), two further conditions were included in Experiment 7. In these conditions, graphic-components and graphic-radicals, which cannot be used as words in isolation in contemporary Chinese, were also used in isolation as pseudowords (termed "pseudosimples").

Although a pseudocompound in the word-pseudocompound pairs virtually has no correct pronunciation in Chinese, subjects may 'guess' the pronunciation of the pseudocompound *by derivation* from its non-radical component, which can also be a word in isolation. For instance, in the word-pseudocompound pair 交 & 𠂇] subjects may judge the pronunciation of the pseudocompound (e.g., 𠂇]) by derivation from its component (e.g., 交), which in isolation is identical to the real word 交 in the same pair. If, however, the pseudoword is a graphic-radical or a graphic-component in isolation (i.e., an unpronounceable pseudosimple), the pronunciation of the pseudosimple has to be achieved *by analogy* from a real word in which the pseudosimple is a constituent component. For example, 𠂇 & 报, subjects may 'guess' the pronunciation of 𠂇 by analogy from 报 [baɔ]4 . The critical point here is whether such "derivation" for word-pseudocompound pairs and "analogy" for pseudosimple-word pairs (inferred from 'Same' response rates) are different between stimulus-pairs matching on the lexical radical and those matching on the non-radical component.

Stimulus Conditions

A total of 60 word-pseudoword pairs were included as stimuli. The stimulus conditions are illustrated in Figure 6.7. There were 30 pseudosimple pairs and 30 pseudocompound pairs. In each type of pseudoword, stimuli were divided equally into two conditions according to the nature of the matching component (radical or non-radical).

Pseudosimples

A pseudosimple is a Graphic-Radical or a Graphic-Component in isolation, which cannot be used as a word in isolation in Chinese. These were two conditions for pseudosimples:

Graphic-Radical match: The Graphic-Radical match in isolation as a pseudosimple is identical to the lexical radical of the compound word in the same pair.

Graphic-Component match: The Graphic-Component match in isolation as a pseudosimple is identical to the non-radical component of the compound word in the same pair.

Pseudocompounds

Pseudocompounds were constructed by a Word-Radical and a Word-Component in a nonsense combination in Chinese. The lexical radicals of pseudocompounds were always in their legitimate position. The two conditions for pseudocompounds are further described as follows:

Word-Radical match: The lexical radical of a pseudocompound was identical to the simple word in the same pair.

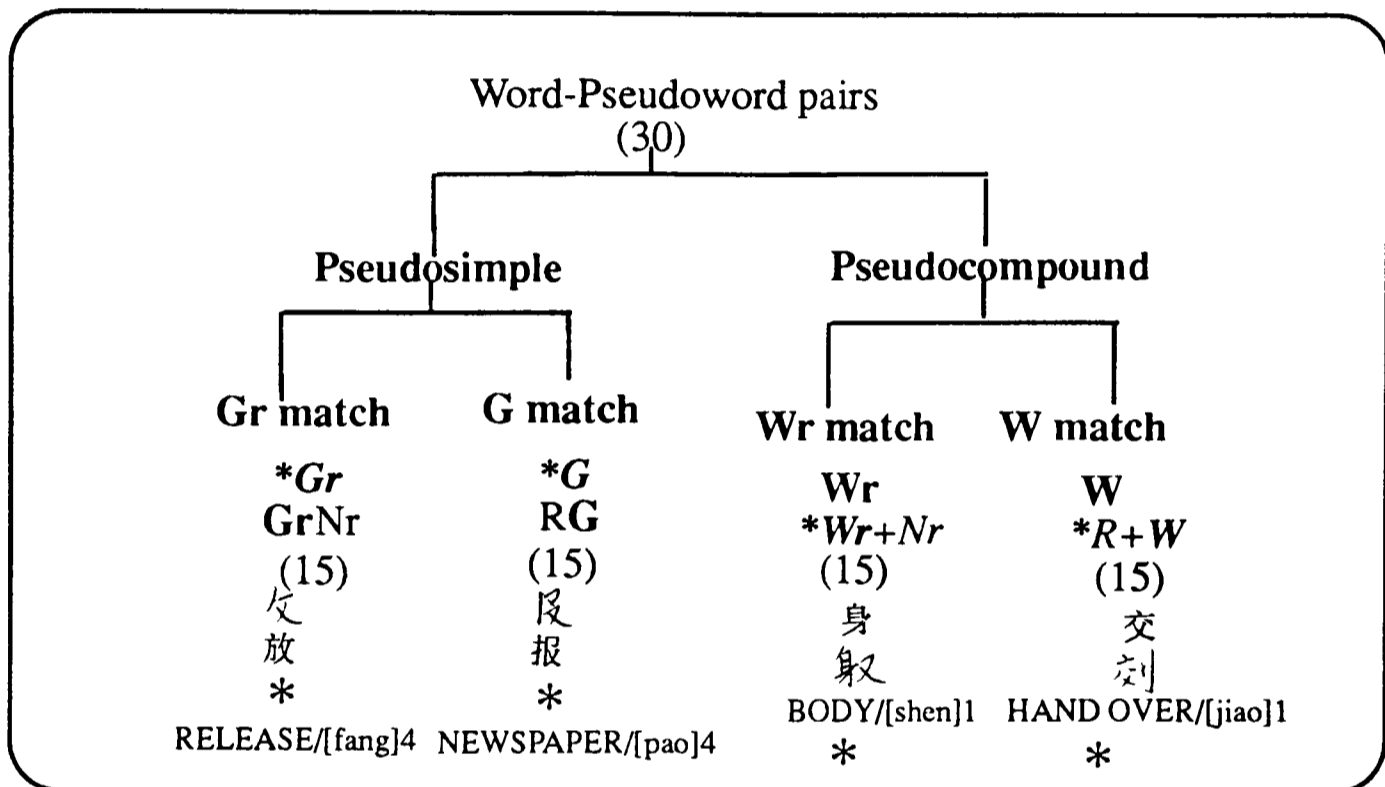
Word-Component match: The non-radical component of a pseudocompound was identical to the simple word in the same pair.

The stimuli were presented in random sequence. The experimental procedures and statistical methods used were the same as those described in the General Methods section.

Experiment 7 & 10

Tasks: Ex 7: Does the pair of words have 'Same' or 'Different' pronunciations?
(Phonological comparison)

Ex 10: Does the pair of words have 'Same' or 'Different' meanings?
(Semantic comparison)



*Note: R: and Nr refer to lexical radical and non-radical component.
W and Wr refer to Word-component and Word-radical
Gr and G refer to Graphic-radical and Graphic-component*

Bold typeface indicates the matching component in each stimulus pair

**: Pseudoword*

Figure 6.7 Stimulus conditions in Experiment 7 and 10

Prior to the experiment, the subjects were informed that there were unknown words (pseudowords) in the stimuli. They were encouraged to guess the sound of those unknown words just as they would do in reading without a dictionary to hand.

Results

Given that no correct response can be defined for pseudowords (pseudowords have no correct pronunciation in Chinese), only the 'Same' response rate was used for the analysis.

Mean 'Same' response rates for both by-subjects and by-items analyses are presented in Table 6.4a, with the corresponding RT data given in Figure 6.4b. In both pseudocompounds and pseudosimples, the 'Same' response rates were generally higher for the non-radical matching conditions than for the radical matching condition.

1) Effects of the matching component (radical vs non-radical)

ANOVA (2 x 2) and ACOVA (2 x 2) were performed respectively on the by-subjects and the by-items error rate data. The main factors in both ANOVA and ACOVA were types of pseudowords (pseudosimples and pseudocompound) and conditions (matching on the radical vs non-radical components). In the ANOVA, both main factors were highly significant (*Pseudoword type*: $F_{(1,11)}=9.2$, $P=0.01$, *Condition*: $F_{(1,11)}=26.0$, $P<0.001$). There was also a significant interaction between pseudoword type and condition ($F_{(1,22)}=15.4$, $P<0.001$). In the by-items analysis (ACOVA), the condition effect remained significant ($F_{(1,11)}=57.1$, $P<0,001$), but the effect of pseudoword-type and the interaction between pseudoword type and condition (P x C) became non-significant (*pseudoword type*: $F_{(1,29)}=0.30$, $P=0.1$; *P x C*: $F_{(1,58)}=1.82$, $P=0.2$) after taking the word frequency of the real word in each pair into account.

Table 6.4a Mean Same response rates (%) both by subjects and by-items, in Experiment 7, with results of statistical analyses

Pseudoword type	Condition	No. of stimuli	By Subjects		By items		By Subjects	By items
			Mean	(SD)	Mean	(SD)	(1) vs (2)	(1) vs (2)
Pseudosimple								
(1) Graphic-Radical		15	4.9	(5.0)	13.8	(13.4)	t ₍₁₁₎ =5.86 P=0.00	t ₍₁₄₎ =3.73 P=0.00
(2) Graphic-Component		15	30.8	(9.9)	44.2	(21.2)		
All		30	17.8	(7.4)	29.0	(17.3)		
Pseudocompound								
(1) Word-Radical		15	8.0	(5.7)	13.8	(12.8)	t ₍₁₁₎ =4.10 P=0.00	t ₍₁₄₎ =5.44 P=0.00
(2) Word-Component		15	21.3	(10.2)	62.4	(22.1)		
All		30	14.6	(7.9)	38.1	(17.4)		
			ANOVA		ACOVA		Co-variate †	
Main effects			F value	P value	F value	P value	F value	P value
Pseudoword-type (P)			9.2	0.01	0.30	0.10	2.46	0.14
Conditions (C)			26.0	0.00	57.1	0.00	0.54	0.47
Interactions								
P x C			15.4	0.00	1.8	0.20		

† Co-variate refers to word frequency of the real word in word-pseudoword pairs

Table 6.4b. Mean reaction times (msec.) both by subjects (By Ss) and by-items in Experiment 7 along with results of statistical analyses

Pseudoword type	Condition	No. of stimuli	By Subjects		By items		By Subjects	By items
			Mean	(SD)	Mean	(SD)	(1) vs (2)	(1) vs (2)
Pseudosimples								
(1) Graphic-Radical		15	1536	(917)	1452	(342)	t ₍₁₁₎ =0.07 P=0.94	t ₍₁₄₎ =0.19 P=0.84
(2) Graphic-Component		15	1544	(828)	1433	(193)		
All		30	1540	(872)	1442	(267)		
Pseudocompounds								
(1) Word-Radical		15	1422	(580)	1758	(348)	t ₍₁₁₎ =0.21 P=0.83	t ₍₁₄₎ =0.12 P=0.90
(2) Word-Component		15	1403	(586)	1528	(251)		
All		30	1412	(583)	1643	(299)		

In Table 6.4, the 'Same' response rates were further compared between the two different conditions for each of the two types of pseudowords. For pseudocompounds there was a significant difference in the 'Same' response rates between Word-Radical match and Word-Component match (*by subjects*: $t_{(11)}=4.10$, $P<0.001$; *by items*: $t_{(14)}=5.44$, $P<0.001$). Likewise for pseudosimples, there was a significant difference in the 'Same' response rates between the Graphic-Radical match and the Graphic-Component match (*by subjects*: $t_{(11)}=5.86$, $P<0.001$; *by items*: $t_{(14)}=3.73$, $P<0.001$).

Discussion

The present experiment shows that the phonological comparison of word-pseudoword pairs, like that of real word-pairs, is influenced by the nature of the matching component, i.e., the distractor. The 'Same' response rates toward the Graphic-Component match condition were significantly higher than in the Graphic-Radical match condition. This indicates that pronunciation of a 'pseudosimple', achieved by analogy from the real word containing that same graphic element, is more likely to be elicited by word-pseudoword pairs with a shared non-radical component, than with a shared lexical radical. This further supports the idea that skilled Chinese readers may have a selective attentional bias toward the non-radical component, the rule-defined phonetic component, in a phonological judgment task, irrespective of the lexical status of the radical component. Given that the matching graphic-radical and graphic-component (i.e., the distractors), are both unpronounceable and meaningless in isolation, the significant influence of the graphic-component suggests that, like "derived pronunciation", the use of "analogy" for lexical decomposition in pronunciation is similarly rule-governed. Both procedures are based on the rule-defined phonetic component—the non-radical component.

General Discussion

In this general discussion, the findings of the three experiments are discussed in the context of a possible functional architecture of phonological processing for Chinese words, compared with current theories of phonological recoding of written words in English.

1) a Direct route from script to sound

The first experiment on phonological comparison demonstrates the independence of phonological recoding from semantic mediation in Chinese. Pairs of Chinese words which are different in pronunciation but similar in meanings such as 和 and 合, were not significantly more difficult to categorize as "Different" in terms of pronunciation than pairs of words which were different in both pronunciation and meaning (Experiment 5). This indicates that phonological comparison may be based on phonological recoding without necessarily involving any semantic recoding. This finding therefore challenges the traditional assumption that the sound of supposedly ideographic or logographic Chinese words may not be assigned without first consulting the semantics. More importantly, the demonstration of a direct mapping from orthography to phonology in Chinese orthography may have an interesting theoretical implication. Although Chinese orthography uses whole words as syllabic 'sound symbols' (in that the non-radical component can generally also be used as a word in isolation), this system may not be so fundamentally different from the abstract sound-symbol in alphabetic script, i.e., the letters of English words. As with letters, so in Chinese, the sound of the Word-Component can be assigned directly from the orthography without reference to meaning.

2) Rule-governed lexical decomposition-*the phonetic component*

The most important finding of the present study on phonological recoding is that the pronunciation of over 90% of Chinese words may be achieved via a rule-governed lexical decomposition process. Studies on so-called 'shallow' orthographies regarding orthographic-phonological correspondences (e.g., Japanese Kana (*Besner & Hildebrandt, 1987*) and Serbo-Croatian (*Katz and Feldman, 1983; Seidenberg & Vidanovic, 198bc, Besner, 1987, 1991*) have suggested that the nature of the relationship between orthography and phonology can be independent of the levels of phonological transcoding, i.e., via words as wholes or via components (through lexical decomposition). The present study on a 'deep' orthography (Chinese) further supports this view that lexical decomposition may be regarded as a general feature of phonological processing of written words.

Many studies on English have accepted that the word-ending of monosyllabic English words, sometimes referred to as the orthographic 'body', plays an important part in the phonological processing of English words (e.g., *Brown, 1987; Henderson, 1982; Parkin, 1982; Patterson & Coltheart, 1987; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Treiman & Chafetz, 1987; Kay, 1986; Kay & Bishop, 1987*). These word-endings in English refer to the middle vowel and following consonants, a monosyllabic structure. Other subword unit, the so-called Basic Orthographic Syllabic Structure (BOSS) has been also proposed for lexical decomposition in English (*Taft, 1979, 1986, 1987; Taft and Forster, 1976*). However, Seidenberg & McClelland (1989) take a different approach in their recent connectionist model for naming English words. They asserted that the pronunciation of English words was unlikely to be related to an unique orthographic unit, for English orthography is a so-called quasi-irregular system, which encodes semantic and phonological information simultaneously in the words (*Chomsky, 1968*).

The debate over the existence of unique orthographic units relevant to the pronunciation of English words still continues. However, the present study has demonstrated that in skilled readers of Chinese there is a selective attentional bias toward the non-radical component (the rule-defined phonetic component) in a phonological judgment task. This was shown by the finding that a physical match between members of a stimulus-pair on the non-radical component acted as a distractor toward a false 'Same' response, exerting a significant effect on both speed and accuracy of phonological judgments. However, a similar physical match on the lexical radical had no effect on the phonological judgment. This may possibly be an orthography-specific phenomenon in Chinese, since in Chinese orthography semantic and phonological information are encoded independently, in two distinct orthographic components—the lexical radical and the non-radical component, respectively. Alternatively, it may be a rather general phenomenon. That is, a selective attentional bias may naturally tend toward segmented monosyllabic structures in the phonological processing of written words (such as the monosyllabic word-ending in English and the monosyllabic Word-Component in Chinese).

The process of phonological recoding via orthographic components in Chinese can be achieved either (1) by "derivation", based on phonological congruency between a word and its own non-radical component in isolation (for phonologically regular words) or (2) by "analogy", based on consistency among words with the same non-radical component. Both procedures are applied to the non-radical component—the phonetic component—as defined by the lexical decomposition rule (i.e., Part-to-Whole phonological correspondence rule in Chapter 2). The operation of an analogy procedure was inferred from 'Same' response rates to word-pseudoword pairs with matching graphic-radical and graphic-component. Both types of element in isolation, called "pseudosimples", are unpronounceable and meaningless. Therefore 'Same' responses to these word-

pseudoword pairs must be due to the matching component (i.e., physical distractor) cueing the analogy in pronunciation from the real word in the pair. Consistent with the Word-Component condition, the Graphic-Component match caused significantly higher false 'Same' responses than the Graphic-radical match in these word-pseudoword pairs (Experiment 7). This result suggests that the selective attentional bias toward the non-radical component in phonological recoding of Chinese words is unlikely to be affected by the lexical status of the non-radical component. Furthermore, the influence of the matching component as a distractor was shown to be a function of the different type-frequency indices of the matching component. The type-frequency indices of a component (radical frequency index, congruency index and consistency index) were measured according to different concordance ratios between the component and all the words (in the dictionary) in which that component occurs. Individually, all three indices including radical frequency, (phonological) Congruency and Consistency were significantly associated with the speed and accuracy of phonological comparison. After appropriate controls, the Congruency Index, but not the Consistency Index was significantly associated with the comparisons. This indicates that lexical decomposition and phonological recoding in Chinese may be based mainly on the phonological Congruency between a word and its own non-radical component. Derivation is thus the major procedure for pronouncing Chinese compound words, following lexical decomposition.

Conclusions

To sum up, script to the sound transcoding in Chinese can be obtained independent of semantic mediation. This process may be qualitatively similar to that found in English. Moreover, there is a selective attentional bias toward the non-radical component—the phonetic component as defined by the lexical decomposition rule called Part-to-Whole Phonological Correspondence. Hence, pronunciation of Chinese words is based on orthographic units smaller than the

word as a whole, at least for large majority Chinese words, namely the compound words. Lexical decomposition does exist in phonological recoding of Chinese words. Such lexical decomposition is also rule-governed and directed to the non-radical component of a word. In contrast to "assembled phonology" in English, the procedure for pronunciation using lexical decomposition in Chinese is "derivation". As in English, the pronunciation of Chinese words can also be achieved via analogy from other words with the same non-radical components. These two procedures for pronunciation of Chinese words via lexical decomposition are based on the same lexical decomposition rule. The speed and accuracy of both "Derivation" and "Analogy" via the non-radical component also depends on the type-frequency (the relative weights) of the correspondences between that component and all other words in the lexicon that contain it.

Chapter 7

From script to meaning in Chinese

Theoretical issues

This chapter reports three experiments on semantic comparisons of pairs of Chinese words. These three experiments correspond respectively to the three experiments reported in the last chapter. Two main issues are addressed in this study:

- 1) What is the role of phonological recoding in the semantic processing of Chinese ?*
- 2) Does the lexical radical, the rule-defined semantic component, receive attentional priority in the semantic processing of Chinese words?*

To address these two issues, three experiments were conducted. With the same subjects and in part the same stimuli, these three experiments on semantic comparisons correspond to the three experiments on phonological processing (Experiments 5, 6 and 7 respectively) reported in the preceding chapter.

Experiment 8: The semantic comparison of pairs of Chinese words

Introduction

This experiment was designed to address the first issue: What is the role of phonological recoding in the semantic processing of Chinese. Analogous to Experiment 5, this question was investigated by examining the homophone effect in semantic comparisons (Experiment 8). The rationale for using homophones in this experiment is similar to the rationale for using synonyms in Experiment 5. If phonological recoding plays any role in semantic judgments, homophone pairs in which the two words have 'Different' meanings but the

'Same' pronunciation (e.g., ICE [bin] 1, SOLDIER [bin]1) should be judged more slowly and/or provoke more false 'Same' responses than the control pairs in which the two words are different in both pronunciation and meaning. On the contrary, the absence of homophone effects would suggest that prior phonological recoding is not necessary for the semantic recoding of Chinese words.

Stimulus Conditions

There were 120 pairs of stimuli used in this experiment. The distribution of all stimuli in the four conditions is illustrated in Figure 7.1, with corresponding examples. There were equal numbers of 'Same' (synonym) and 'Different' (non-synonym) pairs. There were two conditions for the 'Different' judgments, (i.e., control and homophones) with 30 pairs in each condition. For the 'Same' judgments, there was only one condition, i.e., synonyms. For the purpose of analysis, the 60 synonyms pairs were divided into two groups of equal size, in order to match the two conditions to the 'Different' judgment. The stimuli for each condition are described further as follows:

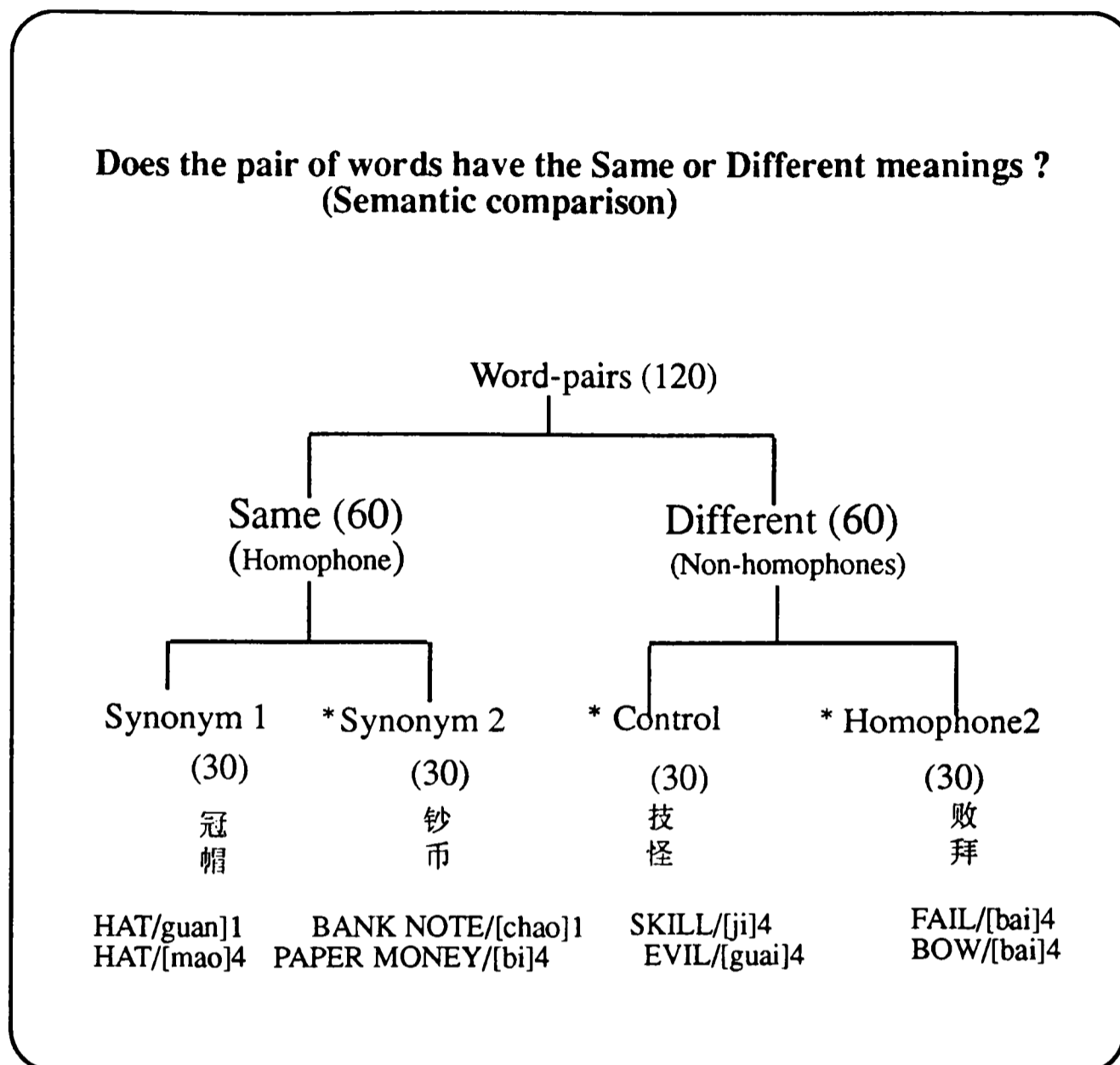
'Same' judgments (Synonym pairs)

Synonym: In this condition, pairs of different words have similar meanings. The actual stimuli used in Synonym 2 were identical to those for the 'Different' judgments of the phonological comparison task in Experiment 5, in the synonym condition.

'Different' judgements (Non-homophone & non-synonym pairs)

Control condition: In this condition, pairs of words have different meanings and pronunciations. The actual stimuli were identical to those used for the 'Different' judgments of the phonological comparison task in Experiment 5, in the control condition.

Experiment 8



** the same set of stimuli was also used in Experiment 5*

Figure 7.1. Stimulus conditions in Experiment 8

Homophone condition: In this condition, pairs of different words have 'Different' meanings but the 'Same' pronunciation, including tone. The stimuli were identical to the second group of stimuli used in the homophone condition for the 'Same' judgments of the phonological comparison task, in Experiment 5.

The task in the following three experiments is to compare meaning rather than pronunciation. The paradigm and procedure were as described in Chapter 5 (General Method).

Results

Mean RTs (both by subjects and by items) are presented in Table 7.1a; The corresponding data for error rates are presented in Table 7.1b. In general, 'Same' (synonym) judgments appear faster than 'Different' (non-synonym) judgments. ANOVA and ACOVA were conducted on the by-subject and by-item data respectively. The two co-variates in the ACOVA were the word frequencies of both words in each pair. The main effects examined were judgment and condition. In the ANOVA, the main effect of condition was significant ($F_{(1,11)}=6.2, P<0.05$) but not that of judgment. There was no significant interaction between judgment and condition. In the ACOVA (with adjustment for word frequency), the main factor of judgment was highly significant while the main factor of condition was only marginally significant (*Judgment:* $F_{(1,29)}=7.6, P=0.01$; *Condition:* $F_{(1,29)}=4.4, P=0.05$). However, the interaction between condition and judgment was again not significant (Table 7.1a). The co-variates individually or together did not significantly affect either the main factors or their interactions (Table 7.1a).

1) Homophone effect on 'Different' judgments

The condition effect was further examined on the 'Different' judgements only. There was no significant difference in RT between homophones and control words by paired t test ($t_{(11)}=1.45, P>0.05$). This was further confirmed

Table 7.1a Mean reaction times (msec.) in Experiment 8, analysed by-subjects and by-items, along with the results of the statistical analyses

Judgment	Condition	No. of stimuli	By Subjects		By items		By Subjects	By items																																																							
			Mean	(SD)	Mean	(SD)	(1) vs (2)	(1) vs (2)																																																							
Same																																																															
	Group 1	30	1097	(164)	1123	(154)																																																									
	Group 2	30	1039	(141)	1076	(91)																																																									
	All	60	1068	(143)	1049	(122)																																																									
Different																																																															
	(1) Control	30	1133	(238)	1159	(98)	$t_{(11)}=1.45$ P=0.17	$t_{(29)}=1.18$ P=0.25																																																							
	(2) Homophone	30	1092	(175)	1128	(94)																																																									
	All	60	1112	(206)	1143	(96)																																																									
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">ANOVA</th> <th colspan="2">ACOVA</th> <th>Co-1§</th> <th>Co-2§</th> <th>Co-1&2</th> </tr> <tr> <th>F value</th> <th>P value</th> <th>F value</th> <th>P value</th> <th>F value</th> <th>F value</th> <th>F value</th> </tr> </thead> <tbody> <tr> <td colspan="8">Main effects</td> </tr> <tr> <td>Judgments (J)</td> <td>2.2</td> <td>0.16</td> <td>4.4</td> <td>0.05</td> <td>1.67</td> <td>0.00</td> <td>0.95</td> </tr> <tr> <td>Conditions (C)</td> <td>6.2</td> <td>0.02</td> <td>7.6</td> <td>0.01</td> <td>0.20</td> <td>0.94</td> <td>0.40</td> </tr> <tr> <td colspan="8">Interactions</td> </tr> <tr> <td>J x C</td> <td>0.4</td> <td>0.55</td> <td>1.00</td> <td>0.32</td> <td>1.15</td> <td>0.37</td> <td>0.93</td> </tr> </tbody> </table>										ANOVA		ACOVA		Co-1§	Co-2§	Co-1&2	F value	P value	F value	P value	F value	F value	F value	Main effects								Judgments (J)	2.2	0.16	4.4	0.05	1.67	0.00	0.95	Conditions (C)	6.2	0.02	7.6	0.01	0.20	0.94	0.40	Interactions								J x C	0.4	0.55	1.00	0.32	1.15	0.37	0.93
	ANOVA		ACOVA		Co-1§	Co-2§	Co-1&2																																																								
	F value	P value	F value	P value	F value	F value	F value																																																								
Main effects																																																															
Judgments (J)	2.2	0.16	4.4	0.05	1.67	0.00	0.95																																																								
Conditions (C)	6.2	0.02	7.6	0.01	0.20	0.94	0.40																																																								
Interactions																																																															
J x C	0.4	0.55	1.00	0.32	1.15	0.37	0.93																																																								

§ Co-1 and 2 refers to co-variate of word frequency of the top and the bottom word respectively

Table 7.1b Mean error rates (%) in Experiment 8, analysed by subjects (By Ss) and by- Items

Judgment	Condition	No. of stimuli	By Subjects		By items		By Subjects	By items
			Mean	(SD)	Mean	(SD)	(1) vs (2)	(1) vs (2)
Same								
	Group 1	30	7.2	(5.1)	15.2	(16.3)		
	Group 2	30	6.9	(4.3)	13.7	(14.4)		
	All	60	7.0	(4.7)	14.4	(15.3)		
Different								
	(1) Control	30	4.7	(3.9)	6.6	(9.3)	Z=1.62 P>0.05	Z=1.75 P>0.05
	(2) Homophone	30	4.3	(3.4)	9.9	(14.2)		
	All	60	4.5	(3.6)	8.2	(11.7)		

Z values were estimated from Wilcoxon signed rank test

by item analysis on the 'Different' judgments ($t_{(29)}=1.18, P>0.05$). Clearly the difference (30-40 msec) in mean RT between homophone and control conditions is, in any case, in the direction opposite to that predicted by the hypothesis of phonological mediation, both in by-subject and by-item analyses.

On error rates, the false 'Same' rate for homophones was higher than that for the control items on the by-item analysis; the difference was in the opposite direction on the by-subject analysis. However, neither of these differences was statistically significant by Wilcoxon signed rank test (by-subjects: $Z=1.62, P>0.05$; by-item: $Z=1.75, P>0.05$).

Discussion

The results of this experiment seem very clear. There was absolutely no 'homophone effect' on RTs, and no significant difference in false 'Same' responses in the homophone condition. This result suggests that the semantic recoding of Chinese words is, in practice, independent from phonological recoding. It is possible, and indeed plausible, that phonological and semantic processing may be automatically activated in the semantic processing task. Although these processes can be accomplished independent of each other, some subjects, on some trials, may nevertheless inadvertently attend to the phonological information in the semantic comparison task. This would account for the finding that synonyms had a (non-significantly) higher error rate than the control items (9.9% vs 6.6% in the by-item data). If so, this may be considered as an optional individual strategy in reading. The results are clearly incompatible with the hypothesis that phonological recoding plays a necessary role in semantic judgments on Chinese words.

Experiment 9: The semantic comparison of Chinese words with matching components

Introduction

This experiment on semantic comparison was designed in parallel with Experiment 6, on phonological comparison. The primary concern in this experiment is the functional role of the lexical radical, in comparison with that of the non-radical component, in the semantic recoding of Chinese words. Studies of word recognition in English orthography have provided less knowledge about semantic processing than about the phonological recoding of written words. There is little doubt that the pronunciation of English words is determined by orthographic units smaller than whole words. However, the role of lexical decomposition in the semantic recoding of written English is far less clear.

As described in Chapter 2, all written Chinese words that can be decomposed (that is, all words with more than one unit) contain a unique orthographic unit, termed the lexical radical. In etymology, the lexical radical was used to represent the semantic category of the words formed therefrom. Each lexical radical possesses a distinctive structure and occupies a specific position in all the different words containing it (the positional regularity; see Chapter 2). It is therefore a unique orthographic unit of most Chinese words, distinguished both by its form and its position. The structure and legal positions of the lexical radicals can be found in Chinese dictionaries (*e.g.*, *XIN-HUA*, 1979).

The preceding study of the visual processing of Chinese words (Chapter 4) has shown that the lexical radical and its positional regularity can be recognized and used by native Chinese readers. The present experiment aims to address further the issue of the functional role of lexical radicals in a semantic processing task. Can subjects recognize the lexical radicals as the semantic component in

accordance with the part-to-whole semantic correspondence rule in Chinese orthography (Chapter 2)? In a pilot study (Experiment 4), this issue was explored by comparing the effects of word pairs with matching components (as a distractor toward false 'Same' responses), in both semantic and phonological comparisons. In Experiment 4, word-pairs in which the lexical radical of one word physically matched the other (simple) word in the pair had no significant influence on the phonological comparison task; however, in Experiment 4 this condition yielded significantly higher false 'Same' responses when the task was changed to the comparison of meanings (semantic comparison), even though the subjects and stimuli were identical. The lexical radical (the rule-defined semantic component) is apparently recognized and used by native readers of Chinese in semantic judgments, as is the non-radical phonetic component in phonological comparisons. Subjects allocate attentional priority toward the lexical radical component in semantic judgment tasks. However, the cross-over pattern on the Word-Radical match and Word-Component match conditions in semantic and phonological comparison tasks, respectively, was only manifested in error rates and not in RTs. Using better controlled stimuli with a computerised task, the current experiment is intended to re-examine the critical Word-Radical match and Word-Component match conditions in the semantic comparison task, to be contrasted with the same stimulus pairs in the phonological comparison task (Experiment 6).

In this experiment the subjects and stimuli (in these two critical stimulus-conditions) were identical to those in Experiment 6. The rationale is as described in Chapter 5. If the semantic nature of the lexical radical is implicitly recognized, and so given attentional priority, subjects should be more biased toward false 'Same' responses when the word pairs have a matching lexical radical (Word-Radical match) than when they match in respect of the non-radical component (Word-Component match). Conversely, in the phonological comparison task,

with identical stimulus pairs, subjects were biased toward false 'Same' responses only in the Word-Component match condition (Experiment 6). In this Experiment, word-pairs with radical matching included the Graphic-Radical match as well as Word-Radical match. The actual stimuli used in the Word-Radical match and Word-Component match in this experiment did not overlap with those used in the pilot study (Experiment 4).

Stimulus Conditions

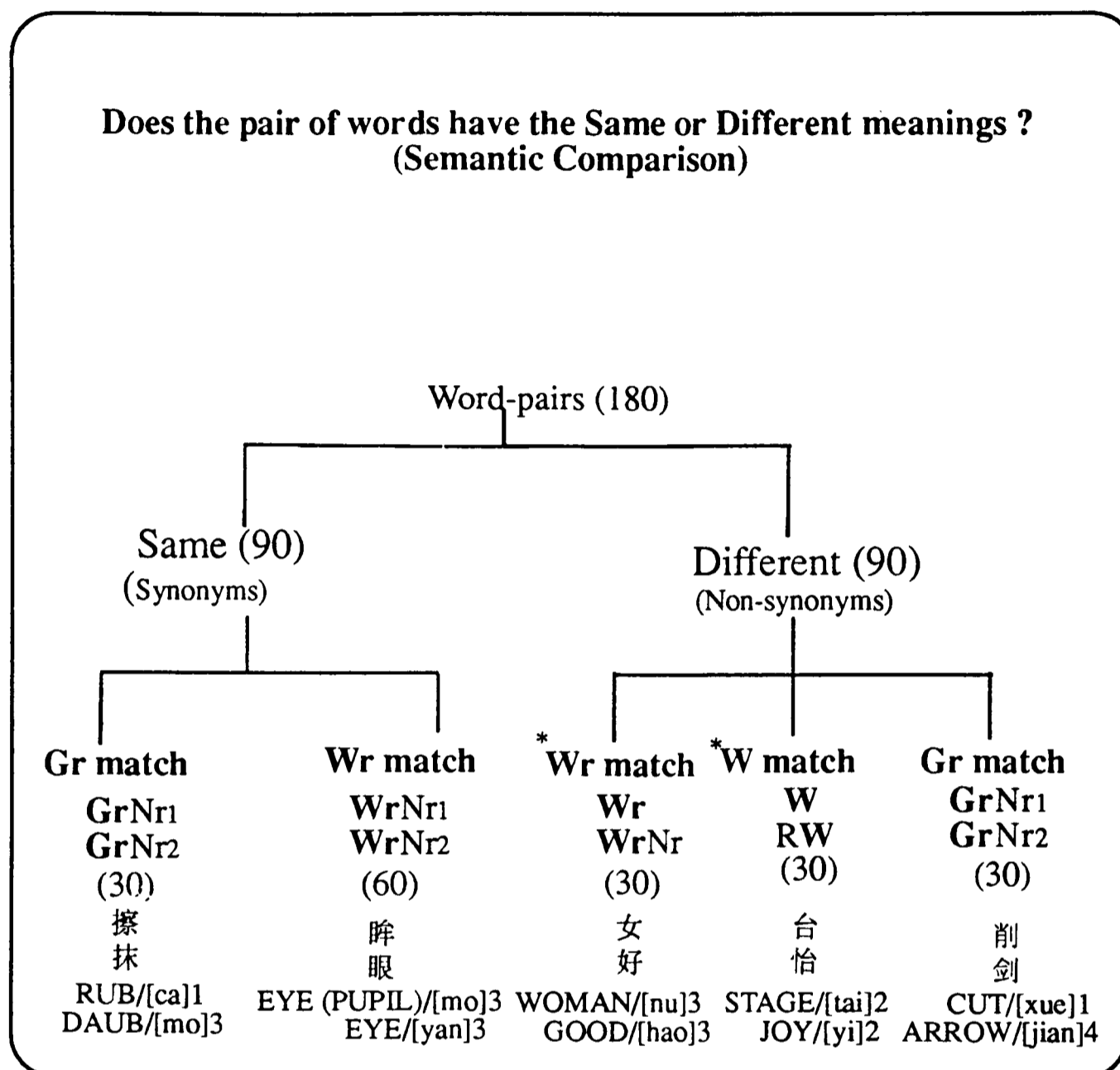
Figure 7.2 illustrates the distribution of the total pairs of stimuli, with equal numbers of 'Same' and 'Different' pairs in the five conditions, with examples of each condition. Stimulus-pairs were classified according to the matching component in each pair of stimuli. The conditions for the 'Different' judgments are Word-Component match, Word-Radical match and Graphic-Radical match. Each condition had 30 stimulus pairs. For the 'Same' (synonym) judgments, there were only two different conditions, Word-Radical match and Graphic-Radical match, with 60 pairs of Word-Radical match and 30 pairs of Graphic-Radical match. For both 'Same' or 'Different' pairs, the two words in the same pair are not meaningful two-character CI in Chinese. Stimuli in each condition are described in detail below:

'Different' judgments (non-homophone & non-synonym pairs)

Word-Radical & Word-Component match conditions: The stimuli in these two critical conditions were identical to those used in Experiment 6. That is, pairs of words in which the lexical radical or non-radical component of one word is identical to another whole (simple) word in the pairs. The pairs of words are not synonyms nor homophones.

Graphic-Radical match: The two words have different meanings but share the same graphic-radical. Unlike a word-radical, a graphic-radical is unpronounceable and can be meaningless in isolation.

Experiment 9



** the same set of stimuli was also used in Experiment 6*

Note: R: and Nr refer to lexical radical and non-radical component.

W and Wr refer to Word-component and Word-radical

Gr and G refer to Graphic-radical and Graphic-component

Bold typeface indicates the matching component in each stimulus pair

Figure 7.2. Stimulus conditions in Experiment 9

'Same' judgment (synonym pairs)

Word-Radical match: Corresponding to the Word-Radical match condition for the 'Different' judgments, in this condition pairs of words have closely similar meanings, i.e., the matching word-radical (the simple word) has a meaning congruent with the compound word in which the same element occurs as a lexical radical.

Graphic-Radical match: Corresponding to the Graphic-Radical match condition for the 'Different' judgments, in this condition two words share the same graphic-radical and are also congruent in their meanings, e.g.,

Baseline data: Baseline data for the 'Same' judgments were taken from those of Synonym 2; for 'Different' judgments they were taken from those of the control items (non-synonym) in Experiment 8.

Results

1) 'Same' and 'Different' judgments

Table 7.2a presents the mean RTs for both the by-subject and by-item data, with the results of the statistical analyses. The corresponding data for error rates are shown in Table 7.2b. In the by-subject ANOVA on RTs, the main effect of condition was highly significant ($F_{(2,11)}=27.3$, $P<0.001$), but there was no significant effect of judgment (Table 7.2a). The interaction of condition and judgment was also significant ($F_{(1,22)}=5.6$, $P=0.01$). After adjustment for the word frequencies of both words, as two co-variates, ACOVA on the by-items RTs showed consistent results. There was a significant condition effect and a significant interaction of condition and judgment (*Condition*: $F_{(1,29)}=3.8$, $P<0.05$; *Interaction of J x C*: $F_{(2,58)}=12.0$, $P<0.001$). There was no significant judgment effect (Table 7.2a). The frequency of either word (or both together) did not significantly affect either the main effects or their interactions.

Further analyses of the condition effect were applied to RT and error rates in 'Same' and 'Different' judgments separately, with the focus on the critical conditions of Word-Radical match and Word-Component match on 'Different' judgments. As in Experiment 6, 'Same' and 'Different' judgments in each type of stimulus-condition were also compared to the 'baseline' data.

2) 'Different' judgements: effects of radical and non-radical matching

For the 'Different' judgments, the stimulus conditions with matching components were Word-Radical match, Graphic-Radical match and Word-Component match. RTs and error rates for 'Different' judgments in these three conditions are shown in Figure 7.3 together with the baseline data. RTs and error rates are shown separately in by-subjects and by-items data .

On RTs, both types of stimulus pairs with matching radical components yielded longer times to response than the Word-Component match pairs, for 'Different' judgments. This difference in RTs was highly significant for the Graphic-Radical match condition, in both by-subject ($t_d=9.23$, $P<0.01$) and by-items analyses ($t_d=9.04$, $P<0.01$). However, for the Word-Radical match condition, the difference was significant in the by-subjects analysis ($t_d=4.75$, $P<0.01$) but not by items. The important point here is that performance on the word-pairs which matched respectively on radical and non-radical components in *semantic* comparisons showed a pattern opposite to that found in *phonological* comparisons (See Figure 7.4). The same Word-Radical pairs which were faster than the Word-Component match pairs (and also the baseline data) in the phonological comparison task (Experiment 6) were slower than the Word-Component match pairs in the semantic comparison task (Experiment 9), even though the stimuli and subjects used in these two tasks were identical. However, the difference between these two conditions was apparently smaller in the semantic comparison task than that in the phonological comparison task. An

Table 7.2a Mean reaction times (msec.) both by-subjects and by-items in Experiment 9, with the statistical results

Judgment	Condition	No. of stimuli	By Subjects		By items		Comparison group	By Subjects	By items
			Mean	(SD)	Mean	(SD)		t or Q value§	t or Q value§
Same									
	(1)Baseline	30	1039	(141)	1076	(91)	(1) vs (2)	1.75	0.29
	(2)Word-Radical	60	1078	(172)	1107	(104)	(1) vs (3)	3.46*	1.97
	(3)Graphic-Radical	30	1117	(149)	1153	(122)	(2) vs (3)	1.74	1.37
	All	120	1078	(154)	1112	(106)			
Different									
	(1)Baseline	30	1133	(238)	1159	(98)	(1) vs (2)	1.50	1.86
	(2)Word-Radical	30	1093	(181)	1100	(134)	(1) vs (3)	3.60*	3.12*
	(3)Word-Component	30	1037	(165)	1060	(81)	(1) vs (4)	3.04*	2.70*
	(4)Graphic-Radical	30	1215	(233)	1237	(120)	(2) vs (3)	4.75**	1.78
							(2) vs (4)	8.07**	5.86**
							(3) vs (4)	9.23**	9.04**
	All	120	1119	(204)	1139	(108)			

Main effects	ANOVA		ACOVA		Co-1†	Co-2†	Co-1&2
	F value	P value	F value	P value	F value	F value	F value
Judgments (J)	1.2	0.29	0.1	0.72	0.05	0.08	0.05
Conditions (C)	27.3	0.00	3.8	0.02	0.85	3.10	1.86
Interactions							
J x C	5.6	0.01	12.0	0.00	0.09	1.44	0.81

§ Dunette t tests were used for comparisons against the baseline data while Newman-Keuls Q tests were used for all other comparisons

† Co-1 and 2 refer to word frequency of the top and the bottom word respectively

*, ** P<0.05 and P<0.01 respectively

Table 7.2b Mean error rates (%) both by-subjects and by-items in Experiment 9, together with the statistical results.

Judgment	Condition	No. of stimuli	By Subjects		By items		Comparison group	By Subjects	By items
			Mean	(SD)	Mean	(SD)		t or Q value§	t or Q value§
Same									
	(1)Baseline	30	6.9	(4.3)	13.7	(14.4)	(1) vs (2)	4.75*	2.25
	(2)Word-Radical	60	12.2	(6.0)	24.0	(23.1)	(1) vs (3)	3.81*	1.80
	(3)Graphic-Radical	30	11.0	(4.8)	22.1	(18.5)	(2) vs (3)	1.42	0.37
	All	120	10.0	(5.0)	19.9	(18.7)			
Different									
	(1)Baseline	30	4.7	(3.9)	6.6	(9.3)	(1) vs (2)	5.87**	3.30*
	(2)Word-Radical	30	10.2	(5.1)	19.4	(25.7)	(1) vs (3)	0.23	0.12
	(3)Word-Component	30	2.7	(2.4)	5.7	(11.4)	(1) vs (4)	3.21**	1.79
	(4)Graphic-Radical	30	6.9	(4.8)	13.6	(16.7)	(2) vs (3)	7.51**	4.41**
							(2) vs (4)	3.94*	2.01
							(3) vs (4)	4.30*	3.22*
	All	120	6.1	(4.0)	11.2	(15.8)			

§ Dunette t tests were used for comparisons against the baseline data while Newman-Keuls Q tests were used for all other comparisons

*, ** P<0.05 and P<0.01 respectively

'Different' Judgements: Component matching effects in semantic comparison (Experiment 9)

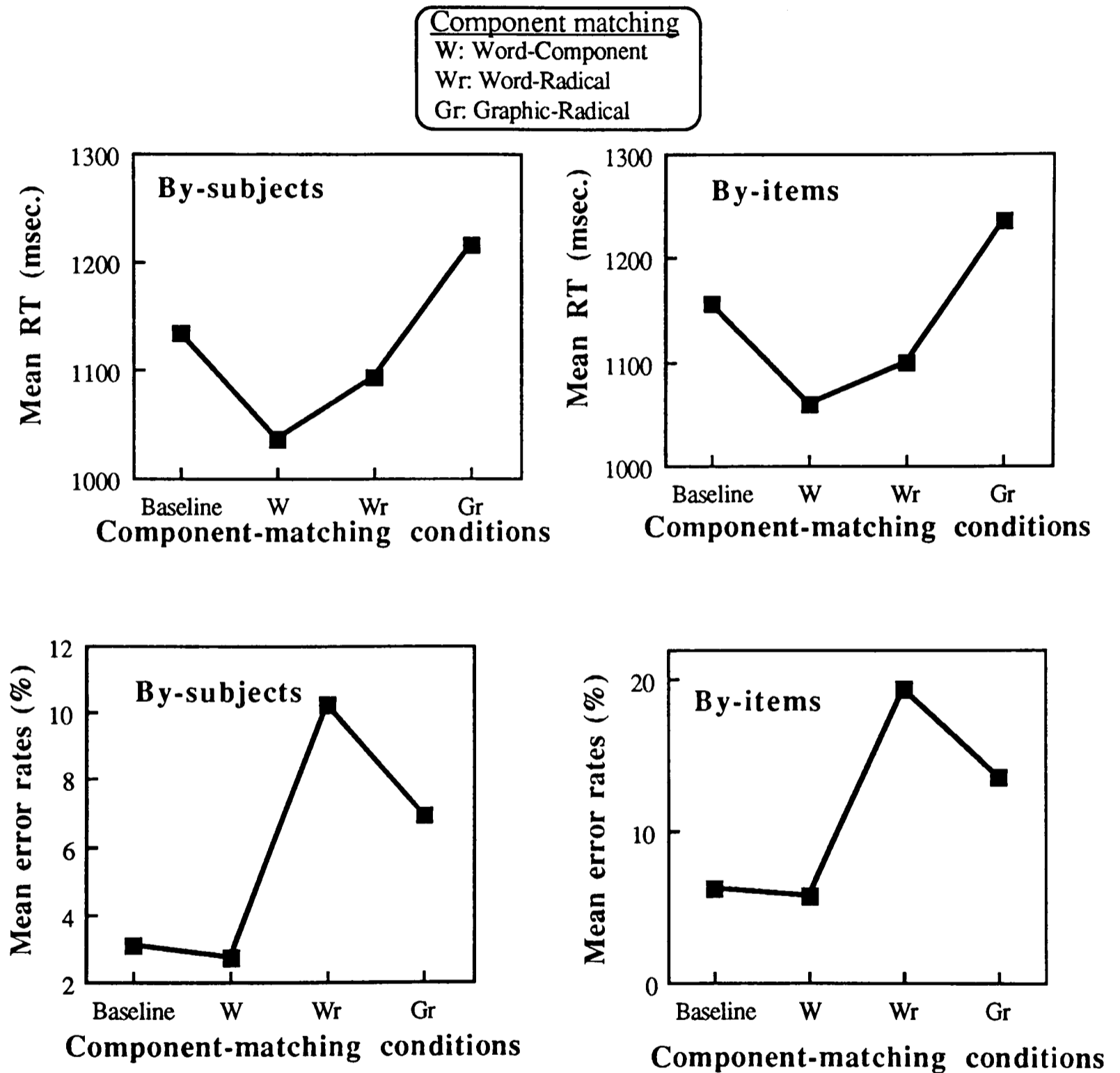


Figure 7.3 Mean RTs and error rates (by subject and by items) in semantic comparison

'Different' Judgments: Component matching effects in both phonological and semantic comparison (Experiment 6 & 9)

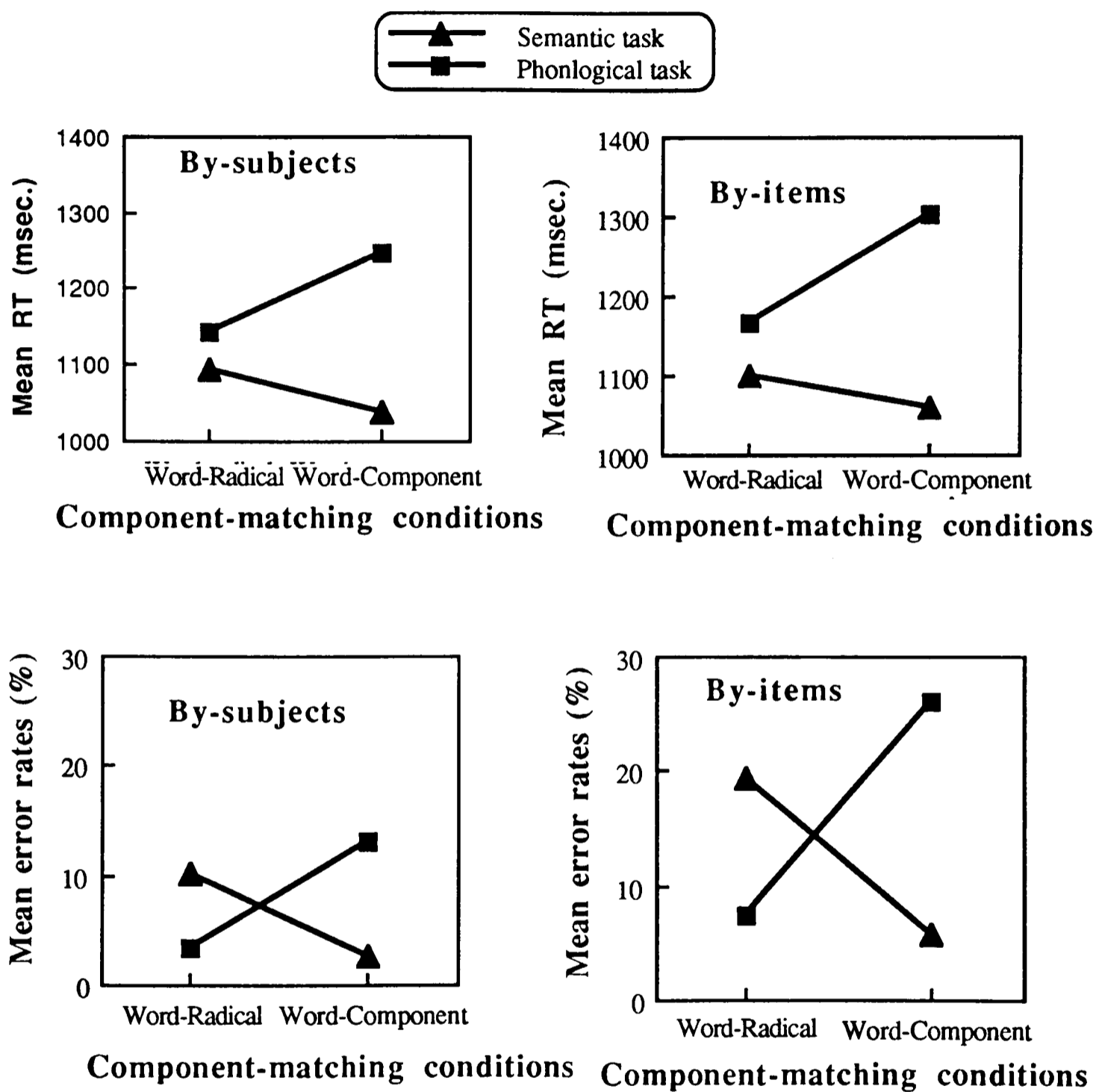


Figure 7.4 Mean RTs and error rates (by subject and by items) in both phonological and semantic comparison

ANOVA and ACOVA were performed for by subjects and by items analysis, using joint data from semantic and phonological comparisons on these two types of component-matching pairs, on Different judgments only. The interaction between task and condition was highly significant on RT (*By subjects: $F_{(1,11)}=17.47, P<0.01$; By items: $F_{(1,29)}=9.20, P<0.01$*) and error rate (*By subjects: $F_{(1,11)}=64.5, P<0.001$; By items: $F_{(1,29)}=20.01, P<0.001$*)

Compared to the baseline condition (with no matching elements), both the Word-Radical match and Word-Component match conditions were slightly faster both by-subjects and by-items. However, the contrast was significant only for the Word-Component match condition (*by-subject analysis: $t_d=3.04, P<0.05$; by items: $t_d=2.70, P<0.05$*). The Word-Radical match condition was not significantly different from the baseline data by Dunnett t test.

The most clear-cut and striking results are found in the error rate data. False Same rates in the Radical match conditions (Word-Radical and Graphic-Radical match) were much higher than in the Word-Component match condition (See Figure 7.3). This difference in error rates was highly significant by Dunnett test both in the by-subject analysis (*Word-Radical: $t_d=7.51, P<0.01$; Graphic-Radical: $t_d=4.30, P<0.05$*) and in the by-items analysis (*Word-Radical: $t_d=4.41, P<0.01$; Graphic-Radical: $t_d=3.22, P<0.05$*). The difference between Word-Radical match and the baseline was also significant (*By subject: $t_d=5.87, P<0.01$; by items: $t_d=3.30, P<0.05$*). Error rates in the Word-Component condition were close to the baseline value; the comparison did not even approach significance. Figure 7.4 shows the error rate data, in the semantic comparison task, contrasted with the results from Experiment 6, with the identical stimulus-pairs, in the phonological comparison task. The cross-over interaction between task and condition is very striking.

3) 'Same' judgments: Effects of radical matching

On 'Same' judgements, the two conditions with a matching component were the Graphical-Radical and the Word-Radical match. As seen in Table 7.2a, no significant differences between these two radical matching conditions were found in either by-subjects or by-items analyses, either in RTs or in error rates. However, both these conditions had significantly higher error rates than the baseline in the by-subject analysis (*Word-Radical match*: $t_d=3.46$, $P<0.05$; *Graphic-Radical match*: $t_d=3.81$, $P<0.05$). In the by-item analysis, this difference was significant only for Word-Radical match ($t=2.25$, $P<0.05$).

4) Influence of radical frequency index upon the semantic comparisons

The radical frequency index of components in the Word-Radical match and Word-Component match conditions was also examined in this experiment for its effect on semantic comparison. With the word frequency of both words in each pair controlled in a multiple regression analysis, the radical frequency index showed an association opposite to that in the phonological comparison task in Experiment 6. Instead of a negative association, there was a small, but significant, positive association between radical frequency index and errors in the semantic comparison task. The higher the radical frequency, the higher the false 'Same' response rate for comparisons on the pairs of stimuli that contain the component. This was significant only for error rates and not for RTs. Figure 7.5 depicts the regression of the radical frequency index over the 60 stimuli in terms of comparison speed and accuracy. (For comparison, see Figure 6.4 for this regression in the phonological comparison task).

Discussion

The most important finding in this experiment is that semantic comparison of the word-pairs was significantly affected by the lexical-radical (Word-Radical

**Radical frequency index: effect on a semantic judgment task
(Experiment 9)**

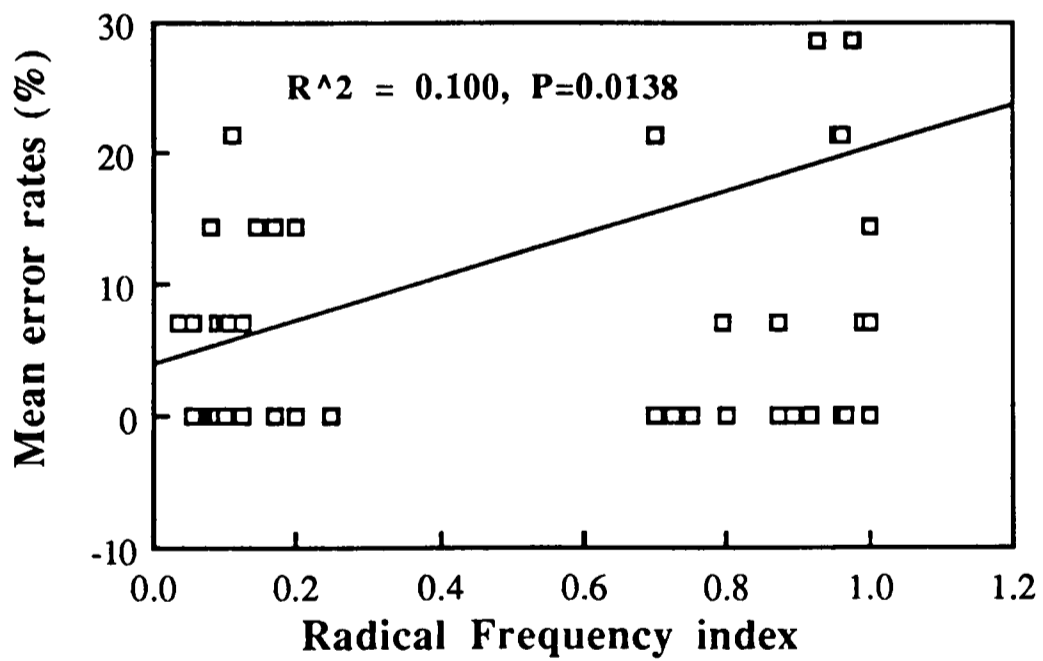
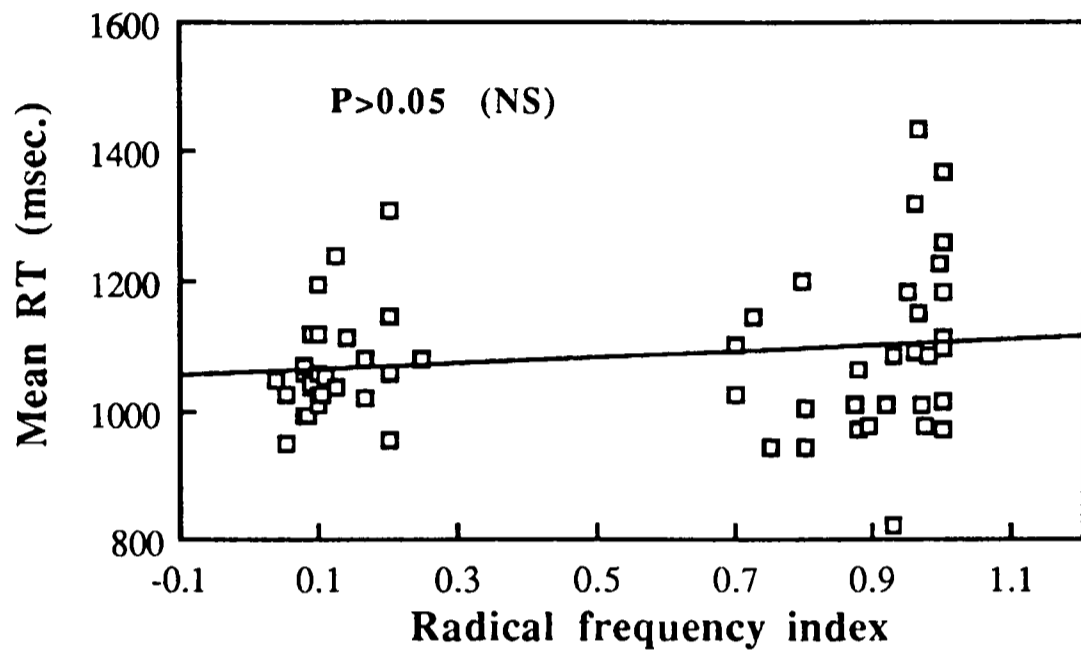


Figure 7.5 Correlation between mean RT and error rate and radical frequency index

match and Graphic-Radical match), as a distractor toward false 'Same' responses, but not by the non-radical component (Word-Component match). This is precisely the opposite result to that found in the phonological comparison task with the same stimuli and the same subjects. These results together (Experiment 6 & 9) suggest that subjects (skilled Chinese readers) have a selective attentional bias toward the lexical radical in semantic judgments and towards the non-radical component in the phonological judgments. That is, the lexical radical is likely to receive attentional priority in semantic tasks, while the non-radical component may become salient in phonological tasks. The significant positive association between the radical frequency index and false 'Same' errors in semantic comparisons, in this experiment, further supports the hypothesis that this preferential attentional bias toward the lexical radical increases as a function of the relative frequency with which that orthographic component functions as the lexical radical in *other* words in the lexicon. However, the attentional bias toward the lexical radical in this semantic comparison task (Experiment 8) was somewhat weaker than the attentional bias toward the non-radical component in the phonological judgment task (Figure 7.4).

Experiment 10: The semantic comparison of word-pseudoword pairs

Introduction

This experiment aimed to examine the radical and non-radical matching conditions extended to word-pseudoword stimulus pairs. The design of this experiment corresponds to Experiment 7 (phonological comparison). The subjects and the whole set of stimuli were identical to those used in Experiment 7. The sole difference between Experiments 7 and 10 is the task. In this experiment, instead of a judgment on pronunciation (as in Experiment 7), subjects were asked to judge whether the stimulus pairs, consisting of one word and one pseudoword, have a similar meaning or not. The rationale is otherwise the same as in Experiment 9. If the (regular) lexical radical in a pseudoword is likewise found to have attentional salience in a semantic judgment task, as it did with real word-pairs, this would further support the *rule-based* nature of the inferred lexical decomposition process, regardless of the lexical or non-lexical status of the whole word. This result would thus mirror that found in Experiment 7 (phonological comparison).

The subjects, stimulus conditions and procedure were identical to those in Experiment 7.

Results

Given that pseudowords (at least in Chinese) have virtually no meaning, all 'Same' responses were regarded as 'errors' for the purpose of the analyses. The false 'Same' response rates were thus used, as the main independent variable.

Table 7.3a presents the mean 'Same' rates in both by-subjects and by-items analyses. The corresponding RT data are also given in Table 7.3b. An ANOVA was performed to examine the two main factors of pseudoword-type and component matching conditions (radical vs non-radical match). Both main

Table 7.3a Semantic comparison of word-pseudoword pairs: Mean Same response rates(%) in Experiment 10, with the results of the statistical analyses

Pseudoword type	Condition	No. of stimuli	By Subjects		By items		By Subjects	By items
			Mean	(SD)	Mean	(SD)	(1) vs (2)	(1) vs (2)
Pseudosimple								
(1) Graphic-radical		15	43.6	(4.9)	12.3	(23.3)	t=8.7 P=0.00	t=0.79 P=0.44
(2) Graphic-component		15	9.9	(10.3)	18.0	(10.1)		
All		30	26.7	(7.6)	15.1	(16.7)		
Pseudocompound								
(1) Word-radical		15	31.9	(14.8)	31.0	(10.7)	t=2.9 P=0.01	t=2.5 P=0.02
(2) Word-component		15	11.8	(9.5)	20.9	(13.3)		
All		30	21.8	(12.1)	25.9	(12)		
			ANOVA		ACOVA		Co-variate†	
Main effects			F value	P value	F value	P value	F value	P value
Pseudotype (P)			5.4	0.04	5.7	0.03	0.06	0.81
Conditions (C)			27.5	0.00	0.2	0.67	0.20	0.63
Interactions								
J x C			8.7	0.01	1.6	0.22	0.26	0.62

† Co-variate refers to word frequency of the real word in the word-pseudoword pairs

Table 7.3b Semantic comparison of word-pseudoword pairs: Mean reaction times (msec.) in Experiment, along with statistical results.

Pseudoword type	Condition	No. of stimuli	By Subjects		By items		By Subjects	By items
			Mean	(SD)	Mean	(SD)	(1) vs (2)	(1) vs (2)
Pseudosimple								
(1) Graphic-radical		15	1132	(176)	969	(142)	t=0.4 P=0.66	t=0.9 P=0.35
(2) Graphic-component		15	1178	(352)	1275	(579)		
All		30	1155	(264)	1122	(360)		
Pseudocompound								
(1) Word-radical		15	1254	(311)	1139	(216)	t=1.1 P=0.27	t=0.1 P=0.88
(2) Word-component		15	1159	(316)	1158	(142)		
All		30	1206	(313)	1148	(179)		

factors were significant (*Pseudoword type*: $F_{(1,11)}=5.4$, $P<0.05$; *Stimulus-Condition*: $F_{(1,11)}=27.5$, $P<0.001$) in the ANOVA. The interaction of pseudoword type and condition was also significant ($F_{(1,22)}=8.7$, $P=0.01$). However, after adjustment for word frequency, only the effect of pseudoword-type was significant ($F_{(1,11)}=5.7$, $P<0.05$). The interaction of pseudoword-type and condition was not significant (Table 7.3a). The word frequency of real words as a co-variate individually had no significant influence upon the main effects or their interactions. Further analyses were performed to examine the conditions effect in each type of pseudoword.

1) Word-Radical and Word-Component Match in word-pseudoword pairs

False 'Same' rates in the different stimulus conditions are illustrated in Figure 7.6. In the pseudo-compound pairs, a false 'Same' rate of 31.9% for Word-Radical match and 11.8% for the Word-Component match condition, was significantly different by paired t test ($t_{(11)}=2.9$, $P=0.01$). This persisted after the adjustment of word frequency in the by-item data ($t_{(14)}=2.5$, $P<0.05$).

2) Graphic-Radical match and Graphic-Component match in word-pseudoword pairs

In the pseudosimple stimulus pairs, the difference in 'Same' response rates between Graphic-Radical and Graphic-Component conditions was highly significant in by-subjects analysis ($t_{(11)}=8.7$, $P<0.001$). However, in the by-items analysis, this difference actually reverses, and was far short of significance. Evidently, the effect depends on a subset of stimulus items.

It is noteworthy that these effects of matching components are the reverse of those found in Experiment 7, with phonological comparison of the identical set of word-pseudoword pairs, with the same subjects (See Figure 7.6). The cross-

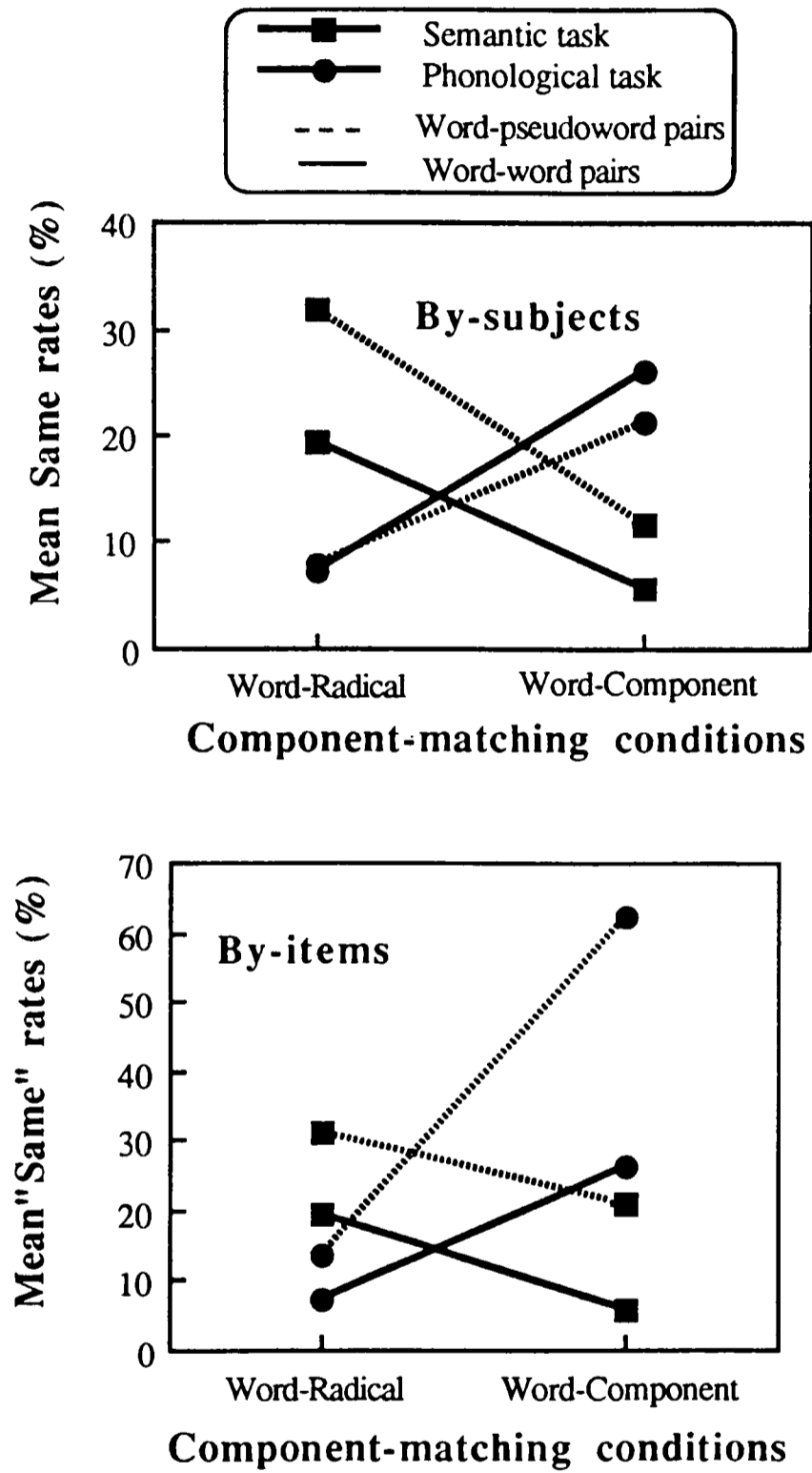


Figure 7.6 Mean "Same" rates (by subjects and by-items) in Experiment 7 and 10

over was consistent with the findings in the semantic and phonological comparisons that used real word-pairs, in Experiment 6 & 9 (See Figure 7.4).

Discussion

This experiment has demonstrated that in guessing the meaning of an 'unknown' word (a pseudoword), subjects were also more likely to be biased by the lexical radical than by the non-radical component of the pseudoword. They were far less likely to rely on the potential semantic meaning of the matching non-radical component of the pseudoword, which was also identical to the adjacent simple word and also meaningful in isolation. When a matching component in isolation is both unpronounceable and meaningless (as are the graphic-radical and graphic-component in the pseudosimple pairs), subjects were more likely to rely on the graphic-radical for a semantic judgment. This further suggests that the lexical radical, irrespective of its strictly lexical status, may be treated as the salient source of semantic information in the semantic recoding of Chinese words.

General Discussion and Conclusion

The general discussion and conclusion focuses on the following two aspects:

1) A direct route from script to meaning

Current word recognition models, including connectionist accounts (*e.g.*, *Seidenberg and McClelland, 1989*), tend to provide less specific accounts of semantic processing than of phonological processing in reading written words. Nevertheless, there are some general features found in most word recognition models. One is that phonological mediation is not necessary for the semantic recoding of written words. In Chinese, direct access from script to meaning is implied by the terms, ideography, logography or lexigraphy, conventionally used for Chinese orthography. However, the traditional story of picture-like reading

in Chinese (e.g., *Barron, 1978, Wang, 1973*) has been modified and, in more recent accounts, phonological structure is considered to be activated automatically in reading Chinese and may also play an important role in short term memory for Chinese text (*Tzeng & Hung, 1980; Tzeng, Hung & Wang, 1977; Baddeley, 1979; Perfetti & McCutchen, 1982; Perfetti & Zhang, 1991*). Here it is worth pointing out that the activation and the utilization of information can be very different questions. Phonological and semantic information from Chinese words may be activated automatically and independently of each other. However, the utilization of semantic and phonological information may tend to be linked. In the present study, synonyms (in the pronunciation task) as well as homophones (in the semantic task) tended to produce higher false 'Same' response rates than the control (Experiment 5 & 8). These results imply that although phonological recoding may be unnecessary in order to do the semantic task, subjects (or some of them) may nevertheless sometimes attend to the phonological information. The implications of individual reading strategies, and their influence upon the reading of Chinese, will be investigated in the next chapter.

2) The functional role of the lexical radical

The focus of the current series of experiments was the functional role of the lexical radical compared to that of the non-radical component in semantic processing. Experiment 9 has shown that the lexical radical receives attentional salience in semantic judgments. This was demonstrated by the fact that subjects took longer to reject, as semantically 'Different', pairs of words which were matched on the lexical radical, and they made significantly more false 'Same' responses to these Word-Radical match pairs compared both to control (non-matching) items and to word pairs matching on the non-radical component. This effect was the opposite to that found with the same set of stimuli and the same set of subjects in the phonological comparison task (Experiment 6). The radical

frequency index also had a significant association with accuracy, but in the inverse sense to that on the same set of stimuli in the phonological comparison (Experiment 6). Subjects' performance was not only influenced by the nature of the orthographic components in the stimuli used in the experiment, but also by the frequency of these components as lexical radicals in other words in the lexicon. This favours the view that implementation of a rule is influenced by the strength of association between that component and all words in the lexicon that contain that same component.

A similar attentional bias toward the lexical radical within a compound character was found, in Experiment 10, even when the character was not a true lexical item (i.e., was a pseudoword), indicating that lexical decomposition, and selective attentional focus on potentially appropriate components of a compound words occurs logically prior to, or independent of, full lexical access. At this point, however, a note of caution is in order concerning the precise role of the lexical radical in semantic processing. It must be remembered that the lexical radical typically indicates a semantic category or semantic association rather than a precise identity as does the non-radical component representing phonological information. For instance, 日 SUN [ri]4 the semantic radical indicates the semantic field of SUN. The individual words that contain 日 are not synonyms of SUN; rather they have a semantic association with SUN, e.g., 时 TIME, 晚 LATE, 早 EARLY. The question as to the precise role of the lexical radical in semantic processing of Chinese words remains open.

Chapter 8

Reading two-character *CI* in Chinese

Introduction

So far the present study has focused on single word recognition. In the preceding four chapters, the visual, phonological and semantic processing of single-character words in Chinese has been studied. This chapter describes an experiment with two-word phrases, called *CI* in Chinese. The experiment is concerned with visual, semantic and phonological aspects of reading two-character *CI*. The relationship between reading efficiency and reading strategies is also investigated in this experiment.

Experiment 11: Lexical decision with two-character CI

Background and Rationale

1) Paradigm: Lexical decision in Chinese with two-character *CI*

Two-character *CI* play an important lexical role in Chinese. As introduced before (see Chapter 1), multiple-word *CI*, especially the two-character *CI*, are the most frequently used lexical items in both spoken and written Chinese in modern times (since 1919). The two-character *CI* is, to some extent, analogous to the English compound word containing two free morphemes (carhorse, armchair, toothbrush, headache). They both consist of independent words and are used as more or less fixed expressions in the language. The lexical decision task for two-character *CI* requires the subject to judge *whether or not the two words together form a CI in Chinese*. In contrast to the comparison task with visual, phonological, or semantic criterion, employed in the preceding experiments, this task does not specify the nature of the information on which the judgments are made.

It should be pointed out that lexical decision with two-character *CI* is different from the conventional lexical decision task in English (i.e., whether or not a written symbol is a word) in at least two ways. First, lexical decision with two-character *CI* involves two

words written separately in Chinese. Apart from recognition of the individual words, the task demands the recognition of a specific lexical association between the two words. Second, for "No" responses, 'non-CI' or 'pseudo-CI' consist of two legitimate Chinese words, but in a nonsense combination. It is the combination, not the individual elements, that must be judged.

In this experiment, consistent with the arrangement of words in contemporary Chinese print in mainland China, the two characters are arranged horizontally side by side, instead of vertically one above the other (as in the preceding comparison-task). In this way, lexical decision with two-character CI involves not only word recognition but also recognition of the lexical association between words, which thus more closely approximates real reading.

The aim of this experiment is twofold. First, the experiment is concerned with the contribution of different aspects of information-processing (visual, semantic and phonological) involved in reading two-character CI. This was approached by the examination of different experimental effects, assumed to be associated with specific components of the reading process, as follows.

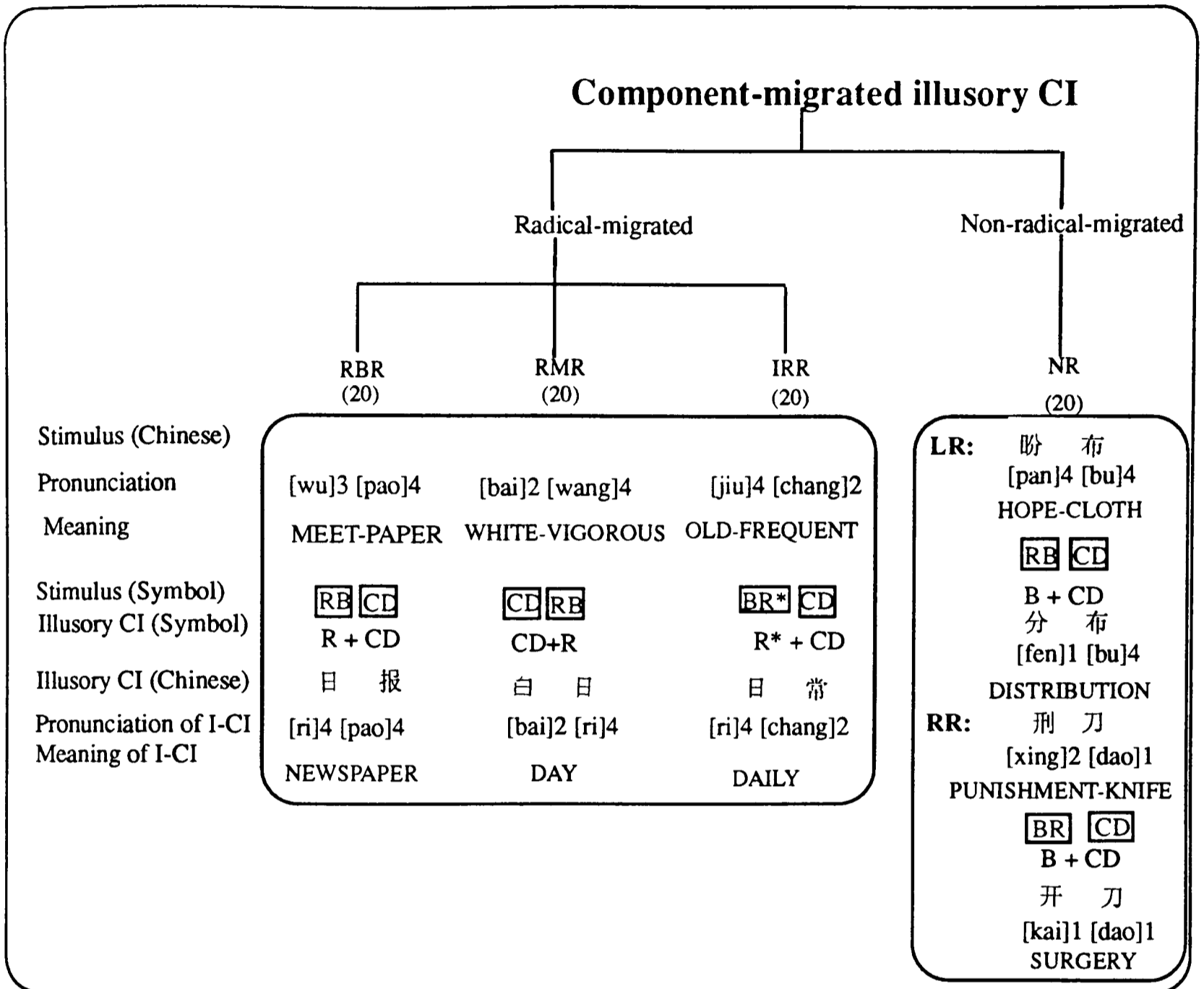
First, the extent to which orthographic-lexical processing is the limiting factor in subjects' performance may be inferred from frequency effects on word recognition (*Besner & McCann, 1987*). The frequencies of two-character CI were derived from the frequency dictionary in Chinese (*Beijing University, 1988*). In this experiment, the effects of CI frequency on lexical decisions of real CI, and of the frequencies of CI-homophones, in which the two words are homophones of the two words in the real CI, were examined.

Second, the meaning of a two-character CI may be classified as concrete or abstract in meaning. A concreteness effect may thus reflect semantic processing. In this experiment, both high and low frequency CI were further divided into three categories of concreteness, using the following criteria. CI such as TELEPHONE, SHOE, THREAD, etc., whose referents are primarily physical objects, were classified as Concrete CI. Almost all of these are object names. By contrast, CI which refer to internal feelings or

mental states or processes, such as LOVE, REGRET, HOPE were classified as Low Concrete CI. CI that could not easily be classified into these two categories but which were also of low imageability, e.g., ORIGIN, SUCCESS, AGAIN, are treated as a third, 'non-specific' category. The complete set of CI in these three categories, are listed in Appendix II.

Third, phonological processing was investigated by constructing CI-homophones, in which two legitimate words together sound like a meaningful two-character CI, but are visually, or orthographically meaningless. For instance, 店化 [dian]4 [hua]4, in which the individual words mean SHOP and TURN, sounds exactly the same (including tone) as the real two-character CI电话[dian]4 [hua]4, meaning TELEPHONE. In addition, the functional role of tone in Chinese is also examined. It is well known that the pronunciations of individual words in Chinese can be distinguished also by their tones. Tone-changed CI-homophones were constructed by substituting a different tone from one of the words, in two-character CI-homophones. For instance, the CI-homophone [dian]4 [hua]4 might be changed to the tone-changed CI-homophone 颠划 [dian]1 [hua]4, in which the individual words mean CROWN and DELIMIT. In this example, the tone of the first word has been changed from tone 4 in the original CI-homophone to tone 1 in the tone-changed version. If tone is not encoded separately from the other aspects of pronunciation, one may expect a difference in the latency and/or accuracy of lexical decisions between the CI-homophone and tone-changed CI-homophone condition. If, however, there is no significant difference between these two conditions, differing solely in their tones, (supposing that a CI-homophone effect is found), then tone is likely to be encoded separately from other aspects of pronunciation.

Finally, lexical decomposition was also examined, based on manipulations of the lexical radicals and the non-radical components of the individual words. This was addressed by constructing stimuli capable of forming "component-migrated illusory CI". Figure 8.1a illustrates the four types of stimuli containing potential component-migrated illusory CI. The stimuli for the component-migrated illusory CI were constructed from two legitimate words (RB & CD in Figure 8.1a), in which the two words together neither sound like nor look like a meaningful two-character CI. Nevertheless, one of the words



Note: R refers to 'migrating' radical; B refers to 'migrating' non-radical component

LR refer to 'migrating' radical' s legitimate position is always in the left side of the different words containing it.

RR refer to 'migrating' radical' s legitimate position is always in the right side of the different words containing it.

(CD) can combine with a component (either R or B) of the other word (RB) to form a meaningful two-character CI, e.g., R + CD or B + CD. The resulting recombination will be referred to here as an "illusory CI" (I-CI), and the stimuli, designed to contain potential I-CI will be referred to as I-CI stimuli. For instance, 晤报 MEET PAPER (in Figure 8.1a) is not a meaningful two-character CI in Chinese. However, 日, a component of 晤, in conjunction with the other whole word 报 forms a potential component-migrated illusory CI 日报, meaning NEWSPAPER (See Figure 8.1a).

If lexical decomposition of the single-character words occurs in lexical decision with two-character CI, so that one component is free to recombine perceptually with other elements, such I-CI stimuli may result in longer RTs and/or give rise to more errors than the control stimuli. Given that lexical decomposition in Chinese concerns primarily the two principal orthographic constituents, the lexical radical and the non-radical component, the I-CI stimuli were divided into (i) potential *radical-migrated* and (ii) potential *non-radical-migrated* illusory CI. These are, respectively, (i) stimuli in which the lexical radical (R) from one word could potentially recombine with the other whole word to form an "illusory CI" (R + CD), and (ii) stimuli in which the non-radical component (B) could do so (B + CD). (Again see Figure 8.1a.) Furthermore, as discussed earlier in this thesis (Chapter 4), lexical decomposition may be strongly influenced by the positional regularity of the lexical radical. The lexical radical is the essential element of most Chinese words (i.e., those 90% or more of Chinese words with more than one stroke-pattern). The lexical radicals have both a specific, fixed form and a fixed position, in the different words in which they appear. The importance of positional regularity has already been shown in the pseudoword advantage over nonwords (See Experiment 2).

In the present experiment, the positions of the critical elements were manipulated in two different ways. First, two types of potential *radical-migrated* illusory CI were selected, as illustrated in Figure 8.1a. Twenty radicals selected from the radical list (XIN HUA, 1979) were employed as potentially migrating lexical radicals (represented by R in Figure 8.1a). The position of the potentially migrating lexical radicals (R) were further classified as regular *beginning* radical (RBR) and regular *middle* radical (RMR) conditions, according to the position of the potentially migrating radical in the

two-character CI as a whole. For instance, 日 is a regular radical in 旺 晴 . It appears at the beginning of the pseudo-CI, 晴 报 , MEET PAPER (RBR) and in the middle of the pseudo-CI, 白 旺 , WHITE VIGOROUS (RMR).

Stimuli capable of forming *non-radical-migrated* illusory CI were formed in a similar way to the potential radical-migrated illusory CI, except that the potentially 'migrating' component was the *non-radical component*. These stimuli were also of two types. In the first type, called "pseudo-radical-migrated "I-CI" (R*), the critical non-radical component, was, also in its graphic form, one of 20 listed lexical radicals (the same as these used in the "radical-migrated" I-CI stimuli). For instance, 日 in isolation is a radical form, however, in 旧 OLD [jiu]⁴ it is not in a regular position (in fact, in 旧 , 丨 is the true lexical radical), since the radical position for 日 is on the left. Thus in 旧 , 日 has the graphic form of a radical, but, in terms of regular orthographic structure 日 is here a non-radical component. Hence, 日 is a pseudo-radical (R*) in 旧 常. Suppose that reading Chinese is based on lexical decomposition of single-character words. The orthographic status of the potentially 'migrating' component within a word, i.e., either lexical radical or non-radical component or the intrinsic graphic form (non-radical or "pseudo-radical") may have significant effects. If, however, the CI is processed as a whole as a single orthographic-lexical unit, the position of the potentially 'migrating' lexical radical in terms of CI as a whole (e.g., RBR or RMR) will be also important. In the second type of non-radical-migrated I-CI stimuli, the critical element was a typical non-radical component, which does not belong to the 189 listed radicals. In this non-radical-migrated illusory CI condition (NR), the lexical radical was always in its regular position in the words containing it. Moreover, the potentially 'migrating' non-radical component could also be either at the beginning or in the middle of the CI. This variable was balanced across the stimuli. As can be seen in Figure 8.1a, apart from the words with left radical (LR), which have typically been used in previous studies in Chinese (e.g., *Seidenberg, 1985a; Flores d'Arcais, 1991*), words with 'right' radicals (RR) were also selected to form the non-radical-migrated illusory CI. 'Left' and 'right' here refer to the legitimate positions of the lexical radical in the different words, as defined in Chinese orthography. Ten left and ten right radicals were employed.

2) Reading strategies

The second objective of this experiment was to explore individual reading strategies. This was investigated by means of a number of different processing indices. These were estimated for individual subjects, separately in terms of both RTs and error rates, as follows:

- 1) Orthographic-lexical frequency-dependence was indexed for each individual subject by the difference in RT (or error rates) between low frequency and high frequency CI, divided by the mean RT (or error rate) for all 180 CI: $(\text{low frequency} - \text{high frequency}) / \text{total CI}$.
- 2) Semantic recoding strategy was indexed for each individual subject by the difference between low concrete CI and high concrete CI in mean RT (or error rates), divided by the mean RT (or error rate) of total CI: $(\text{high concrete} - \text{low concrete}) / \text{total CI}$.
- 3) Phonological recoding strategy was indexed for each individual subject by the difference in RT (or error rate) between the CI-homophone condition and control condition, divided by the control: $(\text{CI-homophone} - \text{control}) / \text{Control}$.
- 4) Lexical decomposition strategy was inferred from both the radical and non-radical component migration indices, based on responses to the I-CI stimuli. (a) Radical-migration was indexed by the difference between regular radical-migration, i.e., $(\text{RBR} + \text{RMR})/2$ and the control, divided by the control, on RT and error rates. (b) Non-radical component migration (NR) was indexed by the difference between non-radical-migration and the control, divided by the control, on RT and error rates. (The R* condition was not included in the latter index, since it is possible that the "pseudo-radical" component may not behave like regular, non-radical component.)
- 3) Reading efficiency was indexed for each individual subject by the subject's overall mean RT (and error rate) in responding to all the CI.

Method

Subjects

16 skilled native Chinese readers and speakers, 8 male and 8 female, aged 20 to 35 years, were paid to participate in the experiment. All of them had graduated from

university in mainland China and were reading for doctorates at the University of Oxford. All subjects were right handed and had normal or corrected to normal vision. Seven of them also participated in the previous 10 experiments with single-character words.

Stimulus Conditions

There was a total of 360 stimulus-pairs in this experiment. The distribution of stimuli over the individual conditions is shown in Figure 8.1b (with examples). There was an equal number of YES-pairs (two-character CI) and NO-pairs (two-character Non-CI), each with 180 stimuli. All stimuli consisted of two legitimate words in the simplified version. The total of 360 pairs of words were presented in randomised order with the constraint that no more than three consecutive trials occurred with the same response (YES or NO). The randomization was determined by a computer programme and was different for individual subjects. The stimuli in each condition are described as follows.

1) Two-character CI

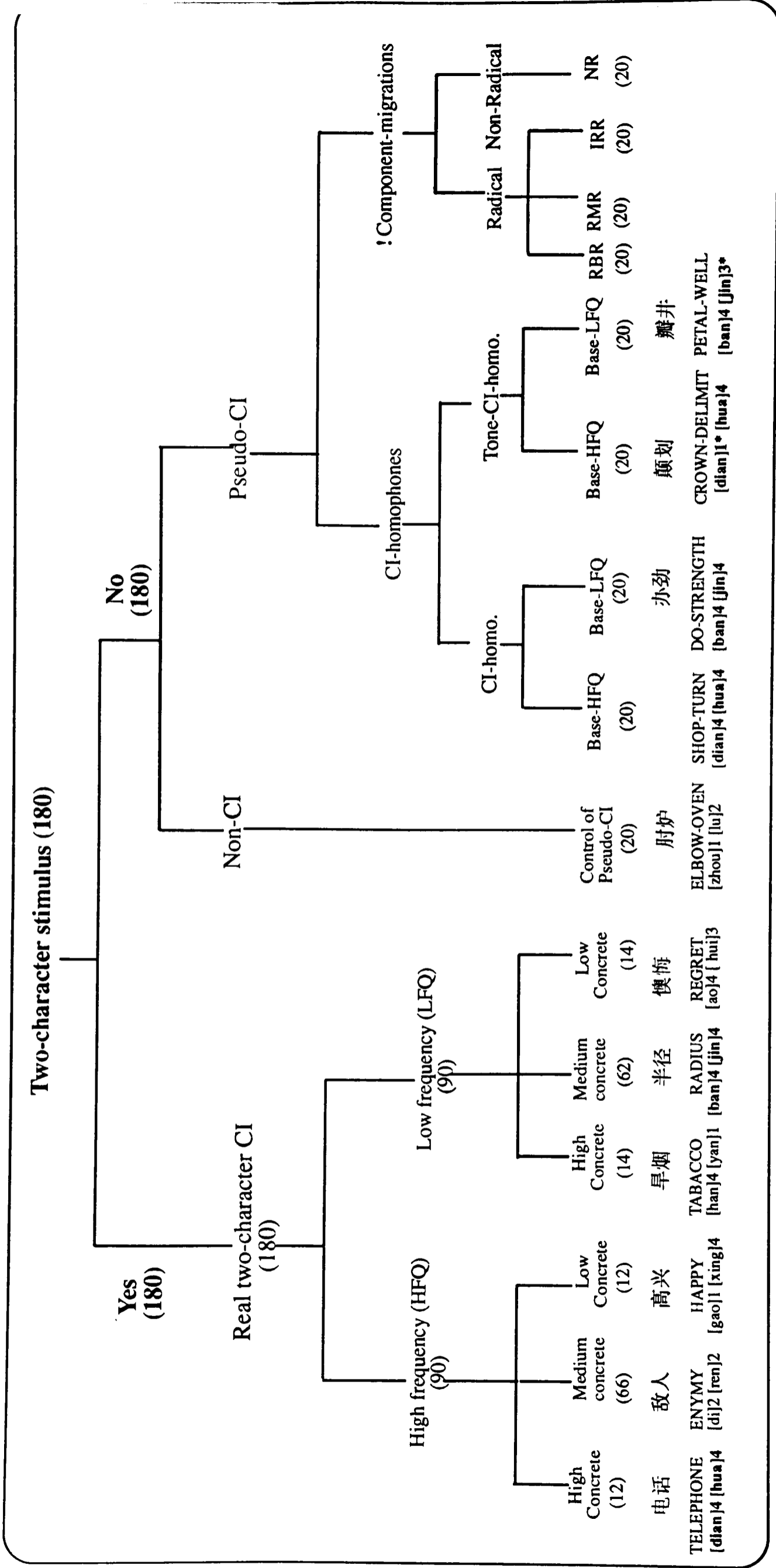
A total of 180 two-character real CI was selected, 90 from the high frequency band (ranging from 1000 to 8200 per hundred thousand) and 90 from the low frequency band (ranging from 23 to 99 per hundred thousand). CI frequency was based on a count of the occurrences of two-character CI in print. The CI frequency count used in this experiment was based on the updated word (ZI) and CI frequency dictionary, published by Beijing university in 1988. Two-character CI in each frequency group were further classified into Low Concrete (LC), non-specific and High Concrete CI (HC), according to the criterion described earlier. The high and low frequency groups contained respectively 12 and 14 HC CI, and another 12 and 14 LC CI (Figure 8.1b). The complete set of stimuli in the three categories of concreteness in each frequency group is given in Appendix II of this thesis.

2) Control for the pseudo-CI

Control non-CI were constructed by combining two legitimate Chinese words in a nonsense combination. Neither the two single words together nor any component of

Experiment 11

Task: Do the two characters together form a meaningful CI?



Note: * changed tone

Boldface print indicates 'Same' pronunciation

! See Figure 8.1a (in the text) for examples in component-migration conditions

Fig.8.1. Stimulus conditions in Experiment 11

them in conjunction with other whole word either look like or sound like a two-character CI. These non-CI were therefore used as the control for the other Pseudo-CI conditions described as follows:

3) Pseudo-CI (Homophone conditions)

CI-homophones: Two words, which together cannot be a meaningful two-character CI, but nonetheless sound like a real two-character CI (including tone). (See examples in Figure 8.1b.) CI-homophones were constructed by using homophones of the two words in the real two-character CI in this experiment. The base-CI for the CI-homophones were therefore the real two-character CI used in this experiment (see the example in Figure 8.1b). They were also divided into a high and a low frequency group, according to the frequency of the base-CI, i.e., the real CI used in this experiment.

Tone-changed CI-homophones: These were formed by changing the tone of one constituent word in the CI-homophones as described above. Either the tone of the first or the second word might be changed, and this was counterbalanced among the stimuli in this condition.

4) Component-migrated illusory-CI stimuli

The four subtypes of component-migrated illusory-CI stimuli (I-CI stimuli) were already described earlier, in the introduction of this chapter, and are also illustrated in Figure 8.1a.

Apparatus

The experiment was carried out on a Packard Bell IBM PC-AT286 monograph computer. A programme was designed to control the presentation of the stimuli, record the response time (RT) in msec, and the accuracy (error rate), and to calculate the mean value of RTs (of correct responses) and error rates in each condition. The programme was also connected with a Chinese word processing package-ET, to retrieve Chinese words. Two words were presented simultaneously side by side in white on a black background. Each word was 1 cm x 1 cm and the two words were spaced 1.5 cm centre to centre. Each pair of stimuli covered a visual angle of about 1° horizontally by 2° vertically from a viewing distance of 30 cm.

Procedure

The task was to judge whether or not the two words together formed a meaningful two-character CI. The Z and / keys, located in the left and right corner of the keyboard, were used for responses. They were labelled YES and NO in Chinese. Subjects were asked to respond as accurately and quickly as possible, using the index finger of each hand. All subjects were asked to use the left hand for "Yes" responses and the right hand for "No" responses.

Each trial began with a fixation cross, "+", sized 0.2 cm x 0.2 cm. The fixation cross appeared in the centre of the screen to indicate the position of the stimulus. It lasted 0.5 seconds and then disappeared. After a further 0.5 second interval, two words would appear immediately on the left and right side of the fixation point. They remained on the screen until a response key was pressed. Once a response key was pressed, the stimuli would disappear immediately from the screen. After a 0.5 second interval, the fixation cross " + " would appear again to start the next trial.

Written instructions were given on the screen, as well as oral explanations by the experimenter. There were 30 pairs of warm-up trials. At the end of this practice session, the mean reaction time and error rate was shown on the screen. If the mean error rate of the practice trials was equal to or larger than 20%, more practice using different stimuli was given until an accuracy of at least 85% was reached.

After every 120 trials, there was a break of 2 minutes or so, accompanied by a piece of quiet electronic music. After the experiment, the subjects were asked to report their strategies.

Results

1) CI-type effects

Table 8.1 presents the mean reaction times and error rates of the 16 Chinese subjects. An analysis of variance (ANOVA) with repeated measures and a within-subjects design was performed on both RT and error rate data. For RTs, both the within-subjects effect,

Table 8.1. "Yes" and "No" responses in Experiment 11: Mean reaction times (RTs) and error rates, with statistical analyses

CI-type	No. of stimuli	RT (msec.)		Error rates (%)	
		Mean	SD	Mean	SD
"Yes" response					
All	180	686	(71)	6.0	(2.3)
"No" responses					
Non-CI (NCI)	20	778	(74)	4.1	(4.2)
Pseudo-CI (PCI)	160	802	(101)	8.3	(7.1)
All	180	790	(87)	6.2	(5.6)
<i>ANOVA (2 x 3)</i>		<u>F value</u>	<u>P value</u>	<u>F value</u>	<u>P value</u>
CI-type		34.8	0.00	3.2	0.05
Between subjects		3.19	0.03	1.1	0.41
<i>Newman-Keuls test</i>		<u>Q value</u>	<u>P value</u>	<u>Q value</u>	<u>P value</u>
CI vs NCI		5.93	<0.01	1.47	>0.05
CI vs PCI		11.53	<0.01	1.30	>0.05
PCI vs NCI (control)		2.35	>0.05	3.21	<0.05

i.e., the types of CI, and the between-subjects effect were significant (CI-type: $F_{(2,47)}=34.8$, $P<0.001$; between subjects: $F_{(15,47)}=3.2$, $P<0.05$). For error rates, the effect of CI-type was also significant ($F_{(2,47)}=3.2$, $P=0.05$). In further analyses on both RTs and error rates, using Newman-Keuls Q tests, real CI were found to be judged significantly faster than pseudo-CI ($Q_{(2,3)}=11.53$, $P<0.01$) and non-CI ($Q_{(3,3)}=2.35$, $P<0.05$). Overall, the pseudo-CI were not significantly slower than the non-CI, i.e., the control items (Figure 8.1). However, this may be attributed to the unequal size of pseudo-CI (160) and non-CI (20) samples in this overall comparison.

On error rates, there was no significant difference between CI and non-CI nor between CI and pseudo-CI (Table 8.1). However, the pseudo-CI caused overall significantly higher error rates than non-CI, i.e., the control items ($Q_{(2,3)}=3.21$, $P<0.05$).

2) CI frequency and concreteness effects

Table 8.2 presents the mean RTs and error rates of the 16 subjects by CI-frequency and concreteness for the real two-character CI only. An ANOVA was performed on the RTs, to examine the main effects of CI frequency and concreteness. Both CI frequency and concreteness effects were highly significant in the ANOVA on RTs (*Frequency*: $F_{(1,15)}=121.9$, $P<0.001$; *Concreteness*: $F_{(2,15)}=12.7$, $P<0.001$). There was no significant interaction between the CI frequency and concreteness effects (Table 8.2). Both effects are illustrated in Figure 8.2, separately for RTs and error rates. As seen in Figure 8.2, high frequency CI were judged significantly faster than low frequency CI in all three semantic concreteness categories. The difference between high frequency and low frequency CI (including all three categories of concreteness) was highly significant on both RT ($t_{(15)}=12.76$, $P<0.001$) and error rates ($t_{(15)}=8.56$, $P<0.001$) by paired t tests. The concreteness effect was examined further by Newman-Keuls Q tests, in high and low frequency CI separately. For high frequency CI (HFQ), high concrete CI were judged significantly slower than both non-specific ($Q_{(2,3)}=3.63$, $P<0.01$) and low concrete CI ($Q_{(3,3)}=3.48$, $P<0.01$). However, there was no significant difference between low concrete words and non-specific CI (Table 8.2). For low frequency CI (LFQ), high concrete CI were significantly slower than low concrete ($Q_{(3,3)}=6.66$, $P<0.01$) and non-

Table 8.2. Effects of concreteness and CI-frequency: Mean reaction times and error rates of 16 native Chinese subjects in Experiment 11 with statistical results.

CI-type	No. of stimuli	RT (msec.)		Error rates (%)	
		Mean	(SD)	Mean	(SD)
High frequency CI (HFQ)					
High concrete	14	654.2	(78.9)	1.6	(3.3)
Non-specific	62	618.5	(62.1)	1.7	(1.8)
Low concrete	14	619.0	(63.2)	1.0	(2.8)
All	90	630.5	(68.0)	1.4	(2.6)
Low frequency CI (LFQ)					
High concrete	12	768.5	(74.2)	14.7	(10.3)
Non-specific	66	748.1	(89.6)	8.1	(4.9)
Low concrete	12	718.2	(95.5)	10.7	(8.7)
All	90	744.9	(86.4)	11.1	(8.0)
Statistical analysis					
<i>ANOVA (2 x 3)</i>		<u>F value</u>		<u>P value</u>	
Main effect					
Word frequency (W)		127.9		0.00	
Concreteness (C)		12.7		0.00	
Interaction					
W x C		1.58		0.22	
<i>Newman-Keuls test</i>		<u>Q value</u>	<u>P value</u>	<u>Q value</u>	<u>P value</u>
Concreteness in HFQ					
High vs Low		3.48	<0.01	0.62	>0.05
High vs Medium		3.63	<0.01	0.19	>0.05
Non-specific vs Low		0.09	>0.05	1.27	>0.05
Concreteness in LFQ					
High vs Low		6.66	<0.01	1.74	>0.05
High vs Medium		2.27	<0.05	3.23	<0.05
Non-specific vs Low		3.60	<0.01	1.84	>0.05

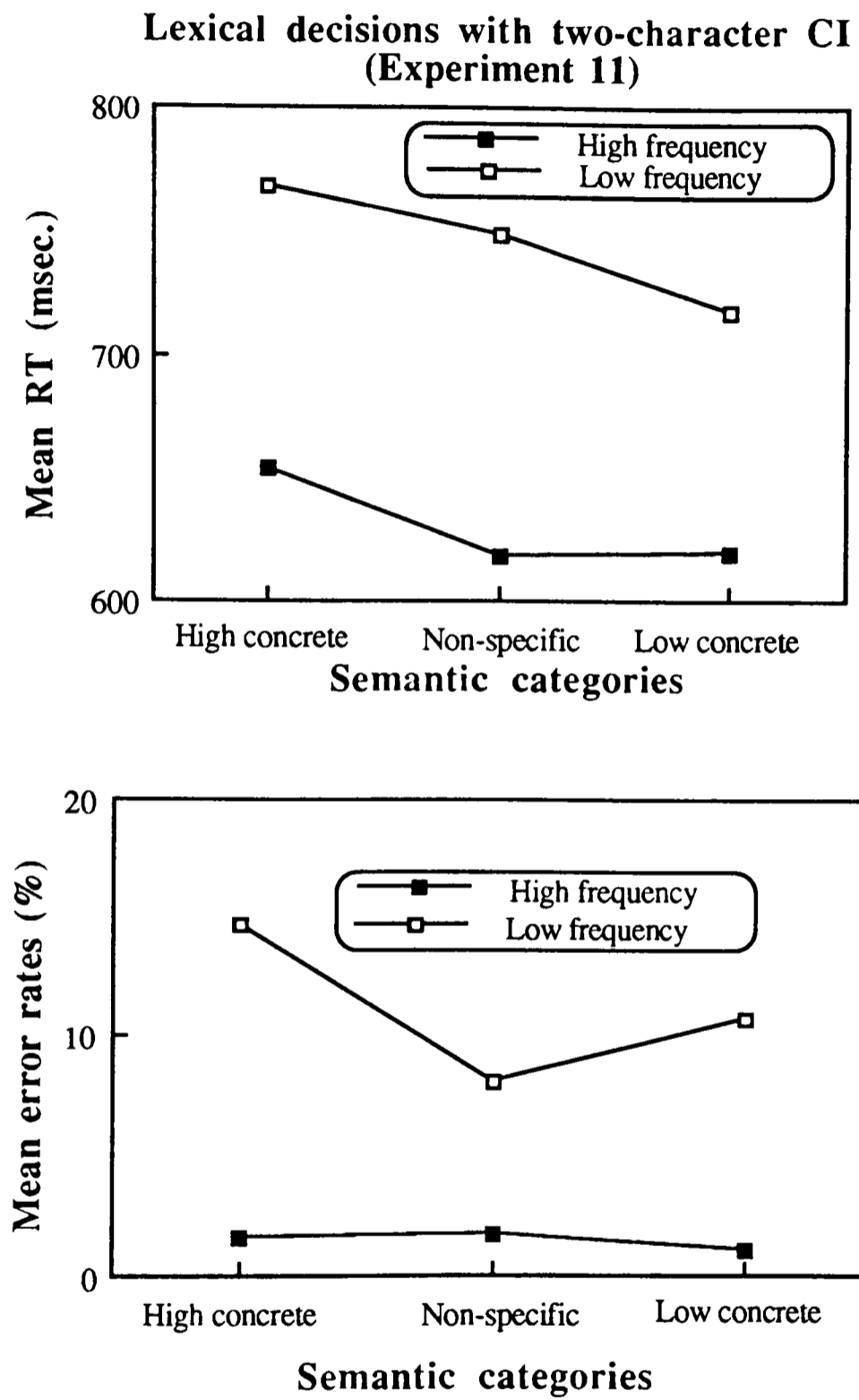


Figure.8.2. Mean RT and error rates by three semantic categories of concreteness for high and low frequency CI

specific CI ($Q_{(2,3)}=2.27$, $P<0.05$). Moreover, non-specific CI were also significantly slower than the low concrete CI ($Q_{(3,3)}=3.60$, $P<0.01$). For error rates, however, the effect of concreteness may be marked by a floor effect for high frequency CI, where error rates did not differ significantly between any two concreteness categories. For low-frequency CI, there was a significant difference in error rates between the high concrete and non-specific CI but not between non-specific CI and the low concrete CI (Table 8.2).

3) Pseudo-CI: analysis of the main stimulus-categories

We now turn to the “No”-response data, that is, the performance in response to the two main categories of two-character pseudo-CI: CI-homophones and the component-migrated illusory CI stimuli. Simple non-CI were used as controls for these two types of pseudo-CI. Table 8.3 presents the mean RTs and error rates for these three main categories of pseudo-CI, which are also illustrated in Figure 8.3. Increased RTs, or error rates, in rejecting the CI-homophone stimuli (as not real CI) would be evidence of phonological encoding in the CI-lexical-decision task. Increased RTs, or error rates, in responding to the component-migration (I-CI) stimuli may be taken, likewise, as evidence of lexical decomposition. An ANOVA was performed on the RT and error data. There was a significant effect of stimulus-type on both RTs ($F_{(2,15)}=4.4$, $P<0.05$) and error rates ($F_{(2,15)}=8.5$, $P<0.001$). There was a highly significant between-subjects factor, both in RT ($F_{(15,47)}=15.9$, $P<0.001$) and error rates ($F_{(15,47)}=2.4$, $P<0.05$), indicating substantial individual differences in this task. Further comparisons between these two categories of pseudo-CI and the control (non-CI) were examined by Dunnett t test. In comparison with the control, the component-migrated I-CI stimuli resulted in significantly slower and less accurate “No” responses than control (RT: $t_d=2.92$, $P<0.01$; error rates: $t_d=3.99$, $P<0.01$). Although the CI-homophones were generally slower and less accurate than the control (non-CI), the difference between CI-homophones and controls was not significant in either RT or error rates. This analysis thus fails to find consistent evidence of phonological recoding affecting the CI-lexical decision task. There does, however, appear to be some lexical decomposition within the single-character words.

Table 8.3 "No" responses in Experiment 11: Mean reaction times and error rates of 16 native Chinese subjects for different categories of pseudo-CI, with statistical analyses

Stimulus-types	No. of stimuli	<u>RT (msec.)</u>		<u>Error rates (%)</u>	
		Mean	SD	Mean	SD
Control (non-CI)	20	778	(74)	4.1	(4.2)
CI-homophones	80	790	(100)	5.9	(1.5)
Component-migration	80	814	(103)	8.5	(1.0)
<u>ANOVA (2 x 3)</u>		<u>F value</u>	<u>P value</u>	<u>F value</u>	<u>P value</u>
Stimulus-types		4.4	0.02	8.5	0.001
Between subjects		15.9	0.00	2.4	0.02
<u>Dunnett t test</u>		<u>td value</u>	<u>P value</u>	<u>td value</u>	<u>P value</u>
Component vs Control		2.92	<0.05	3.99	<0.01
CI-homo. vs Control		0.93	>0.05	1.07	>0.05
Component vs CI-homo.		1.99	>0.05	2.92	<0.05

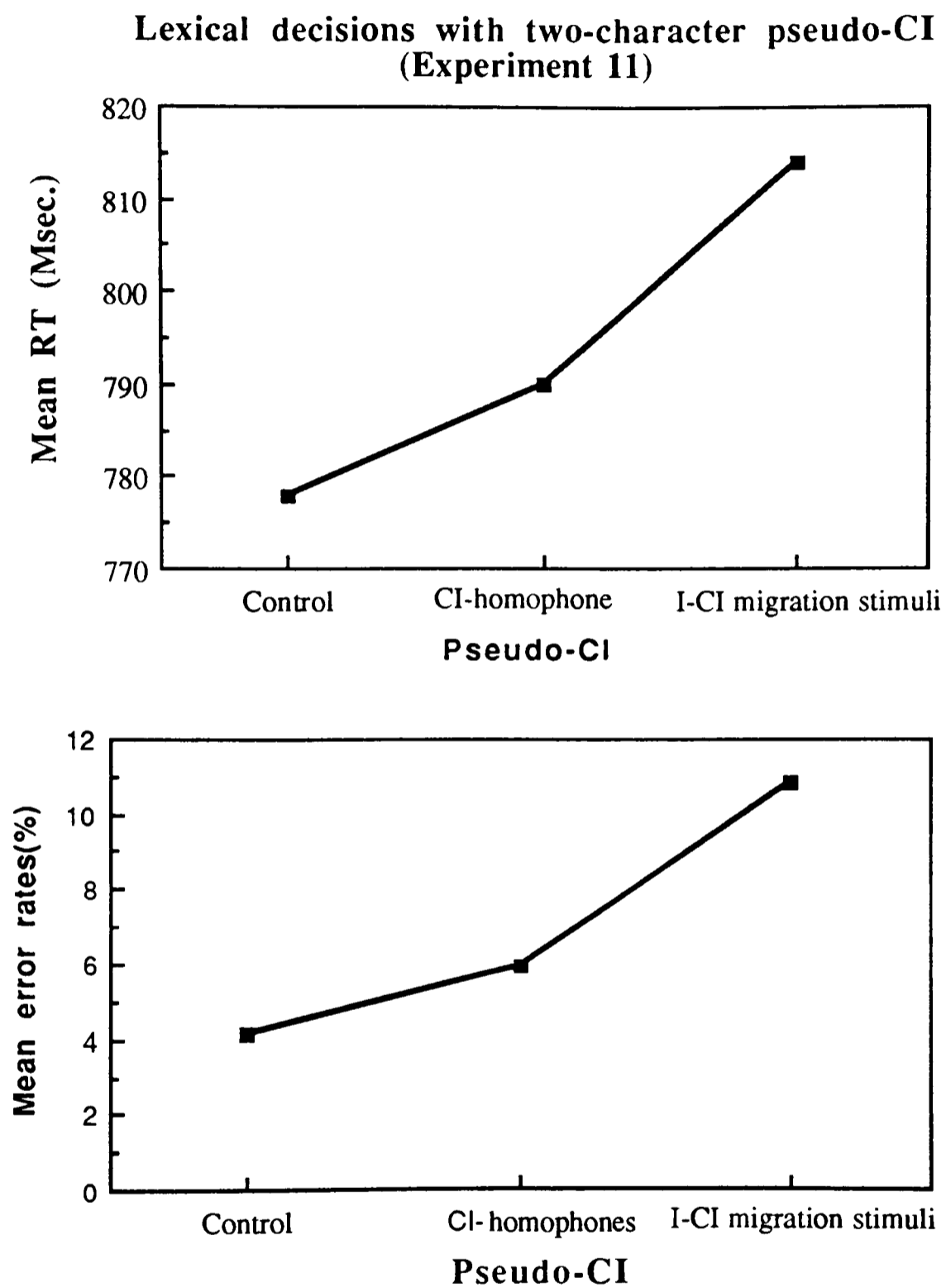


Figure.8.3. Mean RT and error rates of 16 Chinese subjects with two types of pseudo-CI

4) CI-homophones: effect of tone and base-CI frequency

Table 8.4 presents the mean RTs and error rates for the two types of CI-homophones derived from high and low frequency base CI. There were no significant differences between CI-homophones and tone-changed CI-homophones in either RTs or errors (Table 8.4). Base-CI frequency likewise had no significant effect on the speed or accuracy with which these stimuli could be rejected as not real CI (Table 8.4).

5) Component-migration effects (Lexical decomposition)

Table 8.5 presents the mean RTs and error rates for the four types of illusory-CI stimuli and the (non-CI) controls. An overall analysis by ANOVA with repeated measures on both RT and errors was performed. RT and error rates of lexical decisions on the four types of component-migrated illusory-CI stimuli are illustrated in Figure 8.4. The main effect of stimulus type was highly significant on both RT ($F_{(4,15)}=7.8, P<0.001$) and error rates ($F_{(4,15)}=3.3, P<0.001$). Compared to control, both the non-radical-migrated illusory CI (NR) ($t_d=3.60, P<0.01$) and the pseudo-radical-migrated illusory CI (R*) ($t_d=4.29, P<0.01$) were significantly slower by Dunnett t test. This effect was also consistent on error rates analysis (NR: $t_d=4.10, P<0.01$; R*: $t_d=3.98, P<0.01$). There was no significant difference between control and radical-migrated illusory-CI stimuli, regardless of the critical radical position in the two-character CI (RBR or RMR) (Table 8.5).

The differences among the four types of illusory-CI stimuli were also examined by Newman-Keuls Q test. Both irregular radical (R*) and non-radical component (NR) illusory-CI stimuli took significantly longer to reject as non-CI than the stimuli in which the lexical radical, in its regular position, formed the potential illusory-CI element (RBR and RMR). However, there was no significant difference between R* and NR (Table 8.5). Likewise, there was no significant difference between the two regular radical-migration conditions, with the critical radical either at the beginning (RBR) or in the middle (RMR) of a two-character CI.

Table 8.4. CI-homophone type and base-word frequency, in Experiment 11: Mean reaction times and error rates of 16 native Chinese subjects, along with statistical analyses

Base-CI frequency	CI-homophone types	No. of stimuli	RT (msec.)		Error rates (%)	
			Mean	(SD)	Mean	(SD)
Base-high frequency (HFQ)						
	CI-homophone.	20	792	(106)	5.6	(4.0)
	Tone-CI-homophone	20	778	(114)	6.9	(9.5)
	All	40	785	(110)	6.2	(6.7)
Base-low frequency (LFQ)						
	CI-homophone.	20	801	(109)	6.2	(8.1)
	Tone-CI-homophone	20	788	(88)	4.7	(6.4)
	All	40	794	(98)	5.4	(7.2)
<i>Paired t tests</i>			<u>t value</u>	<u>P value</u>	<u>t value</u>	<u>P value</u>
CI-homophone vs Tone-CI-homophone						
	Base-HFQ		0.83	0.42	0.64	0.53
	Base-LFQ		1.25	0.23	1.16	0.26
Base-HFQ vs Base-LFQ						
	CI-homophones		0.90	0.38	0.37	0.72
	Tone-CI-homophone		0.64	0.53	1.20	0.25

Table 8.5 Four types of "illusory CI" stimuli for "No" responses in Experiment 11: Mean reaction times and error rates, with statistical analyses

Component-associated Pseudo-CI	No. of stimuli	RT (msec.)		Error rates (%)	
		Mean	SD	Mean	SD
Control (Non-CI)	20	778	(74)	4.1	(4.2)
Component-migration					
RBR	20	799	(108)	8.8	(11.0)
RMR	20	781	(95)	5.3	(4.3)
"Pseudo-radical" (R*)	20	844	(108)	14.4	(11.5)
NR	20	833	(122)	14.7	(13.0)
All pseudo-CI	80	814	(108)	10.8	(9.9)

<i>ANOVA (2 x 3)</i>	<u>F value</u>	<u>P value</u>	<u>F value</u>	<u>P value</u>
Stimulus-types	7.8	0.00	3.3	0.00
Between subjects	16.9	0.00	7.3	0.00

<i>Dunnett t test</i>	<u>td value</u>	<u>P value</u>	<u>td value</u>	<u>P value</u>
RBR. vs control	1.33	>0.05	1.81	>0.05
RMR vs control	0.16	>0.05	0.48	>0.05
R* vs control	4.29	<0.01	3.98	<0.01
NR vs control	3.60	<0.01	4.10	<0.01

<i>Newman-Keuls Q</i>	<u>Q value</u>	<u>P value</u>	<u>Q value</u>	<u>P value</u>
RBR vs RMR	2.08	>0.05	1.81	>0.05
R* vs NR	0.20	>0.05	1.00	>0.05
RBR vs R*	4.58	<0.01	3.72	<0.05
RMR vs R*	6.56	<0.01	4.94	<0.01
RBR vs NR	2.84	<0.05	4.09	<0.01
RMR vs NR	6.56	<0.01	5.95	<0.01

RBR=Radical in the beginning of two-character CI

RMR=Regular radical in the middle of two-character CI

R*=A component in a radical form but in a non-radical position in a word and serves as a non-radical component of the word

NR=Non-radical component (in the beginning or middle of two-character CI)

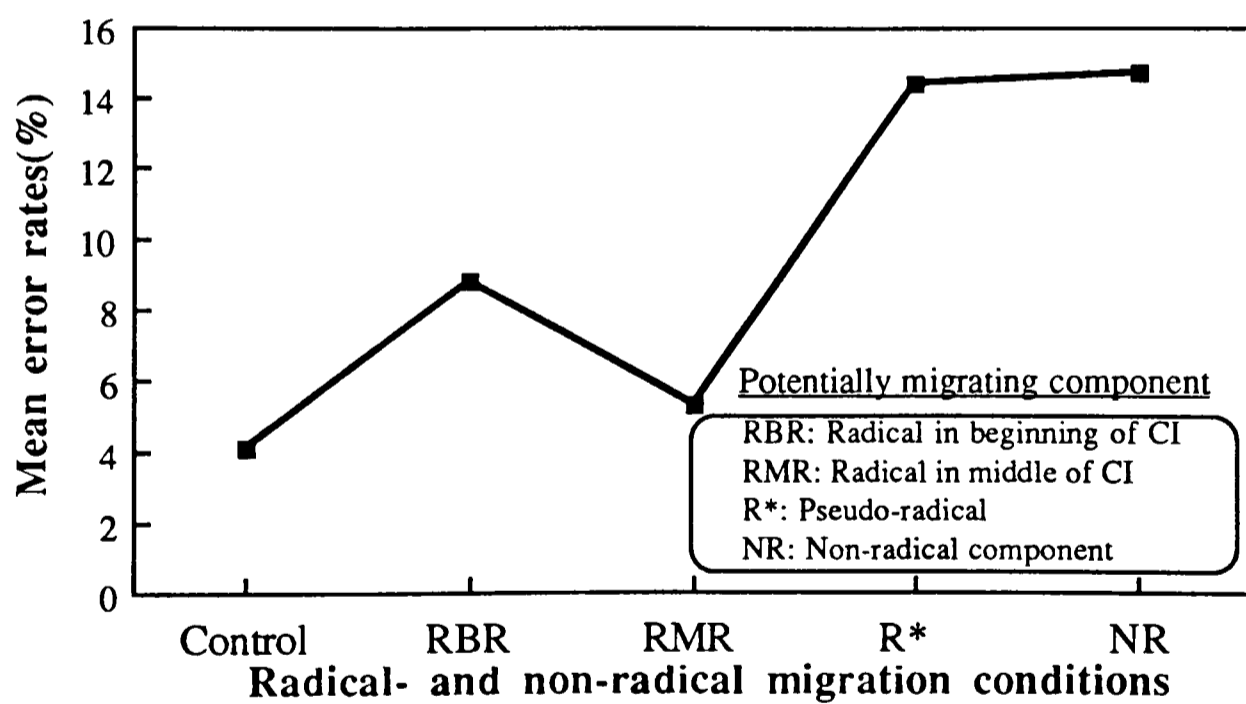
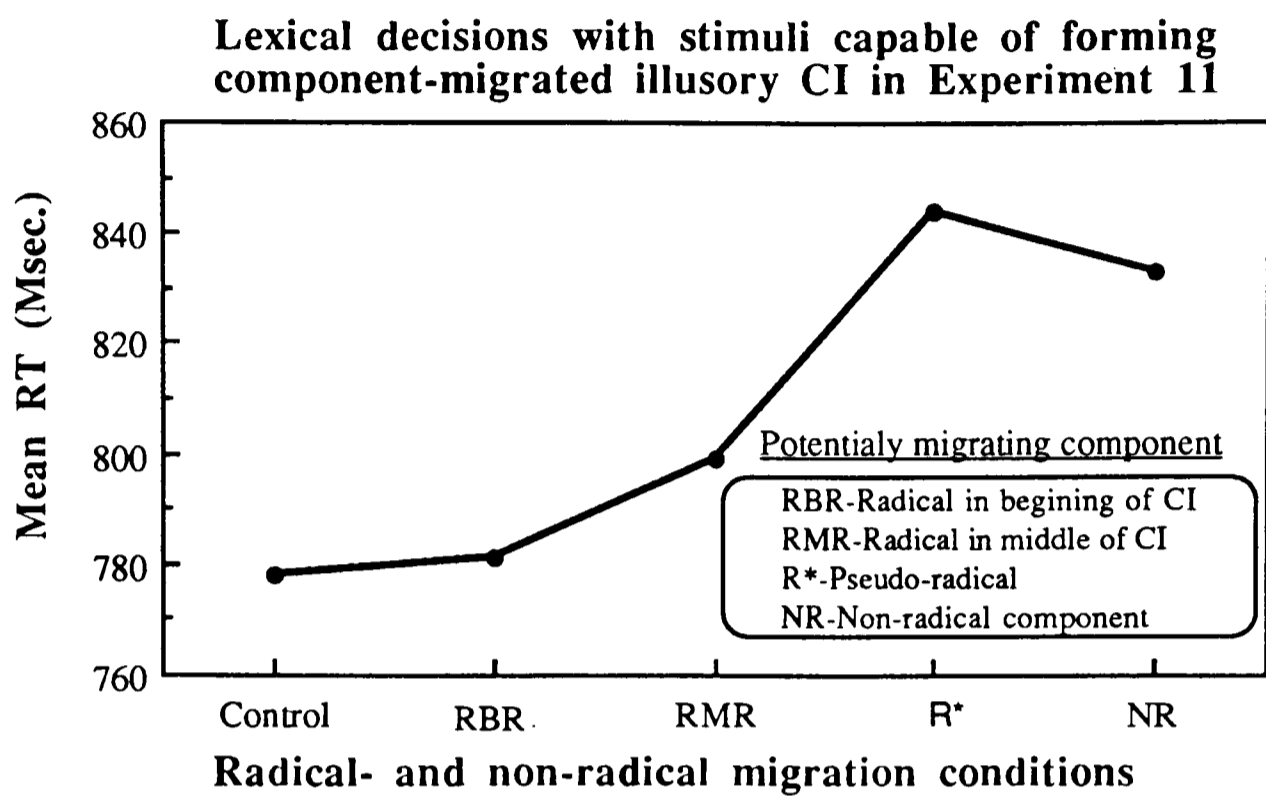


Figure.8.4. Mean RT and error rates in radical- and non-radical migration conditions

6) The speed and accuracy of lexical decision and its association with reading strategy indices

The speed and accuracy of lexical decision ('yes' response) may be taken as a simple indicator of a subject's reading efficiency, in recognizing two-word CI. Simple regression tests were conducted between the four recoding strategy indices, described in the 'Background and Rationale' to this Chapter, and mean RT and error rate of lexical decision in response to the 180 real CI stimuli, for each subject. The results of these regression analyses are summarized in Table 8.6. As a preliminary, it was established that there was no significant speed-accuracy trade-off between lexical decision RT and accuracy on the 180 real CI ($r^2=0.00$, $P=0.99$). Reading speed (RT) and accuracy were hence treated as two independent variables, and were examined separately.

As indicated in Table 8.6a, individual subject's reading speed was significantly associated with each of these different recoding strategy indices, except with the orthographic-lexical-frequency-dependence index. The latter index of orthographic-lexical dependence, inferred from the susceptibility of each subject to CI frequency effects, marginally affected speed of decision for the low-frequency CI. However, the corresponding frequency-dependence index based on *accuracy* data was highly predictive of lexical decision accuracy, but again for the low-frequency CI only ($r^2=0.76$, $P<0.001$, Table 8.6b). Indeed, it was the only one of these indices to show such an association. The reading speed (i.e., lexical decision time) for two-character CI was also significantly associated with both the semantic and phonological processing indices. Faster reading was associated with a higher semantic recoding strategy index but a lower phonological recoding strategy index and this was especially evident in responding to the low frequency CI (Figures 8.5 and 8.6). In other words, (i) the greater the subject's dependence on semantic recoding, indexed by his/her susceptibility to the CI-concreteness effect, the *faster* his/her overall lexical decision times; whereas (ii) the greater the subject's dependence on phonological recoding, indexed by his/her susceptibility to the CI-homophone interference effect, the *slower* his/her overall lexical decision times. Finally (iii) both slower and less accurate lexical decision was significantly associated with higher indices of lexical decomposition in particular the

Table 8.6a. The associations between each reading-strategy index and lexical decision times for 180 real CI in Experiment 11

	Ortho-lex. FQ		Semantic		Phonological		Lexical Decomposition			
	CI-FQ)		(Concreteness)		(CI-homo.)		(Radical)		(Non-radical)	
	r ²	P	r ²	P	r ²	P	r ²	P	r ²	P
CI (180)	0.10	0.23	0.24	0.05	0.37	0.01	0.27	0.04	0.37	0.01
HFQ (90)	0.00	0.79	0.21	0.07	0.34	0.02	0.18	0.10	0.34	0.01
LFQ (90)	0.23	0.06	0.25	0.05	0.35	0.01	0.34	0.02	0.36	0.01

Table 8.6b The associations between each reading strategy indices and accuracy (%) of lexical decisions for 180 real CI in Experiment 11

	Ortho-lex. FQ		Semantic		Phonological		Lexical Decomposition			
	CI-FQ)		(Concreteness)		(CI-homo.)		(Radical)		(Non-radical)	
	r ²	P	r ²	P	r ²	P	r ²	P	r ²	P
CI (180)	0.01	0.67	0.16	0.12	0.15	0.31	0.57	0.02	0.00	0.81
HFQ (90)	0.03	0.49	0.01	0.69	0.07	0.49	0.21	0.21	0.00	0.82
LFQ (90)	0.76	0.00	0.17	0.11	0.13	0.35	0.51	0.03	0.01	0.84

Note: Correlations of CI & Non-CI: RT :r²=0.4,P=0.0086; Error rates:r² =0.085,P=0.2722

Correlations of RT & Error rates for CI: r² =2.314E-6, P=0.99 and non-CI: r² =0.038, P=0.47

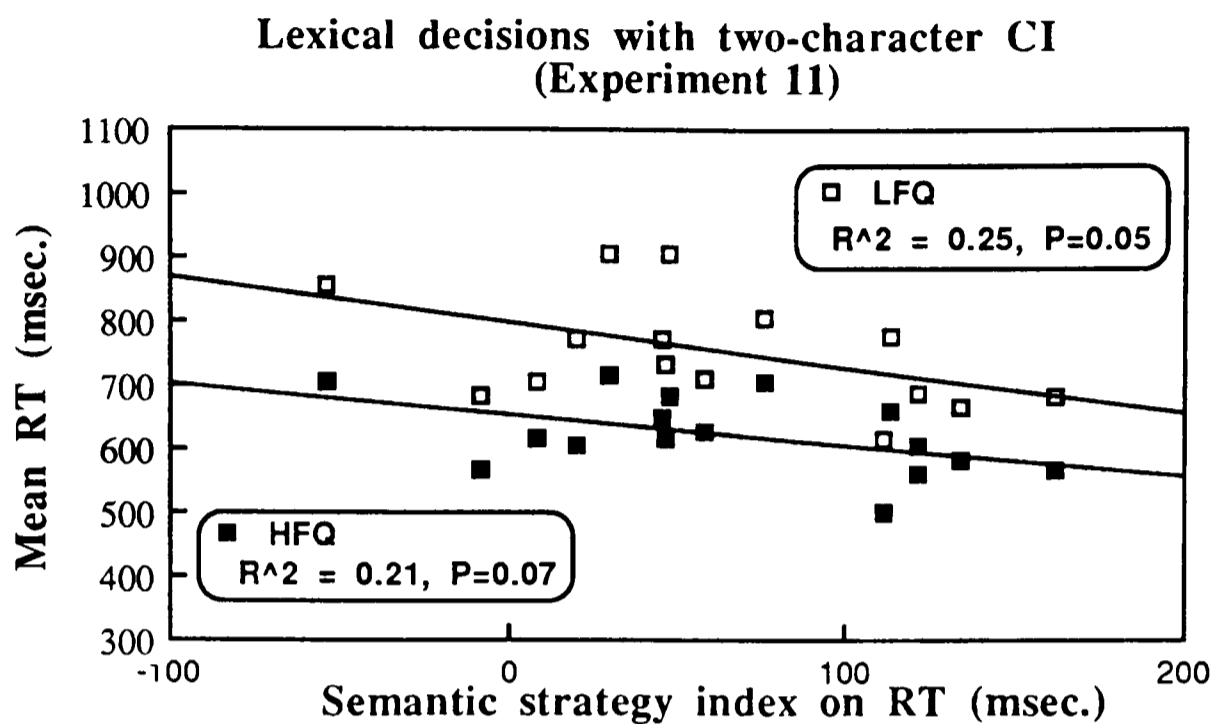


Figure 8.5 The correlation between RT of high or low frequency CI and semantic strategy index, over 16 subjects

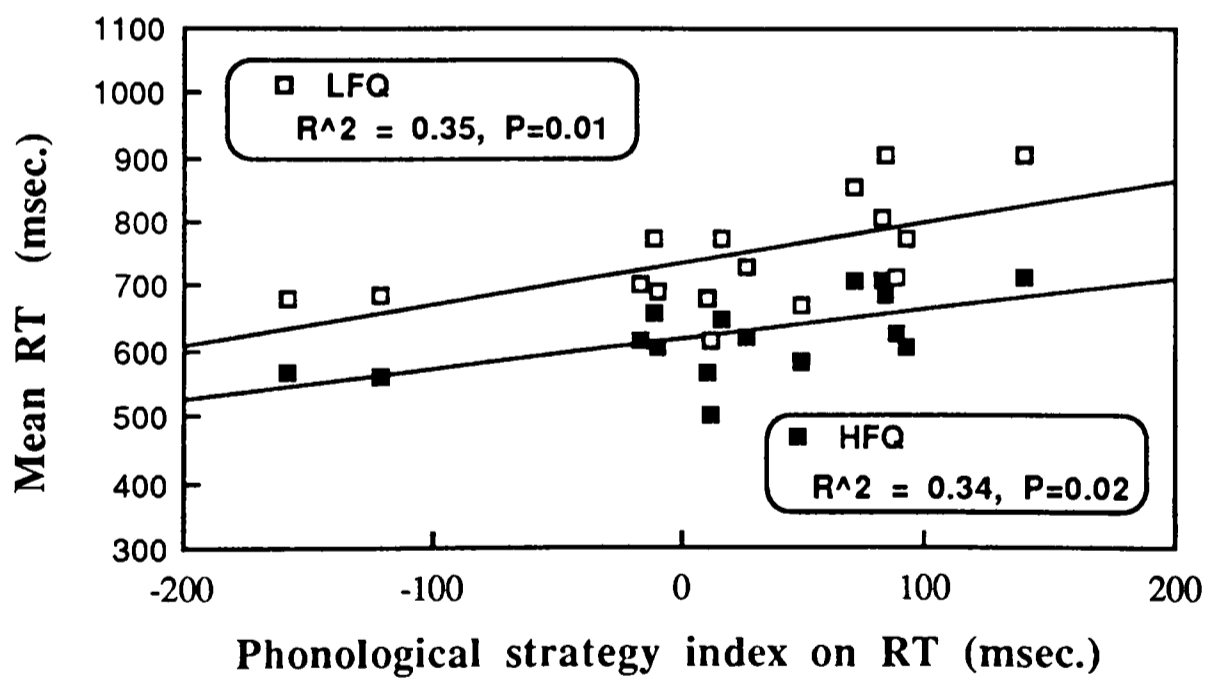


Figure 8.6 The correlation between mean RT of high or low frequency CI and phonological strategy index, over 16 subjects

radical-migration index (see Table 8.6). This was particularly evident in responding to low frequency CI.

7) Reading strategies: Interassociations among orthographic, semantic and phonological processing indices

Table 8.7 presents the results of the correlation and regression analyses *between* the various processing indices. Table 8.7a shows the regression analyses based on the RT data; 8.7b shows the equivalent analyses based on errors.

There was a significant inverse association between semantic and phonological processing in the lexical decision time for two-character CI ($r^2=0.31$, $P=0.02$). The inverse relationship is illustrated in Figure 8.7. The phonological recoding index was also highly significantly associated with both the radical and the non-radical migration indices, in both RT and error rates. The strong, positive correlations between the phonological-recoding-strategy index and the radical and the non-radical-migration indices are illustrated in Figures 8.8a & 8.8b. The indices of radical and non-radical migration were also positively correlated with each other (Table 8.7).

Discussion

1) Lexical decision with two-character CI

The results of this experiment can be summarized as follows: First, lexical decisions with two-character CI were significantly affected by the lexical familiarity of the two-character CI as a whole. High frequency CI were judged substantially faster and more accurately than low frequency CI. Moreover, the concreteness categories also had a significant, though lesser, influence upon lexical decision speed and accuracy. Counterintuitively, however, for both high and low frequency items, high concrete CI (typically object names) were responded to more slowly than the low concrete CI ('mental' words).

Second, this experiment suggests that phonological recoding may be unnecessary for lexical decision on two-character CI. The CI-homophones were not significantly slower than the control. The difference between the CI-homophones and the tone-changed CI-

Table 8.7a. The association among various reading strategy, based on RTs, in Experiment 11, with regression analyses.

	Semantic		Phonological		Lexical Decomposition			
	(Concreteness)		(Ci-homo.)		(R-migration)		(NR-migration)	
	r^2	P	r^2	P	r^2	P	r^2	P
Orth.-Lex. FQ (CI-FQ)	0.05	0.38	0.03	0.50	0.23	0.06	0.03	0.50
Semantic (<i>Concreteness</i>)			0.31	0.02	0.26	0.04	0.09	0.26
Phonological (<i>CI-homo.</i>)					0.68	0.00	0.42	0.01
Lexical Decomposition (<i>Radical</i>)							0.38	0.01

Table 8.7b. The association among various reading strategies, inferred from different processing indices in Experiment 11, with regression analyses on error rates

	Semantic		Phonological		Lexical Decomposition			
	(Concreteness)		(Ci-homo.)		(R-migration)		(NR-migration)	
	r^2	P	r^2	P	r^2	P	r^2	P
Ortho.-Lex. FQ	0.02	0.59	0.01	0.78	0.00	0.99	0.00	0.98
Semantic (<i>Concreteness</i>)			0.06	0.52	0.00	0.88	0.00	0.87
Phonological (<i>CI-homo.</i>)					0.60	0.01	0.72	0.00
Lexical Decomposition (<i>Radical</i>)							0.29	0.13

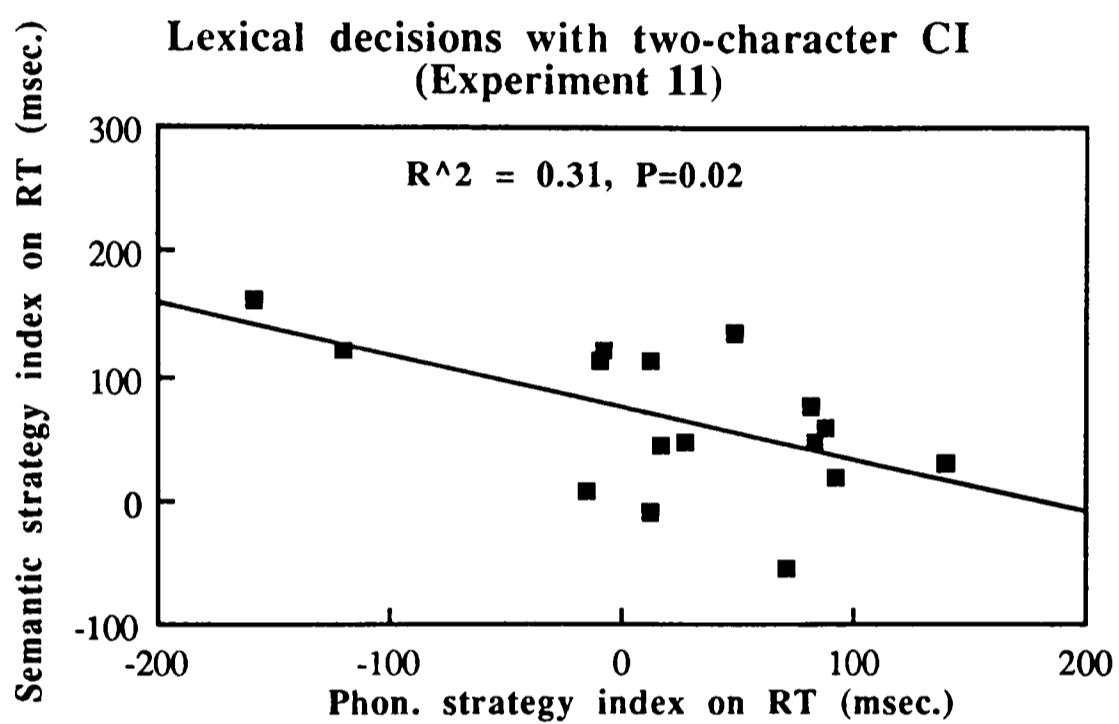


Figure 8.7 The correlation between semantic recoding index and phonological recoding index, over 16 subjects in the lexical decision task

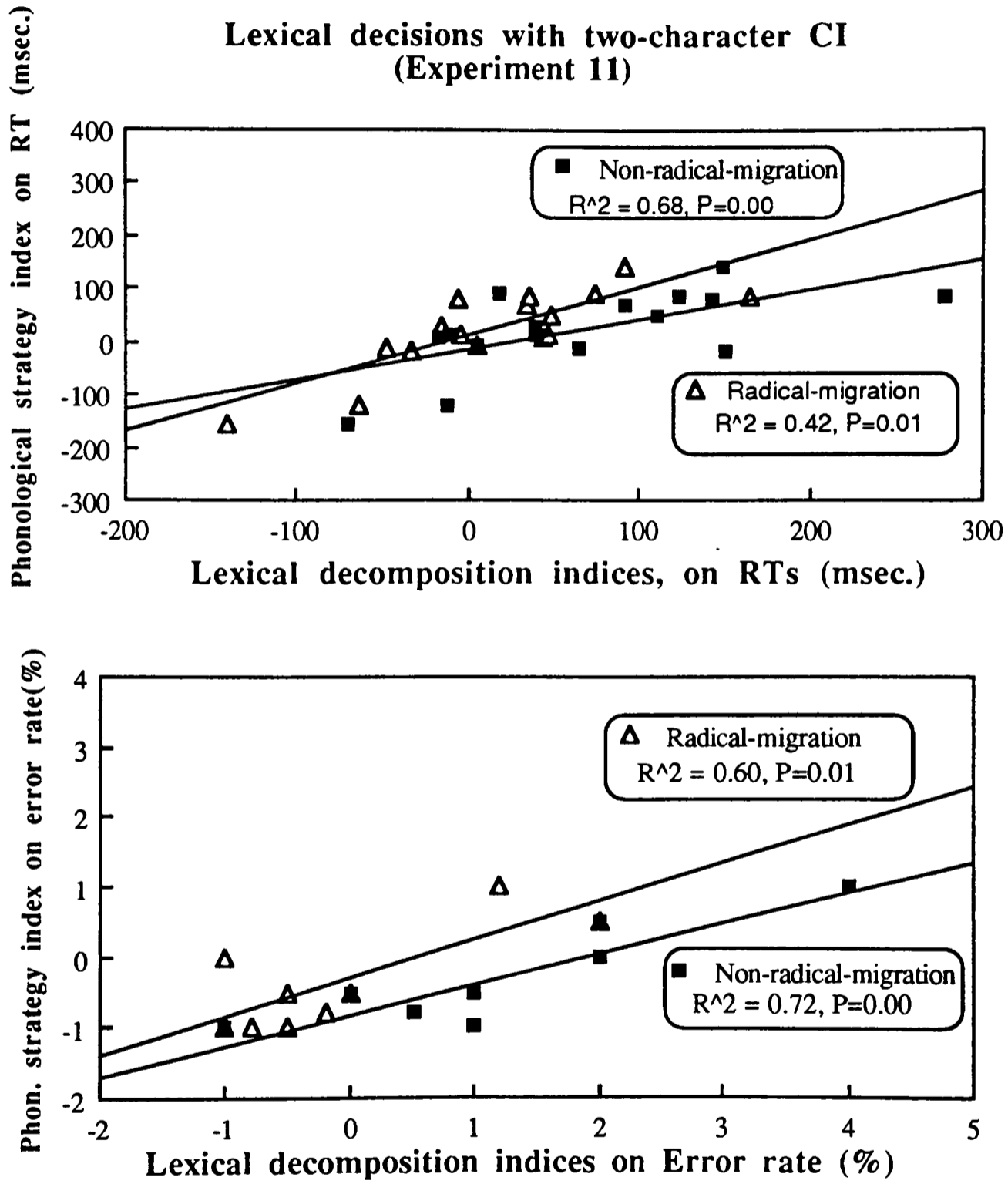


Figure 8.8 The correlation between phonological recoding index and two types of lexical decomposition indices on RT and error rates, over 16 subjects

homophones was also not significant. However, the incidence of phonological recoding appeared to be an individual-difference variable, related in a systematic way to other aspects of the lexical decision task, in particular to lexical decomposition strategies (See below).

Third, there were striking differences between the effect of the four different types of illusory-CI stimuli (Figure 8.4). With the lexical radical in its legitimate position in a single-character word conditions (RBR and RMR), the radical itself showed no systematic tendency to form an illusory CI in conjunction with the other whole word. In contrast, orthographic component appearing *in the non-radical position* in a single-character word, whether this was a regular non-radical component (NR) or in fact a lexical radical but in an irregular (non-radical) position (IRR), showed a very marked tendency to form "illusory conjunctions" with the other word, as indicated by a more-than-threefold increase in false 'Yes' responses, and increased correct 'No' RTs.

The implications of each of these results are discussed further in the following three sections, (a), (b) and (c), below.

a) Orthographic-lexical processing: effects of CI-frequency

Efficiency of orthographic-lexical processing was inferred from the effect of CI frequency effects on lexical decisions of two-character CI. The word frequency effect is a powerful and ubiquitous effect, which has been reported in many different tasks, such as tachistoscopic report (*Humphreys, Besner & Quinlan, 1988; Jacoby & Dallas, 1981*), speeded naming (*Forster & Chambers, 1973*), lexical decision (*Rubenstein, Garfield & Millikan, 1970; Forster & Chambers, 1973*), identification of words in background noise (*Savin, 1963*), and auditory lexical decision (*Slowiaczek & Pisoni, 1986; Taft & Hambly, 1986; Whaley, 1978*). Word frequency is a prominent feature of almost all the word recognition models in the literature, including the recent connectionist accounts of word recognition (*McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989*). The underlying assumption for the word frequency effect in almost all existing models, except for Seidenberg and McClelland (1989), is that the effect is directly "lexical". However, the unitary lexical account for the word frequency effect has been questioned

by some studies in English. In contrast to the lexical account, orthographic familiarity is considered to provide an important non-lexical account for the word frequency effect in word recognition (*Besner & McCann, 1987*). According to some studies, the effects of the frequency of a written symbol on word recognition may occur at three loci (*McCann and Besner, 1987; McCann, Besner and Davelaar, 1988*). These are, 1) Visual familiarity discrimination (*Balota & Chumbley, 1984; Besner, 1980, 1983, 1984; Besner et al., 1984; Besner & McCann, 1987*); 2) Connection to or involvement of the semantic system (*e.g., Morton, 1982*); 3) The lexical-lexical connection (*McCann & Besner, 1987, McCann, Besner & Davelaar, 1988*).

In this experiment, the CI frequency effect was examined in both real CI and CI-homophone conditions. (The real CI were also the base-CI of the CI-homophones in this experiment.) Given that the lexical decisions on two-character CI are based on the individual words (as discussed later in detail), the underlying mechanism for the CI frequency-effect is unlikely to be radically different from the word frequency effect. As with the word frequency effect, the CI-frequency effect may be considered to be related mainly (though not solely) to the visual familiarity of the two-character item (but maybe to its meaning as well). CI frequency had a significant effect on decisions on the real CI, but not on the CI-homophones. This is consistent with studies of the word frequency effect on naming pseudohomophones in English. The present study favours the view that the word frequency effect is sensitive to the visual lexicon but not to the phonological lexicon (*McCann & Besner, 1987; McCann, Besner & Davelaar, 1988*).

b) Phonological recoding

The role of phonological recoding in reading is an old but still controversial issue (*Hung & Tzeng, 1981, for a review*). Some earlier studies demonstrated that phonological recoding is involved in some aspects of reading written words (*Corcoran, 1967; MacKay, 1972; Baddeley & Hitch, 1974; Pynte, 1974; Massaro, 1975*). The role of phonological recoding has been reported in lexical decision tasks on pseudohomophones in English (*Rubenstein, Lewis & Rubenstein, 1971; Meyer, Schvaneveldt & Ruddy, 1974*). By contrast, other well-known studies demonstrated that

phonological recoding is not necessary for reading English (*Baron, 1973; Kleiman, 1975*). Studies of homophone and pseudohomophone effects in English have suggested that phonological recoding is not compulsory but optional (*Davelaar, Coltheart, Besner & Jonasson, 1978*). However, it may be important to distinguish between the activation and the utilisation of phonological information. According to some parallel distributed processing models, the activation of visual, semantic and phonological information are automatic and parallel processes. However, the utilisation of the activated information may be subject to reading strategies, which in turn can be influenced by many factors. The presence or absence of a homophone or pseudohomophone effect in different experiments may reflect the utilisation, rather than simply the activation of the information in different reading tasks. Hence, the presence or absence of homophone and pseudohomophone effects may be subject to many factors, such as the nature of the stimulus ensemble (e.g., whether they include pseudowords), the nature of the task (whether it demands overt phonological output such as naming, vs. lexical decision; or whether or how far it demands working memory, e.g., single word reading or reading a sentence or text), and the linguistic status of the individuals (skilled adults vs. children learning to read). Phonological processing as an individual reading strategy, as already pointed out by *Hung & Tzeng (1981)*, is unlikely to be determined simply by differences in orthographies. As to the role of phonological recoding in reading Chinese, the traditional view of a picture-like direct access process (*Barron & Baron, 1977*) is based on the supposed nature of the so-called ideographic or lexicographic Chinese writing system, which has been re-examined in the present study (Chapter 2). This traditional view of direct access has recently been modified. Many studies on Chinese have demonstrated that phonological information in Chinese, as in English, is automatically activated (*Hung & Tzeng & Hung, 1977; Tzeng & Hung and Wang, 1977; Baddeley, 1979; Tzeng, 1980; Zhang & Simon, 1985; Perfetti & Zhang, 1991*). In the previous chapters, with single character Chinese words, homophone effects were absent in the semantic comparison of Chinese words (Experiment 8). In this experiment with two-character CI, there was likewise no significant CI-homophone effect on the lexical decision task. This suggests that, at least for skilled adult readers of Chinese, phonological processing may be unnecessary. However, it is noteworthy that the CI-

homophone effect showed significant individual differences. The present study favours the view that the use of phonological recoding in lexical decision may be regarded as an optional rather than a compulsory reading strategy, which may be subject to individual control (Tzeng, 1977; Hung & Tzeng, 1981). The implications of phonological strategies for reading Chinese are discussed later in more detail.

c) Lexical decomposition processes

In Chapters 4 to 7, it was reported that the visual, semantic and phonological comparison task were each influenced by manipulations of individual components of single-character words. That is, those experiments found clear evidence of lexical decomposition. First, the visual analysis of Chinese words was shown to be based on the orthographic unit (i.e., the stroke-pattern), rather than on individual strokes (Experiments 1, 2 & 3). Second, for the 90% or more of Chinese words with more than one unit, the evidence suggested that one component, the lexical radical, receives attentional priority in semantic judgements by skilled Chinese readers (Experiments 6 & 9). The non-radical component as a whole likewise receives attentional priority in the phonological recoding of single-character words (Experiments 6 & 9). The biasing effect in favour of the non-radical component in the phonological comparison task (Experiment 6) was stronger than that of the lexical radical in the semantic comparison task (Experiment 8). In the present experiment, the decisions on two-character CI were also shown to be affected by the illusory migration or recombination of components of single-character words. The subjects, the skilled Chinese readers, were significantly slower and less accurate in judging the two-character CI, when a component of one of the words in conjunction with the other whole word constituted a meaningful two-character CI, although the two words as a whole did not. This result suggests that the recognition of two-character CI is based on the analysis and lexical decomposition of the individual (single-character) words, even though the two-character CI are standardly used as a fixed expression in the language. As the two-character CI are constructed out of individual words, it is conceivable that they are represented by an associating connection between the representations of individual words in the mental lexicon. In this case, the multiple-character CI in Chinese should not be regarded as independent 'words' of Chinese. The

functional unit equivalent to words in English is formed by single character words, the so-called 'characters'.

The "illusory CI" stimuli were subdivided into four types. Only two of these caused significant interference in the lexical decision task: (i) the NR condition, in which the non-radical component of one of the two words forms a potential illusory CI in conjunction with the other word in the pair; and (ii) the R* condition, in which the critical component that forms a potential illusory CI, though a member of the set of 189 radicals in terms of its graphic form, is in the non-radical position in the word: i.e., it is a "pseudo-radical". On the definition of the lexical radical proposed in this study (Chapter 2), it is a non-radical component, just like the critical element in the NR condition. In fact, these two types of I-CI stimuli, NR and R*, resulted in equally large "illusory conjunction" effects. Neither of the other two conditions, in which the critical component was in the radical position (i.e., was the true lexical radical for that word) showed any systematic tendency to form illusory conjunctions with the other word in the pair. In spite of this very clear result in the overall group data, effects of the 'radical-migrated' stimuli, as well as the 'non-radical-migrated' stimuli, showed substantial individual variation, as seen in the radical-migration index. (These were discussed further under reading strategies, below.)

This pattern of results with the "illusory CI" stimuli is open to one clear, and very interesting, interpretation. The argument is as follows. First, the occurrence of "component-migration" interference effects constitutes rather compelling evidence for lexical decomposition of the individual words in the two-character CI. However, of the resulting, orthographic components, *only* components occupying the *non-radical position* in a word appear free to "migrate" to form an illusory conjunction with the neighbouring word. This seems to imply a strong asymmetry in the lexical decomposition process: the non-radical component as a whole is segmented away from (or with respect to) the lexical radical, not vice versa. The lexical radical, as the hierarchically dominant or 'essential' element, remains fixed as the base. A hierarchically ordered, or asymmetrical segmentation process of this kind would also make good sense for the purpose of deriving a pronunciation for the word. (Even though, as the CI-homophone conditions

confirm, phonological recoding is not used to any significant extent in the lexical decision task with two-character CI.) The pronunciation of a regular compound word is specified by the non-radical components as a whole. But, to locate (and, then segment) the non-radical component, it is arguable that the reader must first locate the lexical radical. The remainder of the character must then be segmented from the identified lexical radical.

In summary, the results from this experiment are in harmony with the general view that silent reading of Chinese is based on the visual-semantic route, the so-called lexical route, at least for skilled adult readers. Phonological recoding is not necessarily used for skilled adult reading of two-character CI. However, it must be emphasized again that the utilization of information is a different question from its activation. Different codes may be activated independently in parallel. However, the use of a particular type of information may be optional, and subject to strategic-attentional control. In what follows these effects are discussed in terms of different reading strategies and their implication for reading efficiency.

2) Relationships between different reading strategies and their implications for reading efficiency

The second strand of investigation in this experiment concerned the relationship among different reading strategies and their implication for reading efficiency. Indices of four different reading strategies were obtained, based on each individual's performance. Unlike many studies of the relationships between reading strategy and reading skill in English, which have used good and poorer readers, as assessed by standard reading scores (*e.g.*, Barron, 1981, Bryant, 1983), this experiment used only skilled native readers, who all had a university degree in Chinese. The association between different reading strategies and reading efficiency may be understood in the light of the strategies themselves, rather than as constant for any one individual. Reading strategies can, of course, vary consistently from one individual to another. However, for the same individual, they may also vary over time and in different contexts. Unless some disturbance precludes the use of a certain strategy (such as patients with word-meaning

blindness may not be able to use the semantic strategy), a reading strategy is not necessarily constant for an individual.

First, the semantic recoding index varied inversely with RT, i.e., a "semantic" strategy was correlated with faster performance. That is, a "semantic" strategy was correlated with faster performance. By contrast, the phonological recoding index (indicating a "phonological" strategy) was associated with slower performance. This suggests that at least for lexical decision with two-character CI, a phonological strategy is less efficient.

The phonological and semantic strategies were themselves inversely related, suggesting that some features of processes involved in phonological recoding conflicts with efficient semantic recoding (or possibly vice versa). It will be argued, below, that this conflicting feature is lexical decomposition.

With regard to lexical decomposition, the radical-migration and non-radical-migration indices were strongly positively associated with the phonological recoding index. Moreover, both of these lexical decomposition indices were also correlated with slower lexical decision performance with the two-character CI. It seems clear that lexical decomposition (viz; segmenting the non-radical component away from the lexical radical) is associated with a phonological recoding strategy. Less clearly, but very plausibly, lexical decomposition of this kind is not needed for, and may be partially in conflict with, semantic recoding. The plausibility of this inference derives from the fact that the lexical radical (semantic radical) cannot specify the meaning of the whole word in which it appears (Chapter 7). In order to specify the full meaning of a word, both orthographic components (lexical radical and non-radical component) are needed together. It may be that, for skilled readers, lexical decomposition into these two orthographic components is not helpful for the semantic recoding process. Further, a phonological recoding strategy, requiring a selective attentional focus on the segmented, non-radical component *alone*, will be positively unhelpful for semantic recoding.

It is tempting to infer somewhat *different*, lexical decomposition strategies from the radical-migration and non-radical-migration indices, respectively. Two features of the results argue against such an inference, however. These are, first, that the two lexical

decomposition indices are themselves significantly correlated (though only on RT); and second, that while the non-radical-migration index has the slightly stronger relationship with overall speed of performance, accuracy is better predicted by the radical-migration index. Without a much more fully elaborated model of CI lexical decision, this apparent inconsistency discourages more detailed speculation.

As one final note of caution, it should be addressed that any attempt to generalise from these results with skilled adult reading to children learning to read is full of risk. However, direct strategies of children learning to read in Chinese, on the model of Bryant and Bradley (1985), are urgently needed.

Conclusions

The present experiment on lexical decision with two-character CI has several important findings. Firstly from the point of view of process, lexical decision with two-character CI is influenced by the effects of CI-frequency and concreteness but not by CI-homophones. As in English, lexical decision is also unaffected by the CI-frequency of these CI-homophones. However, lexical decision on two-character CI shows clear evidence of lexical decomposition on the individual words (characters). The decomposition process is asymmetrical: it appears to involve segmenting the non-radical component away from the lexical radical, which itself acts as a base. Despite the fact that multiple-character CI occur as fixed lexical items in both spoken and written Chinese, lexical decisions on two-character CI are based on the recognition and lexical decomposition of individual words (characters). The basic functional unit for reading Chinese is thus the (single-character) 'word'.

Secondly, from point of view of reading strategies, in accord with the general view of reading Chinese, efficient lexical decision does seem to be mainly based on orthographic processing and its interface with the semantic domain, i.e., the lexical route, at least for skilled adult readers. A phonological strategy is shown to be less efficient. These two strategies themselves have an inverse relationship. As an optional strategy, lexical decomposition, involving segmenting the non-radical component (i.e., the rule-defined *phonetic* component), is clearly associated with a phonological recoding strategy. This

suggests that orthographic knowledge (such as knowledge of the form and position of the lexical radicals in Chinese orthography) may play an important role in word recognition, *only* in so far as phonological recoding is involved.

Chapter 9

Summary and Implications

This chapter presents a summary of the empirical results of eleven experiments conducted in the present study. An outline functional architecture for word recognition in Chinese is tentatively put forward, in the light of these empirical results. Some further implications of the present study are also discussed.

Summary of the empirical results

1) The visual analysis of Chinese script

The crucial finding in the visual analysis of Chinese words is that Chinese so-called logographic characters can be read analytically on the basis of orthographic units smaller than the words as wholes. The orthographic units defined in this study (i.e., the smallest stroke-patterns) were used by skilled readers of Chinese in visual comparisons of Chinese words, although the subjects were not taught explicitly about such units. Subjects were significantly affected both by the number of mismatching units (Unit-dissimilarity) and by the total number of orthographic units (visual complexity) in the visual comparison of Chinese words (Experiments 1 & 2). The unit-complexity effect was not confounded with stroke-complexity, since the effect of the former remained significant when the number of strokes in the stimuli was held constant. By contrast, the individual strokes had no significant influence upon visual comparisons of Chinese words. Hence, the functional orthographic unit of Chinese words is the smallest stroke-pattern, i.e., the orthographic unit proposed in this study, rather than the individual stroke as previously thought (*Wang, 1971; Martin, 1972*).

The evidence suggests some degree of functional parallelism between the orthographic units of Chinese words and letters in English words. Firstly, as in English, recognition of the orthographic units in Chinese must involve positional information. Pseudowords, with the lexical radical in the legitimate position, were significantly advantaged in comparison speed over nonwords, with the lexical radical in an

illegitimate position. The importance of the positional regularity of lexical radicals was confirmed in a lexical decision study of two-character CI (Experiment 11). Lexical radicals in an irregular position (i.e., in the position of a non-radical component) caused significant interference in a lexical decision task, as did the non-radical components in their normal position. However, the same set of lexical radicals in their regular positions did not have such interference (Experiment 11). Secondly, a word superiority effect was obtained, which interacted with the unit-effect, as inferred from the degree of dissimilarity and visual complexity effects. These latter factors had significant effects on words, but had a smaller influence on pseudowords and no significant effect on nonwords. On the other hand, words were also judged significantly faster than pseudowords and pseudowords faster than nonwords. These two word superiority effects, demonstrated for the first time in the English literature with single character words in Chinese (*Hung & Tzeng, 1981*), support the mutual interactions between orthographic units and words proposed by interactive models of word recognition (*e.g., McClelland & Rumelhart, 1981*).

2) Semantic and phonological processing of Chinese words

The phonological and semantic processing of Chinese words was investigated in the two series of experiments reported in Chapters 5, 6 and 7. The same stimuli in the critical conditions were assessed in both the phonological and the semantic comparison tasks, and with the same subjects. The analysis of semantic and phonological processing for Chinese words focused on the following characteristics:

a) Direct routes from script to sound and from script to meaning

The independence of semantic and phonological transcoding from print in reading constrains the basic functional architecture of most word recognition models in English. The present study suggests that this functional architecture of word recognition is qualitatively similar in Chinese. Accessing the phonological and semantic lexicons are independent processes in Chinese, as in English. The independence of phonological processing from semantic recoding was supported by the absence of a synonym effect in the phonological comparison of paired Chinese words. Paired words that have a

different pronunciation but similar meanings (synonyms) were not judged 'Different' significantly more slowly than the control words, which have different meanings. On the other hand, semantic processing is also independent of phonological recoding: the homophone effect was not significant in semantic comparison task. Here, it is worth emphasizing that the absence of homophone and synonym effects in semantic and phonological comparisons respectively merely implies that phonological and semantic information can be utilized independently without reference to each other. It should not be confused with the role of phonological recoding and semantic recoding in reading. The activation and utilisation of different (visual, semantic and phonological) codes in reading may be quite different questions. According to current parallel distributed models for word recognition, different types of code (e.g., visual, semantic and phonological) pertinent to written words are automatically activated in parallel. Phonological information, although not necessarily used in the semantic comparisons, may nonetheless be activated in parallel to the orthographic and semantic codes in Chinese (Tzeng, Hung, & Wang, 1977; Baddeley, 1979; Tzeng & Hung, 1980; Perfetti & McCutchen, 1982; Perfetti & Zhang, 1991), as in English. Note that although neither the homophone nor the synonym effects were significant on comparison speed, synonym and homophone pairs did have somewhat higher error rates (false Same) than the controls (Experiments 5 & 8). Some subjects may have used phonological information in the semantic comparison task, and vice versa even when this was not strictly necessary. Moreover, in lexical decisions of two-character CI, there were significant individual differences on the homophone effect (although the effect was not significant overall). Some subjects made no errors at all to CI-homophones (sounds meaningful but visually nonsense). Others made as many as 9 errors out of 15, treating CI-homophones (sounds like CI) as real CI in Chinese. Hence, considering phonological processing as an optional strategy, 'Phoenician' reading is by no means unique to English (Barron, 1978). The role of phonological recoding in reading may be considered as an optional strategy (Hung & Tzeng, 1981), influenced by many factors including individual differences, for skilled adult reading. This point will be further discussed later in this chapter.

b) *Rule-governed lexical decomposition process*

Before investigating the influence of rules in reading Chinese, the thesis clarified three types of rule in Chinese: 1) the etymological rules of word-formation, specifying semantic and phonetic components, well-known in the English literature; 2) the orthographic rules of lexical decomposition into semantic "*radical*" and phonetic "*component*", emphasising the structure of the two unique orthographic constituents—the lexical radical and the non-radical component; 3) phonological recoding rules of "*derivation*" and "*analogy*" (Chapter 2). The present study investigated the functional roles of the two major orthographic constituents, i.e., the lexical radical and the non-radical component, in both semantic and phonological processing, by manipulating graphic matching either on the radical or on the non-radical component. Paired experiments studying phonological and semantic comparisons (Experiments 6 & 9, 7 & 10) were conducted, using not only the same subjects but also the same stimuli for these two critical matching conditions (radical and non-radical matching). Radical and non-radical matching showed opposite cross-over effects in the phonological and semantic comparisons, even though the stimuli and subjects were identical. For the semantic comparisons, radical matching provoked longer reaction times and significantly higher error rates than either controls (stimulus-pairs with no matching constituents) or non-radical matching (stimulus pairs matching on the non-radical component). For the phonological comparisons, exactly the opposite results were found on both reaction times and error rates. This result is in accord with the postulated rules of lexical decomposition introduced in Chapter 2. According to these orthographic rules, the non-radical component is the regular "phonetic" component. This orthographic rule supposedly applies to *all* compound words, regardless of their etymological origin. Accordingly some words (ideocompounds) are orthographically *irregular*, since the pronunciation of these words is not congruent with their non-radical component. An important question, therefore, was whether a graphic match between the non-radical component in these *irregular* (ideocompound) words and the other member of the stimulus pair would have the same distractor effect on phonological comparison as it did in *regular* (ideophonetic) words. It did have this effect. This result indicates that skilled

readers of Chinese apply the orthographic rule in lexical decomposition of compound words regardless of their etymology. The non-radical component is treated initially as the phonetic element in *all* compound words. As already pointed out by Seidenberg (1987), it is the orthographic patterns that may be implicitly encoded in the reader's mind. Moreover, the lexical decomposition process, inferred from RT and error rates on the critical radical and non-radical matching conditions is more evident in the phonological processing task (as in English).

Furthermore, the present study has pursued the nature of the rules implemented in phonological recoding via lexical decomposition. By measuring various type-frequency indices of a component over all words containing this component in the lexicon, the efficiency of phonological processing was found to be significantly associated with the type-frequency of the orthographic-phonological correspondences of that component in the lexicon, rather than in specific words. These results favour the connectionist view that the rules are themselves embodied in the connections, and the implementation of the rules is through statistically modifying the weights of the connections. In the mental lexicon, the relative weights of connections may depend on the type-frequency of different part-to-whole orthographic-phonological correspondences in the lexicon. Pronunciation of a compound Chinese word can be achieved either via derivation, from the pronunciation of its own non-radical component from or via analogy from other words containing the same component. Both derivation and analogy may play significant roles in the phonological processing of Chinese words. However, derivation may be the major procedure, using lexical decomposition process, for the phonological recoding of Chinese words.

3) Word frequency effects in reading Chinese

Word frequency is one of the most ubiquitous factors in word recognition. In the present study, word frequency values were obtained from an updated word frequency dictionary in Chinese (*Beijing University, 1988*), and this variable was carefully balanced in all the main experiments. The interaction between lexical decomposition and the word frequency effect has been shown in Chinese as well as in English (*Seidenberg*

1985). This interaction was emphasized in a recent computational model for the pronunciation of English words (*Seidenberg & McClelland, 1989*). In a study on Chinese (*Seidenberg, 1985a*), the naming of Chinese words in Cantonese dialect was slower for low frequency words than for high frequency words. Moreover, 'regular' Chinese words were named faster than 'irregular' Chinese words, but only for the low frequency, not the high frequency words. The results of the present study (with a comparison paradigm) are not consistent with that study. The frequencies of the paired words used as two co-variates in the analysis of variance and co-variance (ACOVA), on by-item data, had no significant influence (individually or together) upon the phonological comparisons of the paired words. Moreover, the frequency effect of CI (two words together as a meaningful CI in Chinese), also had no significant effect on phonological processing, inferred from the homophone condition (Experiment 11). However, the CI-frequency effect was significant in lexical decisions to real CI, which are the base-CI of the CI-homophones. These results favour the recent view that word frequency may affect the orthographic processing but not the phonological processing of written words (*Besner & McCann, 1987; McCann, 1987; McCann & Besner, 1988*). The phonological lexicon in Chinese, as in English, may be insensitive to word frequency.

4) Reading and word recognition in Chinese

As a departure from reading single words (word recognition), the last experiment used two-character CI in a lexical decision task. Firstly, this experiment showed that lexical decisions to two-character CI in Chinese are significantly affected by CI-frequency (visual familiarity) and by concreteness of CI (semantics), but not by CI-homophone effects (phonology of CI). Hence the general view that reading Chinese words is mainly based on the so-called lexical routine, from orthography to semantics, (for silent reading) is supported, although visual, semantic and phonological codes may be automatically activated in parallel, in reading Chinese as in reading English. More importantly, the different orthographic components of single characters (words) in Chinese had significantly different influences upon lexical decisions to two-character CI in Experiment 11. This result suggests that lexical decision of two-character CI is based

on the processing of individual words (single characters), despite the fact that two-character CI are used as fixed expressions in the language (not unlike English idioms and compounds). This was further confirmed by the finding that the influence of the position of the same set of radicals was significant in terms of single characters (words) but not in terms of two-character CI. Hence, reading fixed multiple-word expressions (CI) in Chinese is based on single word units. These word units in Chinese are single characters. Single characters in Chinese should thus be regarded as "words" rather than "morphemes" (*Hung & Tzeng, 1981*). The combination of multiple characters is more appropriately equated to a "phrase" rather than a "word" in English, at least from a psychological point of view. Moreover, the non-radical-migration (i.e., phonetic-component-migration) but not the radical-migration was significant in this experiment. Lexical decomposition is more likely to be associated with phonological processing than with the semantic processing. This is confirmed by the significant association between the non-radical-migration index and the phonological recoding index in this experiment, as discussed below.

5) Reading strategies and reading efficiency in Chinese

The second aspect of the last experiment examined various reading strategies (or skills) and their association with lexical decisions to two-character CI. Visual, semantic, phonological and decomposition strategies were indexed respectively by the magnitude of CI frequency, concreteness, homophone and component-migrated effects in individual subjects. There was an inverse association between the semantic recoding index and the phonological recoding index. The phonological recoding index also had a significant positive association with the lexical decomposition indices, including indices of both radical-migration and non-radical-migration.

Experiment 11 also investigated a number of reading strategies and their association with reading efficiency, as inferred from the latency and accuracy of lexical decisions. Faster reading of two-character CI was shown to be significantly associated with a higher semantic recoding index, a lower phonological recoding index and a lower index of decomposition. Here it is important to emphasize that this relationship between

reading strategies and reading efficiency in Chinese pertains to skilled adult reading. It would be incautious to generalize any of these results to reading by children. The role of phonological recoding can change with increasing reading experience (*Barron & Baron, 1977; Barron, 1978*).

Toward a word recognition model for Chinese

In the light of the empirical results summarized above, some very tentative outlines of a functional architecture for word recognition in Chinese may be proposed. The author fully recognizes the provisional nature of these proposals, which could be only properly evaluated through detailed simulations.

First of all, the present study favours the current view that word recognition in logographic Chinese is unlikely to involve a radically different cognitive architecture from that employed in the recognition of alphabetic words, in English (*Hung & Tzeng, 1981; Flores d'Arcais, 1992*). It is therefore unnecessary to postulate an entirely novel model for reading Chinese. However, to accommodate orthography-specific features, or language-specific or culture-specific features in reading different orthographies, current word recognition models established on English must be 'tuneable' to other orthographies. The analytic approach developed in this thesis is certainly not yet a model, but it may constitute a theoretical basis for the establishment of a Chinese version of Parallel Distributed Processing (PDP) models for word recognition (*Rumelhart & McClelland, 1981; Seidenberg & McClelland, 1989*). Any validation of the proposed framework must await the future development of PDP models for Chinese.

To make this functional architecture more easily understood by non-Chinese readers as well as Chinese readers, orthographic units and words used in this framework are illustrated by geometric figures (as well as by examples in Chinese script). As seen in Figure 9, an **ellipse** represents the orthographic code; a **square** (as a mnemonic, a square is also a word meaning MOUTH in Chinese) and a **triangle** represent **phonological** and **semantic** codes, respectively. The framework illustrated in Figure 9.1 has the following characteristics:

Toward a word recognition model in Chinese

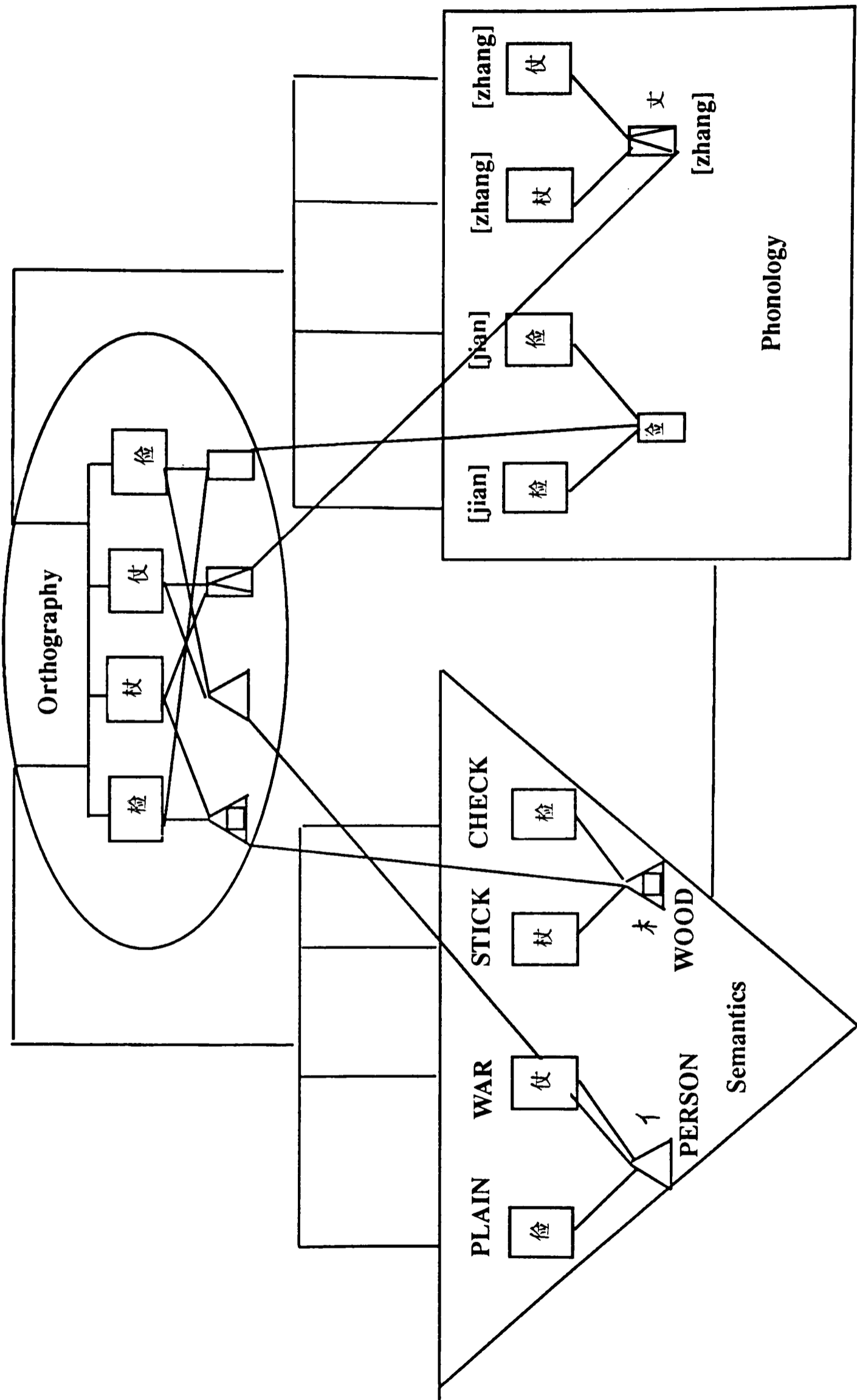


Figure 9.1 A functional architecture as a basis for a word recognition model in Chinese. See text for explanation.

1) Interface between the visual pattern and its linguistic domain

Recognition of written Chinese words is based on the recognition of the orthographic units, and their mappings to the appropriate linguistic domain. In Chinese, orthographic units can be divided into two principal categories: the lexical radical and the non-radical component. In many cases, the non-radical component can be further sub-divided into two or more constituent orthographic units (the 'smallest stroke-patterns'). However, for the purpose of phonetic or semantic recoding this further subdivision is not important. Each of the two principal categories can be classified as either word-components or graphic-components according to their lexical status (i.e., whether the component can be used as word in isolation, in modern Chinese). Figure 9.1 illustrates these four types of orthographic unit (木 WOOD/mu-word-radical, 攵 DEPEND/zhang-word-component, 扌 -graphic radical, 金 -graphic component), along with four words (e.g., 检 俭 仗 杖 CHECK/[jian]3, PLAIN/[jian]3, WAR/[zhang]4, STICK/[zhang]4) made up from these four categories. According to this framework, the important characteristic of the orthographic lexicon is that words are not considered just as graphic patterns, but as graphic patterns with associated linguistic features. As seen in Figure 9.1, the embracing figure represents the dominant linguistic pattern, as defined by the two rules of lexical decomposition (Chapter 2). The word-radical is the semantic component, according to the semantic rule, and it is represented by a square embedded in a **triangle**. The non-radical word-component is phonetic component, according to the phonetic rule, and it is represented by a triangle embedded in a **square**. A square and a triangle with no figures embedded represent the graphic component (e.g., 金) and the graphic radical (e.g., 扌), respectively, implying that these two components have no lexical status (i.e., cannot be used as words in isolation in modern Chinese), and hence have no *independent* phonetic or semantic interpretation, respectively. The importance of having these two principal orthographic bodies represented specifically in the orthographic lexicon of Chinese is that the orthographic processing of Chinese words may depend on distinguishing the lexical radicals from the non-radical components. The activation of distinct cognitive codes representing the lexical radical and non-radical component may be needed to access the semantic and the phonological lexicon,

respectively, in Chinese. In connectionist terms, larger weights should be given (a) to the connections between the codes for the lexical radical and coding unit in the semantic lexicon and (b) to those between the codes for the non-radical component and coding unit in the phonological lexicon.

2) *Tree-like structure for the organization of words and orthographic units*

The computational model developed by Seidenberg and McClelland (1989) dispensed with an explicit word-level but nevertheless maintained the power to simulate many experimental results. The necessity of the word unit in theoretical frameworks for word recognition is still debatable (McCann & Besner, 1990; Seidenberg & McClelland, 1990). This also needs to be examined in future studies of Chinese. The present framework preserves both the unit and word levels, as in earlier versions of PDP models (e.g., McClelland & Rumelhart, 1981). Given that all Chinese words (with more than one unit) have a lexical radical and a non-radical component, it may be reasonable to propose that in the mental lexicon of Chinese, the words and components are organized in a tree-like structure, with the component as trunk and the individual words as branches. Given the characteristics of Chinese orthography, in which the semantic and phonological information is encoded in two separate orthographic components, the lexical radical and the non-radical component, respectively, the trunk of the trees in the semantic and phonological lexicons are the lexical radical and non-radical component respectively. The semantic and phonologic lexicons consist of many such trees. A semantic tree may represent a semantic field or a category, while a phonetic tree may represent a syllable in Chinese. Multiple-word combinations (i.e., the CI in Chinese) may be represented by the connections between branches of the same or different trees. Figure 9.1 illustrates two such tree-like structures. (The tree in the mental lexicon obviously has more than two branches, as illustrated in Figure 9.1.)

In such tree-like structures, information processing in reading (orthographic, semantic and phonological) involves the activation of both the component (trunk) and the words (branches). There is mutual interaction between the activation of the component (the trunk) and its constituent words (the branches).

Activation of the component (the trunk) may feedforward a positive activation to all the words (branches) containing the same component (connected with the same trunk). The trunk (i.e., the lexical radical or non-radical component) also receives feedback from its branches. This feedback from branches can be positive or negative. The weight of connections in a phonetic tree may depend on the phonological congruency between the non-radical component and the constituent words. A regular word may have positive feedback while an irregular word may have negative feedback. The weight or strength of the trunk in particular, for a phonetic tree, is determined by the combined weights from all its branches, which will be influenced by the proportionate phonological congruency and non-congruency between the non-radical component (the trunk) and all of its constituent words in the lexicon (i.e., the branches). (See Experiment 6.) If the trunk is a graphic component, the connection strengths may be determined by the phonological consistency among its branches (i.e., the constituent words). Moreover, the connection weight of the trunk in the phonological lexicon (from the non-radical component) is also influenced by its own phonetic values. A graphic-component with no phonetic value in isolation should have less weight than a word-component with its own phonological value in isolation. Both the frequency indices of congruency and consistency individually appeared to play a significant role in the phonological processing of Chinese words (Experiment 6). However, the congruency index may play the major role in phonological recoding.

The activation of individual words depends on the activations from their own orthographic component (the trunk) and the activations from other lexicons at the word-level. The activation at the word level is one-to-one and therefore is word-specific. Word-specific activation may make the code for the target word become the strongest candidate for output. If such word-specific activation is disturbed in dyslexic conditions, activation of the words would depend mainly on the activation from the trunk (the component). In this case, different types of reading errors associated with the different components (the trunk) may appear, as discussed later in the implications of this study.

The tree-like structure proposed here may explain many important experimental effects, such as neighbourhood effects (*Coltheart, Davelaar, Jonasson & Besner, 1977; Laxon, Coltheart & Keating, 1988; Grainger, 1990; Grainger & Segui, 1990; Andrews, 1989, 1992*) and phonological congruency and consistency effects in word recognition. The orthographic neighbourhood in Chinese can be defined as the set of words with either the same lexical radical or the same non-radical components. In the present framework, the activation of a word can be influenced by other words connected to the same trunk (i.e., the non-radical component or lexical radical). The words connected to the same trunk would influence each other by contributing their own activation to modify the activation of the trunk. The phonological congruency (regularity) effect may be explained by interactions between the activation of the word and of its own non-radical component (trunk). Words that are phonologically congruent with their non-radical component (the trunk) may receive positive feedback activation of the trunk and therefore increase the total activation of the trunk. By contrast, words that are phonologically incongruent with their own non-radical component (trunk) would have negative weight for the trunk.

The implications of the present study

The present study has some implications for future studies on the following aspects of reading research in Chinese.

1) On experimental studies: some methodological concerns

a) Language statistics on word-stimuli

In order to read, we need to know the language. The present study has shown an important role for orthographic knowledge ("rules") of Chinese orthography in reading Chinese, demonstrated by the highly selective rule-like effects of the "semantic radical" and the "phonetic component". The results suggest that the phonetic rule is a generalization of the correspondence between a component and all the words containing that component in the lexicon. The implementation of the rule depends on

the relative frequency of this correspondence in the lexicon. In future studies of reading in Chinese, it is important to take account of the language statistics on the core lexicon. Some language statistics have been collected in Chinese (*e.g.*, Zhou, 1978; Yin, 1991). However, the future language statistics should be based on well-defined orthographic constituents in Chinese (in particular the lexical radical and the non-radical component) rather than on units defined by their etymological function (the "semantic radical" and "phonetic radical"). Apart from such well-known word attributes as word frequency, other measures such as the type-frequency of components, phonological regularity, and consistency should also be considered in future language-statistical studies. A wider range of language statistics, regarding various types of frequency of components in the lexicon, is particularly important for research on lexical decomposition (analytic processing) in reading Chinese.

b) Orthographic attributes of word-stimuli

Orthographic attributes of Chinese words, such as unit-complexity and the positional regularity of the lexical radical, had a significant influence upon the visual analysis of Chinese words, in Experiments 1 to 3. Compared to previous studies, the present research controlled these orthographic attributes of the words in the selection of word-stimuli in all experiments. Moreover, different orthographic patterns (such as left-right and top-bottom) have been carefully matched in the present study. Word frequency has also been carefully controlled by using by-item co-variance analyses in most experiments of this study. The present study emphasizes that the proper selection of stimuli in such experiments must be based on the clarification of orthographic structure. For instance, measures of the visual complexity of Chinese words should be based on counting the number of orthographic units (stroke-patterns) rather than individual strokes. Moreover, the distinction between different functional components of Chinese words should refer to the distinction between the major orthographic bodies (*i.e.*, the lexical radical and the non-radical component) rather than their etymological function (*i.e.*, semantic and phonetic).

c) *More stimuli*

The last methodological point concerns the sample sizes of the stimuli and subjects used in the experiments. The principle of the present study is 'more stimuli but fewer subjects'. The number of stimuli used for each experiment in the present study ranged from 360 to 720 pairs (double the number in terms of words). The subject numbers in these experiments were usually 12 to 16. Such numbers of subjects and particularly their cultural homogeneity may have minimized the experimental variance between subjects. The large sample of stimuli, covering a wider range of words in the language, provides more convincing evidence for the stimulus-condition effects within subjects. Another characteristic of the present study is that the same set of word-stimuli, as well as the same subjects, were used in the different tasks. This is a good method for the examination of condition effects and needs to be further explored in future experimental studies.

2) On future neuropsychological studies of dyslexia in Chinese

According to some reports, the incidence of developmental dyslexia in Chinese is about one tenth that in English. This could be attributed to several reasons apart from the difference between orthographies (*Flores d'Arcais, 1992*). However, the possibility that some potential reading difficulties may not be well recognized in Chinese cannot be excluded. Without stringent epidemiological studies on dyslexia in Chinese, based on proper criteria for diagnosis, the claim of a low incidence of dyslexia in Chinese cannot be very meaningful. The present study has implications for identifying potential reading errors especially component-associated reading errors .

a) *Component-associated phonological reading errors—"Half-word" reading and reversed reading errors in Chinese*

There is an old proverb in Chinese: the low level scholar only reads half of a word. A recent neuropsychological study in Chinese (*Butterworth & Yin, 1991*) has interpreted half-word reading as an example of so-called phonological regularization errors. The study refers to the error of reading a "phonetic radical" (half-word) of the target word when the target word as a whole should be read. These phonological regularization errors were used as an important criterion for distinguishing two types of dyslexia, i.e.,

surface and deep dyslexia in Chinese, analogous to the classification of dyslexia in English (*Marshall & Newcombe, 1966; 1973; Coltheart, Patterson & Marshall, 1980; Patterson, Coltheart & Marshall, 1985*). In Butterworth and Yin's study, patients who made regularization errors were classified as surface dyslexics. Patients who did not make such errors were regarded as deep dyslexics. Both types of dyslexics in Chinese can make semantic errors. However, in Butterworth and Yin's classification, the orthographic nature of the phonetic radical was seriously underspecified. From the limited information about reading errors reported in that study, the so-called "phonetic radical" of the target words selected in that study seems to be always on the right side of the words (*Butterworth & Yin, 1991*). This is not the real situation in Chinese language, however. Firstly, a left-right pattern for words is only one of three patterns allowable in Chinese. The phonetic component exists in all these three types of words in Chinese (e.g., the phonetic component of words 固 HARD/[gu]4, 花 FLOWER/[hua]1 and 邻 NEIGHBOUR/[lin]2 are 古 (centre), 化 (bottom) and 令 (left), respectively). Even for left-right words to have the phonetic component on the left side is not uncommon (e.g., 邻).

In the present study, it has been argued that the position of the phonetic component can only be determined relative to the position of the lexical radical, which, for different radicals, can be in any position (left, right, top, bottom, inside or outside) of words. For any specific radical this position is more or less fixed and is specified in the dictionary (see Chapter 2). This is a more appropriate criterion for the selection of compound word stimuli in future studies. Moreover, with regard to "half-word reading", the crucial point is to distinguish carefully the two types of half-word reading: i.e., non-radical component reading and lexical radical reading, as both components may be read aloud (i.e., are pronounceable). Neuropsychological data in Chinese collected by the author (but not included in this thesis) suggests that half-word reading in Chinese can be associated either with the lexical radical (semantic) or with the non-radical component (phonetic). These two types of reading errors are quite distinct. Non-radical component (the phonetic component) reading indicates that readers (patients) may have intact knowledge of Chinese orthography (orthography-phonology correspondence).

Reading aloud the lexical radical suggests, by contrast, the inability to distinguish the lexical radical from the non-radical component. The ability to recognize the lexical radical has been shown to be associated with semantic and phonological access in Chinese and to have direct influence upon reading efficiency in Chinese (Experiment 11).

Furthermore, reversed half-word reading errors have also been found by the author. That is, when a target word (usually a single-unit word) was presented, the patient read aloud another word, which is the target word plus another lexical radical: for instance, 母 as 每, 页 as 顶. This reversed half-word reading was found in a stroke patient with a left hemisphere lesion (including basal ganglia). The pattern may perhaps be explained by mutual interaction between word and unit, in the tree-like structure in the proposed functional architecture for word recognition in Chinese. Given that the single-unit word is usually also a component (trunk) in many, more complicated words, the activation of the trunk would activate all words in the tree. If the word-specific connection is impaired, patients may read the words containing the component instead of the component in isolation as a word. In data collected by the author, a type of writing error showing selective preservation of orthographic knowledge of components was also identified. A patient with a tumour affecting left parietal and occipital region lost his ability to write even his own name. However, he constantly constructed pseudowords by using real components of Chinese words. For instance, 休. More importantly, in the pseudowords he constructed the lexical radical was always in its legitimate position. Hence, this patient seems to have some intact orthographic knowledge (such as the positional regularity of the lexical radical and its structure). The trunk of the lexicon may be well preserved while the branches (i.e., individual words) are impaired..

b) Word-associated and CI-associated semantic errors in Chinese

The proposed tree-like structure suggests that two different types of semantic errors may exist in Chinese dyslexia. They are word-associated and CI-associated errors. Word-associated errors refer to the substitution of another word which has a similar meaning and/or a semantic association with the target word. CI-associated errors refer to the

substitution of a word which may not be semantically similar or even related to the target, but can constitute a CI in conjunction with the target word in Chinese. For instance, the target word 痛 PAIN [tong]⁴ may be read as 快 QUICK [qui]⁴, since PAIN QUICK together is a two-character CI, 痛快, meaning PLEASURE. Such CI-associated errors have already been reported in studies on Chinese dyslexia (*Lyman, Kwan & Chao, 1938 cited in Marshall, 1972*). The present framework provides an appropriate account for these CI-associated semantic errors.

c) Reading tests to study dyslexia in Chinese

The present study has demonstrated the important role of lexical radical in recognition and reading Chinese. It is therefore important to include a test of the ability to recognize the lexical radical in future reading tests on dyslexic patients. Such tests may employ the lexical decision task, i.e., whether or not a written symbol can be a Chinese word, which you may know or may not know? The nonwords should include a lexical radical in an illegitimate position. Skilled Chinese readers (such as the subjects used in the present study) described such nonwords (e.g., ㄨㄥ) as a different symbol system, such as "Japanese Kana". False positive responses on nonwords in the lexical decision task would imply that subjects (or patients) are unable or less able to recognize the lexical radical in Chinese. Nonwords used for lexical decision can consist of two non-radical components with no lexical radical. False positive response to such nonwords would also indicate a lack of orthographic knowledge of Chinese, since the lexical radical is an essential component of all words with more than one unit.

3) On the acquisition of Chinese words

a) A new analytic method for the acquisition of Chinese: the importance of recognition of the lexical radicals

Chinese orthography is acquired traditionally in the wholistic, "look and say" fashion. In order to remember Chinese words, it is thought necessary to trace the individual strokes of the characters in the air by the fingers. This method is also used when learning Chinese characters (Kanji) in Japanese. In the acquisition of Chinese words, word constituents (i.e., the spelling units), are not conventionally taught. The

traditional teaching of Chinese emphasised the need for much practice in the writing of the individual strokes. Conventionally, words are not spelled in the acquisition of reading or writing in school. Nevertheless, it is not uncommon to spell words in speech. For instance, the surname 李 [li]3 is commonly spelled as [shi]2 [ba]1 [zi]3 in speech when one is asked the surname. This spelling unit refers to the word-components 十 八 子 which individually means TEN/[shi]2; EIGHT/[ba]1; SONS/[zi]3. The apparent difficulty in teaching spelling by this method is that the orthographic units of Chinese words have not yet been consistently classified with respect to their graphic structures. Some of the patterns are no longer even pronounceable and are therefore difficult to spell.

The present study has made a first step toward the classification of orthographic units in Chinese words in terms of their structure, independent of etymological origin. The graphic patterns of the orthographic units, but not the individual strokes, can even be identified by non-Chinese readers through practice, independent of lexical information (Experiment 3). The acquisition of the graphic patterns of the orthographic units by non-Chinese was implicit since no instruction in the language was given. This is also the situation for most native adult readers of Chinese. Although the lexical radicals are not taught in learning to read in school, the position of the lexical radical may be learned implicitly by using the radical index dictionary (the most common dictionary used in the mainland). In order to look up an unknown word in a dictionary with a radical index, users must determine the lexical radical of the target word. As suggested by recent studies, explicit teaching of what was previously implicit knowledge may improve learning. In fact, in some experimental schools in mainland China, a new teaching method which involves the identification of stroke-patterns (especially the radicals of words) is used. Whether this new analytic method improves the efficiency of learning awaits evidence from future developmental studies in Chinese.

References

Allport, D.A. (1977). On knowing the meaning of words we are unable to report: the effect of visual masking. In S.Dornic (Ed) *Attention and Performance VI*. Hillsdale, N.J.: Lawrence Erlbaum Associates

Andrews, S. (1989). Frequency and neighbourhood size effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 802-814

Andrews, S. (1992). Frequency and neighbourhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 234-254

Baddeley, A.D. (1979). Working memory and reading. In P.A.Kolers, M.E. Wrolstad, & H. Bouma (Eds.) *Processing of visible language* (pp.355-370). New York: Plenum Press

Baddeley, A.D., & Hitch, C. (1974). Working Memory. In G.H. Bower (Ed.), *The Psychology of learning and motivation* (Vol. 8). New York: Academic Press

Bamber, D. (1969). Reaction time and error rates for "same"-"different" judgments of multidimensional stimuli. *Perception & Psychophysics*, 6, 169-174

Baron, J. (1973). Phonemic stage not necessary for reading. *Quarterly Journal of Experimental Psychology*, 25, 241-246

Baron, J. (1975). Successive stages in word recognition. In P.M.Rabbitt and S.Dornic (Ed) *Attention and Performance V*. London and New York: Academic Press

Barron, R.W. (1978). Accessing to the meaning of printed words: Some implications for reading and learning to read. In F.B.Murray (Ed.) *The recognition of words :IPA series on the development of the reading process* (pp. 34-56). Newark, DE: International Reading Association.

Barron, R.W.C. (1981). Reading skill and reading strategies. In A.M. Lesgold and C.A. Perfetti (Eds.) *Interactive processes in reading*. Hillsdale, NJ: Lawrence Erlbaum Association

Barron, R.N., & Baron, J. (1977). How Children get meaning from printed words. *Child Development*, 48, 487-594

Barron, R.W., & Henderson, L. (1977). The effects of lexical and semantic information on same-different visual comparison of words. *Memory and Cognition*, 5, 566-579

Barron, R.N., & Pittenger, J.B. (1974). The effect of orthographic structure and lexical meaning on "same-different" judgments. *Quarterly Journal of Experimental Psychology*, 26, 566-581.

Balota, D.A., and Chumbley, J.I. (1985). The locus of word-frequency effects in the pronunciation task: lexical access and/or production? *Journal of memory and Language*, 24, 89-106.

Besner, D. (1980) Visual word recognition: codes and procedures for accessing the internal lexicon. Unpublished doctoral dissertation, University of Reading, England.

Besner, D (1983) Basic decoding components in reading: Two dissociable feature extraction processors. *Canadian Journal of Psychology*, 37, 429-438

Besner, D. (1984) Specialized processors subserving visual word recognition: Evidence for local control. *Canadian Journal of Psychology*, 38, 126-138

Besner, D. , Coltheart, M., Davelaar, E. (1984) Basic processes in reading : Computation of abstract letter identities. *Canadian Journal of Psychology*, 38, 126-134.

Besner, D., & Hildebrandt, N. (1987). Orthographic and phonological codes in the oral reading of Japanese Kana. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13, 335-343

Besner, D. (1987). On the relationship between orthographies and phonologies in visual word recognition. In A. Allport, D. MacKay, W. Prinz, E. Scheer (Eds.) *Language perception and production*. London: Academic Press

Besner, D., & McCann, R.S. (1987). Word frequency and pattern distortion in word identification and production: An examination of four classes of models. In

- M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp.201-219). Hillsdale, NJ: Lawrence Erlbaum Associate
- Besner, D., Twilley, L., McCann, R.S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review*, 97, 432-446
- Brown, G.D.A. (1987). Resolving inconsistency: A computational model of naming. *Journal of memory and language*, 26, 1-23
- Bruder, G.A. (1978). Role of visual familiarity in the word-superiority effects obtained with the simultaneous matching task. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 88-100
- Butterworth, B., & Yin, W.G. (1991). The universality of two routines for reading: evidence from Chinese dyslexia. *Proceedings of the Royal Society*, 245, 91-95
- Cattle, J.M. (1886). The time taken up by cerebral operations. *Mind*, 11, 220-242.
- Carr, T.H., & Pollatsek, A. (1985). Recognising printed words: A look at current models. In D. Besner, T.G.Waller and G.E.MacKinnon (Eds.) *Reading Research Vol. 5*. Orlando: Academic Press
- Carr, T.H., Posner, M.I., Pollatsek, A., & Snyder, C.R.R. (1979). Orthography and familiarity effects in word processing. *Journal of Experimental Psychology: General*, 108, 389-414
- Chambers, S.M., & Forster, K.I. (1975). Evidence for lexical access in a simultaneous matching task. *Memory and Cognition*, 3, 549-559
- Chao, Y.R. (1968). *Language and Symbolic Systems*. London: Cambridge University Press
- Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row
- Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible language*, 15, 245-286

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

Coltheart, M., Patterson, K., & Marshall, J.C. (1980). *Deep Dyslexia*. London: Routledge and Kegan Paul

Cornad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, 55, 75-84

Corcoran, D.W., & Weening, D.L. (1967). An acoustic factor in proof reading. *Nature*, 214, 851

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J.T. (1978). Phonological recoding and lexical access. *Memory and Cognition*, 6, 319-402

Egeth, H., & Blecker, D. (1971). Differential effects of familiarity on judgments of sameness and difference. *Perception & Psychophysics*, 9, 321-326

Eichelman, W.H. (1970). Familiarity effects in a simultaneous matching task. *Journal of Experimental Psychology*, 86, 275-282

Fang, S.P., & Wu, P. (1989). Illusory conjunctions in the perception of Chinese characters. *Journal of Experimental Psychology*, 15, 434-447

Flores d'Arcais, G.B. (1992). Graphemic, Phonological and semantic activation processes during the recognition of Chinese characters. In H.C. Chen and O.J.L. Tzeng (Ed.) *Language processing in Chinese (personal communication)*.

Forster, K.I., & Chambers, S.M. (1973). Lexical access and naming time. *Journal of verbal learning and verbal behaviour*, 12, 627-635

Gibson, E.J., Shurcliff, A., & Jonas, A. (1970). Utilization of spelling patterns by deaf and hearing subjects. In H. Levin and J.P. Williams (Eds.) *Basic Studies on Reading*. New York: Basic Books

Glushko, R.J. (1979). The organization of and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 674-691

Grainger, J. (1990). Word frequency and neighbourhood frequency effects in lexical decision and naming. *Journal of memory and language*, 29, 228-244

Grainger, J., & Segui, J. (1990). Neighbourhood frequency effects in visual word recognition: A comparison of lexical decision and mask identification latencies. *Perception & Psychophysics*, 47, 191-198

Henderson, L. (1974). A word superiority effect without orthographic assistance. *Quarterly Journal of Experimental Psychology*, 26, 310-311

Henderson, L. (1982). *Orthography and Word Recognition in Reading*. London: Academic Press

Henderson, L. (1985). Issues in the modelling of pronunciation assembly in normal reading. In K.E. Patterson, J.C. Marshall and M. Coltheart (Eds.), *Surface Dyslexia*. London: Lawrence Erlbaum Associates

Henderson, L. (1987). Word recognition: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: the psychology of reading* (pp. 171-200). London: Lawrence Erlbaum Associates

Hoosain, R. (1991). *Psycholinguistic Implications for Linguistic Relativity. A case study of Chinese*. Hillsdale, N.J.: Lawrence Erlbaum Associates

Hull, C.L. (1920). Quantitative aspects of the evolution of concepts. *Psychological monographs*, 28, No. 123

Humphreys, G.W., Besner, D., & Quinlan, P.T. (1988). Event perception and word repetition effect. *Journal of Experimental Psychology: General*, 117, 51-67

Humphreys, G.W., & Evett, L.J. (1985). Are there independent lexical and non-lexical routes in word processing? A evolution of the dual-route theory in reading. *The behaviour and Brain Sciences*, 8, 689-740

Humphreys, G.W., & Evett, L.J., Quinlan, P.T. & Besner, D. (1987) Orthographic priming: Qualitative difference between priming from identified and unidentified primes. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp. 201-219). Hillsdale, NJ: Lawrence Erlbaum Associate

Hung & Hu (1990) The comprehensive knowledge of Chinese words 汉字学通论 ,
Wu Han: Hua Zhong Normal University Press

Hung, D.L., & Tzeng, O.J.L. (1981). Orthographic variations and visual information processing. *Psychological Bulletin*, 90,377-414

Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 110, 306-340

Johnston, J.C., & McClelland, J.L. (1973). Visual factors in word perception. *Perception & Psychophysics*, 14, 365-370

Kargren (1971). *Sound and Symbols in Chinese*. Oxford:Oxford University Press

Katz, L., & Feldman, L.B. (1983). Relationship between pronunciation and recognition of printed words in deep and shallow orthographies. *Memory and Cognition*, 10, 576-590

Kay, J. (1986). Phonological codes in reading: Assignment of sub-word phonology. In D.A.Allport, D. MacKay, W. Printz and E.Scheer (Eds.) *Language perception and production*. London: Academic Press

Kay, J., & Bishop, D. (1987). Anatomical differences between nose, palm, and foot, or the body in question: Further discussion of the processes of sub-lexical spelling-sound translation. In M.Coltheart (Ed.) *Attention and Performance XII: the psychology of reading* (pp. 449-470). London: Lawrence Erlbaum Associates

Kleiman, G.M. (1975). Speech recoding in reading. *Journal of verbal learning and verbal behaviour*, 14, 323-339

Krueger, L. (1975). Familiarity effects in visual information processing. *Psychology Bulletin*, 82, 949-974

Krueger, L.E. (1978). A theory of perceptual matching. *Psychological Review*, 85, 278-304

Laxon, V.J., Coltheart, V., & Keating, C. (1988). Children find friendly word friendly too: words with many neighbours are easier to read and spell. *British journal of educational psychology*, 58, 103-119

Lyman, R.S., Kwan, S.T., & Chao, W.H. (1938). Left occipital-parietal brain tumor with observation on alexia and agraphia in Chinese and in English. *The Chinese Medical Journal*, 54, 491-516

MacKay, D.G. (1972). Input testing in the detection of misspelling. *American Journal of Psychology*, 85, 121-127

Marshall, J.C., & Newcombe, F. (1966). Syntactic and semantic errors in paralexia. *Neuropsychologia*, 4, 169-176

Marshall, J.C., & Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. *Journal of psycholinguistic Research*, 2, 175-199

Martin, S.E. (1972). Non-alphabetic writing system. In J.F.Kavanagh and I.G. Mattingly (Eds.) *Language by Eye and by Ear* (pp. 81-102). Cambridge: M.I.T. Press

Massaro, D.M. (1975). *Understanding language: An information processing model for speech perception, reading and psycholinguistics*. New York: Academic Press

Massaro, D.M., Venezky, R.L., & Taylor, G.A. (1979). Orthographic regularity, positional frequency and visual processing of letter strings. *Journal of Experimental Psychology*, 108, 107-124

Mayer, D.E., Schvaneveldt, R.W., & Ruddy, M.G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory and Cognition*, 2, 309-321

McCann, R.S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 14-24

McCann, D., Besner, D., & Davelaar, E. (1988). Word recognition and Identification: Do word frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance*, 14, 693-706

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review*, 88, 375-407

Miller, J.M., & Bauer, D.W. (1981). Irrelevant difference in the "Same"-"Different" task. *Journal of Experimental Psychology*, 7, 196-201

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165-178

Morton, J. (1982) Disintegrating the lexicon: An information processing approach. In J. Mahler, S. Franck, E. C. I. Walker, & M. Garrett (Eds.) *Perspectives on mental representation* (pp. 89-109). Hillsdale, NJ: Erlbaum

Newcombe, F., & Marshall, J.C. (1981). On psycholinguistic classifications of acquired dyslexia. *Bulletin of the Orton Society*, 31, 29-46

Nickson, R. (1967). "same"-"different" response times with multiattribute stimulus differences. *Perceptual & Motor Skills*, 24, 543-554

Paradis, M., Hagiwara, H., & Hildebrandt, N. (1985). *Neurolinguistic aspects of the Japanese writing system*. London: Academic Press

Parkin, A.J. (1982). Phonological recoding in lexical decision: Effects of spelling-sound regularity depend on how regularity is defined. *Memory and Cognition*, 10, 43-53

Patterson, K.E., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.) *Attention and Performance XII: the psychology of reading* (pp. 421-447). London: Lawrence Erlbaum Associates

Patterson, K.E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K.E. Patterson, J.C. Marshall, M. Coltheart (Eds.) *Surface Dyslexia*. London: Lawrence Erlbaum Associates

Patterson, K.E., Marshall, J.C., & Coltheart, M. (1985). *Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading*. London: Lawrence Erlbaum Associates

Perfetti, C.A., & McCutchen, D. (1982). Speech and language in reading. In N.Lass (Ed.), *Speech and language: Advances in basic research and practice*. San Diego, CA: Academic Press

Perfetti, C.A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 633-643

Pynte, J. (1974). Readiness for pronunciation during the reading process. *Perception & Psychophysics*, 16, 110-112

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In Coltheart, M (Ed.), *Attention and Performance XII: the psychology of reading* (pp. 327-362). London: Lawrence Erlbaum Associates

Reicher, G.M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 274-280

Rubenstein, H., Garfield, L., & Millikan, J.A. (1970). Homographic entries in the internal lexicon. *Journal of verbal learning and verbal behaviour*, 9, 487-494

Rubenstein, H., Lewis, S.S., & Rubenstein, M.A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of verbal learning and verbal behaviour*, 10, 645-657

Savin, H. (1963). Word frequency effect and errors in the perception of speech. *Journal of the acoustical society of America*, 35, 200-206

Schwartz, M.F., Marin, O.S.M., and Saffran, E.M. (1979). Dissociations of language in dementia: A case study. *Brain and language*, 7, 277-306

Seidenberg, M.S. (1985a). The time-course of phonological code activation in two writing systems. *Cognition*, 19, 1-30

Seidenberg, M.S. (1985b). The time course information activation and utilisation in visual word recognition. In D. Besner, T.G.Waller and G.C.MacKinnon (Eds.), *Reading Research*, vol. 5.Orlando: Academic Press

Seidenberg, M.S. (1987). Subword structure or orthographic redundancy?. In M. Coltheart (Ed.) *Attention and Performance XII: the psychology of reading* (pp. 421-447).London: Lawrence Erlbaum Associates

Seidenberg, M.S., & McClelland, J.L. (1989). Distributed, Developmental Model of word recognition and naming. *Psychological Review*, 4, 523-568

Seidenberg, M.S., & McClelland, J.L. (1990). More words but still no lexicon: Reply to Besner et al (1990). *Psychological Review*, 97, 447-452

Seidenberg, M.S., & Vidanovic, S. (1985c). Word recognition in Serbo-Croatian and English: Do they differ? Paper presented to the Psychonomic Society meeting, Boston

Seidenberg, M.S., Waters, G.S., Barnes, M.A., & Tanehaus, M.K.(1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behaviour*, 23, 383-404

Slowiaczek, L.M., & Pisoni, D.B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory and Cognition*, 14, 230-237

Taft, M. (1979). Lexical access via an orthographic code: The Basic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behaviour*, 18, 21-39

Taft, M. (1985). The decoding of words in lexical access: A review of the morphological approach. In D. Besner, T.G. Waller, & G.E. MacKinnon (Eds), *Reading Research: advances in theory and practice*, vol. V. Orlando, Fla: Academic Press

Taft, M. (1986). Lexical access in visual and auditory word recognition. *Language and Cognitive processes*, 1, 49-60

Taft, M. (1987). Morphological processing: The BOSS re-emerges.In M. Coltheart (Ed.) *Attention and Performance XII: the psychology of reading*. London: Lawrence Erlbaum Associates

- Taft, M., & Forster, K.I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behaviour*, 15, 607-620
- Taft, M., & Hambly, G. (1986). Exploring the cohort model of spoken word recognition. *Cognition*, 22, 259-282
- Taylor, G.A., Miller, T.J., & Juola, J.F. (1977). Isolation of visual units in the perception of words and nonwords. *Perception & Psychophysics*, 21, 377-386
- Treiman, R., & Chafetz, J. (1987). Are there onset and rime-like units in printed words? In M.Coltheart (Ed) *Attention and Performance XII: the psychology of reading* (pp. 281-298). London: Lawrence Erlbaum Associates
- Tzeng, O.J.L., & Hung, D.L. (1980). Reading in a non-alphabetic writing system. In J.F. Kavanagh and R.L. Veneky (Ed). *Orthography, Reading and Dyslexia*. Baltimore: University Park Press
- Tzeng, O.J.L., & Hung, D.L. (1981). Linguistic determinism: A written language perspective. In O.J.L. Tzeng and H. Singers(Eds.), *Perception of Print: reading research in experimental psychology*. Hillsdale, N.J.: Lawrence Erlbaum Associates
- Tzeng, O.J.L., Hung, D.L., & Wang, S.Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 621-630
- XIN HUA Dictionary (1979), Shanghai: Shan Wu Press
- Wang, W.S.-Y. (1971). Review of Liu: Chinese characters and their impact on other languages of east Asia. *Modern language Journal*, 15, 187-188
- Wang, W.S.-Y. (1973). The Chinese language. *Scientific American*, 228, 50-60
- Wheeler, D.D. (1970). Processes in word recognition. *Cognitive Psychology*, 1, 59-85
- Wieger, L.S.J. (1965). *Chinese Characters: Their origin, etymology, history, classification and signification. A thorough study from Chinese documents*. New York: Paragon Book Reprint Corp. and Dover Publications, Inc.

Yin, W.G. Unpublished Doctoral thesis (1991) University of London

Zhang, G., & Simon, H. (1985). STM capacity for Chinese words and idiom: Chunking and the acoustical loop hypothesis. *Memory and Cognition*, 13,193-201

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

Appendix I

**Stimulus-pairs in Word-Radical and Word-Component Match
conditions in Experiments 6 & 9**

- | | | | | | | | |
|-----|--------|------------------|----------------------|-----|--------|-----------------|---------------------|
| 1. | 舌
辞 | [She]
[Ci] | Tongue
Language | 16. | 瓶
瓦 | [Ping]
[Wa] | Bottle
Tile |
| 2. | 齿
龄 | [Chi]
[Ling] | Teeth
Age | 17. | 亩
田 | [Mu]
[Tian] | Acre
Field |
| 3. | 欠
欣 | [Qian]
[Xin] | Owe
Happy | 18. | 般
舟 | [Ban]
[Zhou] | Sort
Boat |
| 4. | 鸟
鸣 | [Niao]
[Ming] | Bird
Bird-singing | 19. | 页
顶 | [Ye]
[Ding] | Page
Top |
| 5. | 山
岩 | [Shan]
[Yan] | Mountain
Cliff | 20. | 次
欠 | [Ci]
[Qian] | Order
Owe |
| 6. | 身
射 | [Shen]
[She] | Body
Shoot | 21. | 需
雨 | [Xu]
[Yu] | Need
Rain |
| 7. | 田
男 | [Tian]
[Nan] | Field
Man | 22. | 解
角 | [Jie]
[Jiao] | Divide
Coin |
| 8. | 身
躬 | [Shen]
[Gong] | Body
Bow | 23. | 资
贝 | [Zi]
[Bei] | Money
Shellfish |
| 9. | 雨
零 | [Yu]
[Ling] | Rain
Zero | 24. | 益
皿 | [Yi]
[Min] | Benefit
Utensils |
| 10. | 目
睡 | [Mu]
[Shui] | Eye
Sleep | 25. | 睁
目 | [Zheng]
[Mu] | Open eyes
Eye |
| 11. | 米
类 | [Mi]
[Lei] | Rice
Type | 26. | 种
禾 | [Zhong]
[He] | Plant
Grain |
| 12. | 白
皇 | [Bai]
[Huang] | White
Royal | 27. | 好
女 | [Hao]
[Nu] | Good
Woman |
| 13. | 日
明 | [Ri]
[Ming] | Sun
Bright | 28. | 墨
黑 | [Mo]
[Hei] | Ink
Black |
| 14. | 羽
翅 | [Yu]
[Chi] | Feather
Wing | 29. | 喝
口 | [He]
[Kou] | Drink
Mouth |
| 15. | 音
韶 | [Yin]
[Shao] | Sound
Splendid | 30. | 负
贝 | [Fu]
[Bei] | Bear
Shellfish |

【 Word-component 】

- | | | | | | | | |
|-----|--------|-------------------|---------------------|-----|--------|-------------------|--------------------------|
| 1. | 甲
押 | [Jia]
[Ya] | First
Mortgage | 16. | 京
掠 | [Jing]
[Lue] | Capital
Plunder |
| 2. | 非
排 | [Fei]
[Pai] | Wrong
Row | 17. | 挑
兆 | [Tiao]
[Zhao] | Select
Sign |
| 3. | 由
抽 | [You]
[Chou] | Reason
Take out | 18. | 执
丸 | [Zhi]
[Huan] | Hold
Pellet, |
| 4. | 台
怡 | [Tai]
[Yi] | Stage
Joyful | 19. | 透
秀 | [Tou]
[Xiou] | Penetrate
Elegant |
| 5. | 京
凉 | [Jing]
[Liang] | Capital
Cold | 20. | 怪
圣 | [Guai]
[Sheng] | Strange
Holy |
| 6. | 里
埋 | [Li]
[Mai] | Inside
Bury | 21. | 错
昔 | [Cuo]
[Xi] | Mistake
Old |
| 7. | 牛
牢 | [Niu]
[Lao] | Cow
Prison | 22. | 除
余 | [Chu]
[Yu] | Divide
Spare |
| 8. | 生
性 | [Shen]
[Xing] | Life
Nature | 23. | 统
充 | [Tong]
[Chong] | Altogether
Sufficient |
| 9. | 平
秤 | [Ping]
[Cheng] | Balance
Scale | 24. | 愁
秋 | [Zhou]
[Qiu] | Worry
Autumn |
| 10. | 周
调 | [Zhou]
[Diao] | Circuit
Transfer | 25. | 件
牛 | [Jian]
[Niu] | A matter
Cow |
| 11. | 干
奸 | [Gan]
[Jian] | Do
Wicked | 26. | 破
皮 | [Po]
[Pi] | Torn
Skin |
| 12. | 完
院 | [Wang]
[Yuan] | Complete
Yard | 27. | 台
治 | [Tai]
[Zhi] | Stage
Treatment |
| 13. | 九
仇 | [Jiu]
[Chou] | Nine
Hate | 28. | 语
吾 | [Yu]
[Wu] | Language
I |
| 14. | 军
挥 | [Jun]
[Hui] | Army
Wave | 29. | 切
彻 | [Qie]
[Che] | Cut
Through |
| 15. | 焰
舀 | [Yan]
[Yao] | Flame
Scoop up | 30. | 坦
旦 | [Tang]
[Dan] | Smooth
Dawn |

Appendix II

Two-character CI in three categories used in Experiment 11

【High Frequency CI】

High Concrete CI:

队伍 [dui] [wu] Troops
 玻璃 [bo] [li] Glass
 火车 [huo] [che] Train
 钢铁 [gang] [tie] Steel

电话 [dian] [hua] Telephone
 城市 [cheng] [shi] City
 鞋子 [xie] [zhi] Shoe
 眼睛 [yang] [jing] Eye

电影 [dian] [ying] Movie
 出去 [chu] [qu] Out
 纤维 [qian] [wei] Fibre
 颜色 [yang] [se] Colour

Low Concrete CI:

调查 [diao] [cha] Investigation
 懂得 [dong] [de] Understand
 坚持 [jian] [chi] Insist
 真理 [zhen] [li] Truth

保护 [bao] [hu] Protection
 机会 [ji] [hui] Opportunity
 相信 [xiang] [xin] Believe
 高兴 [gao] [xin] Happy

剥削 [bo] [que] Exploit
 计划 [ji] [hua] Plan
 希望 [xi] [wang] Hope
 方针 [fang] [zhen] Principle

Non-specific CI:

敌人 [di] [ren] Enemy
 按照 [an] [zhao] According to
 包括 [bao] [kuo] Include
 必须 [bi] [xu] Must
 不管 [bu] [guan] Despite
 产品 [chan] [pin] Product
 成功 [cheng] [gong] Success
 创造 [chuang] [zao] Create
 错误 [cuo] [wu] Mistake
 代替 [dai] [ti] Replace
 故事 [gu] [shi] Story
 过程 [guo] [cheng] Procedure
 活动 [huo] [dong] Activity
 共同 [gong] [tong] Common
 形式 [xing] [shi] Form
 会议 [hui] [yi] Meeting
 真正 [zhen] [zheng] Real
 生产 [sheng] [chan] Manufacture
 指挥 [zhi] [hui] Command
 方向 [fang] [xiang] Direction
 态度 [tai] [du] Attitude
 时代 [shi] [dai] Era

斗争 [dou] [zheng] Struggle
 办法 [ban] [fa] Mean
 本来 [ben] [lai] Origin
 变成 [bian] [cheng] Become
 部分 [bu] [fen] Partial
 长期 [chang] [qi] Long-term
 重新 [chong] [xin] Again
 从来 [cong] [lai] Always
 达到 [da] [dao] Reach
 党员 [dang] [yuan] Party member
 负责 [fu] [ze] Responsible
 过来 [guo] [lai] Come
 集体 [ji] [ti] Collective
 价值 [jia] [zhi] Value
 失败 [shi] [bai] Failure
 全部 [quan] [bu] Entire
 互相 [hu] [xiang] Each other
 教师 [jiao] [shi] Teacher
 制度 [zhi] [du] System
 方法 [fang] [fa] Method
 生存 [sheng] [cun] Survive
 使用 [shi] [yong] Use

多少 [duo] [shao] Many
 帮助 [bang] [zhu] Help
 比较 [be] [jiao] Compare
 表现 [biao] [xian] Manifest
 材料 [cai] [liao] Material
 彻底 [che] [di] Complete
 处理 [chu] [li] Process
 存在 [cun] [zai] Exist
 大家 [da] [jia] Everybody
 代表 [dai] [biao] Delegate
 规定 [gui] [ding] Regulation
 后面 [hou] [mian] Back
 改变 [gai] [bian] Change
 消息 [xiao] [xi] News
 环境 [huan] [jing] Environment
 引起 [yin] [qi] Cause
 东西 [dong] [xi] Goods
 睡觉 [shui] [jiao] Sleep
 制造 [zhi] [zao] Make
 方面 [fang] [mian] Aspect
 师傅 [shi] [fu] Master
 实在 [shi] [zai] Concrete

【Low Frequency CI】

High Concrete CI:

通道 [tong] [dao] Channel	庭院 [ting] [yuan] Yard	图形 [tu] [xing] Picture
旱烟 [han] [yan] Pipe smoke	矮小 [ai] [xiao] Short & small	按钮 [an] [niu] Button
脚趾 [jiao] [zhi] Toe	灯泡 [deng] [pao] bulb	匕首 [bi] [shou] Knife
讲台 [jiang] [tai] Stage	碧空 [bi] [kong] Blue sky	库房 [ku] [fang] Warehouse
礼服 [li] [fu] Suit	条纹 [tiao] [wen] Stripe	

Low concrete CI:

悼念 [dao] [nian] Mourn	哀悼 [ai] [dao] Grieve	傲慢 [ao] [man] Arrogant
懊悔 [ao] [hui] Regret	巴结 [ba] [jie] Flatter	虚伪 [xu] [wei] Hypocritical
屈服 [qu] [fu] Surrender	虚荣 [xu] [rong] Vanity	秀丽 [xiu] [li] Beautiful
道义 [dao] [yi] Moral	猖狂 [chang] [kuang] Savage	宽厚 [kuan] [hou] Generous
酷爱 [ku] [ai] Ardently love	理会 [li] [hui] Comprehend	

Non-specific CI:

坦率 [tan] [shuai] Frank	弹射 [tan] [she] Eject	调和 [tiao] [he] Mediate
听讲 [ting] [jiang] Listen	袒护 [tan] [hu] Shield	微观 [wei] [guan] Microscopic
顺从 [shun] [cong] Obedient	盗用 [dao] [yong] Usurp	倒闭 [dao] [bi] Bankruptcy
爱戴 [ai] [dai] Love & esteem	碍事 [ai] [shi] Hindrance	汗腺 [han] [xian] Sweat gland
昂贵 [ang] [gui] Expensive	拔除 [ba] [chu] Remove	跋涉 [ba] [she] Trudge
拜访 [bai] [fang] Visit	败坏 [bai] [huai] Discredit	失灵 [shi] [ling] Out of order
顺应 [shun] [ying] Conform	胜负 [sheng] [fu] Win or lose	后勤 [hou] [qin] Rear service
失眠 [shi] [mian] Insomnia	延误 [yan] [wu] Delay	沿途 [yan] [tu] On the way
绵延 [mian] [yan] Continuous	延伸 [yan] [shen] Stretch	清算 [qing] [suan] Clear
轰炸 [hong] [zha] Bomb	全程 [quan] [cheng] Whole journey	天堂 [tian] [tang] Paradise
填补 [tian] [bu] Mend	公认 [gong] [ren] Accepted	搬运 [ban] [yun] Transport
导航 [dao] [hang] Navigation	扮相 [ban] [xiang] Pseudo-face	半径 [ban] [jing] Radius
伴奏 [ban] [zou] Accompany	绑架 [bang] [jia] Kidnap	低垂 [di] [chui] Hang down
备战 [bei] [zhan] Prepare for war	崩溃 [beng] [kui] Collapse	逼真 [bi] [zhen] Lifelike
辩解 [bian] [jie] Debate	产地 [chan] [di] Producing area	长远 [chang] [yuan] Long-term
带路 [dai] [lu] Show the way	雇佣 [gu] [yong] Employ	空话 [kong] [hua] Idle talk
散落 [san] [luo] Spread	酷刑 [ku] [xing] Torture	理发 [li] [fa] Hair-cut
谅解 [liang] [jie] Forgive	宁静 [ning] [jing] Quiet	宁愿 [ning] [yuan] Would rather
偏巧 [pian] [qiao] Coincidence	骗局 [pian] [ju] Fraud	相容 [xiang] [rong] Compatible
相邻 [xiang] [lin] Neighbour	消逝 [xiao] [shi] Disappear	销毁 [xiao] [hui] Destroy