CRACKING THE CHINESE ORTHOGRAPHY: TOWARDS A FRAMEWORK FOR ASSESSING INTERVENTIONS IN LEXICAL ACQUISITION

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
EAST ASIAN LANGUAGES AND LITERATURES (CHINESE)

MAY 2012

By

W. Daniel Child

Dissertation Committee:
David Ashworth, Chairperson
Hsin-I Hsieh
Ying-che Li
Alexander Vovin
Martha Crosby

Keywords: Chinese, lexical acquisition, orthography, psychometric model

Copyright @2012 W. Daniel Child
All rights reserved.

In memory of

F. William McElroy

Acknowledgments

I would like to thank the members of the dissertation committee—Hsin-I Hsieh, Yingche Li, Alexander Vovin, Martha Crosby, and David Ashworth—for their patience and support as this dissertation went through its many iterations.

From the beginning, discussions and (friendly) arguments with Dr. Ashworth have been especially stimulating as I worked out the best approach to this research. His continued support throughout this study was a great source of encouragement.

At various stages, each of the committee members has pointed me toward helpful materials that have guided my thinking. Dr. Vovin and Dr. Li, in particular, provided detailed notes and caught numerous errors in earlier drafts, for which I am extremely grateful.

A special thanks must go to Dr. Crosby for her advice on issues to explore (and stay clear of!), as well as her assistance in applying for grants. Although the grants proved not to be forthcoming, the application process itself did much to help the disparate ideas presented here coalesce into something that I hope approaches coherence. Her expertise in statistics was invaluable, and her willingness to assist with last-minute application procedures was certainly beyond the call of duty.

I would also like to express my gratitude to all those who were kind enough to participate in my research experiments. Despite the fact that participation was not compensated, participation mortality was extremely low, and my subjects' patience in undertaking what for them were no doubt difficult and bizarre activities is much appreciated.

Huge thanks go to my sister, Kathleen McElroy, who urged me to enter graduate school and who, along with her husband Bill, helped immensely along the way. I appreciate her undying faith in what I have been trying to do.

Last but not least, I thank my loving wife, Meeishi Child, for her patience during the early years of our marriage, where I spent one foot "in the real world" and another in graduate school. The arrival of our daughter, Anjali, just as I started work on this dissertation no doubt slowed progress during this awkward phase, but the years have been all the happier on her account.

Abstract

Attaining a critical mass of vocabulary and character knowledge poses the greatest challenge to foreign leaners of Chinese wishing to achieve independent reading ability. Curricula that work well for simpler script systems fail to address additional complexities in the Chinese orthography, resulting in a significantly steeper learning curve.

To better understand how to moderate this learning curve, this study analyzes the Chinese writing system from multiple perspectives. In the process, the study articulates lexical acquisition goals, identifies mandatory and potential subcomponents of Chinese lexical acquisition, assesses past and current reading theory, explores relevant facets of information processing theory, and reexamines critical components of memory theory. Next, using insights from these multiple perspectives, this study identifies relevant psychometric variables, posits that all interventions and environmental factors fall under one of two categories, and identifies separate components of memorability.

The subcomponents of lexical acquisition constitute types of data that must, or can, be exploited during the teaching process. The categorization of interventions makes it possible to group positive and negative factors of acquisition into relatively scaled components. The psychometric variables constitute important yardsticks of the acquisition process. And the identification of separate components of memorability enables one to assess their relative contributions toward the acquisition of any given lexical target. Taken together, the acquisition subcomponents, intervention categories, psychometric variables, and components of memorability provide the underlying foundation for a pragmatic psychometric model oriented to testing interventions in lexical acquisition.

As a partial demonstration of how this model can be applied to assessing interventions in lexical acquisition, two experiments were conducted. In terms of the framework presented, the experiments tested the effect of enhancing the extrinsic component of memorability for two acquisition subcomponents (association and discrimination) at the intracharacter level.

In the first experiment, enhancing intracharacter understanding of semantic components showed little immediate effect compared to the control data, but increased retention significantly one week later. This suggests that enhancing extrinsic memorability by providing semantic associations to the recurring elements alters the slope of the forgetting curve. As evidenced by recall speeds, it appears to do this by strengthening the encoding while altering the retrieval mechanism.

In the second experiment, learning the readings of character phonetics and then learning groups of characters sharing those phonetics enhanced retention for character readings relative to the control. There was also a strong correlation between phonetic consistency within the character groups and correct reading recall. This suggests that enhancing discrimination of elements and focusing on their associated readings also aids in retention. Again, an altered retrieval mechanism appears to contribute to the enhanced retention.

Both experiments suggest that altering the processes of perception, attention, encoding, and retrieval creates alternative paths to recall in the network of lexical data stored in the brain. What remains to be done is the establishment of more precise scalars for the various types of amplification and attenuation, and to assess the effect of negative factors such as interference when new data is introduced at regular intervals. Eventually, this should make it possible to fine-tune the manner by which Chinese character and vocabulary data are learned, and to optimize curricula accordingly.

Table of Contents

List of Tables 11

List of Figures 12

Preface 13

Chapter 1 Defining the Problem: How to Fill a Leaky Bucket? 1

1.1 Reading As an Act of Construction 1

1.2 Critical Resources in Reading 4

1.3 Lexical Density and Lexical Targets 6

Chapter 2 Perspectives on the Writing System 15

2.1 Graphical Etymology 16

2.2 Classificatory Analysis 17

2.3 Phonosemantics 22

2.4 Historical Phonology 23

2.5 Morphology 25

2.6 Writing System Taxonomy 25

2.7 Differentiating Character Usage and Compositional Rationale 30

Chapter 3 Mandatory and Potential Facets of Lexical Acquisition 36

3.1 Exposure 37

3.2 Familiarity 38

3.3 Association 39

3.4 Discrimination 40

3.5 Connection 41

3.6 Retention 42

3.7 Interrelations Among the Facets of Lexical Acquisition 42

3.8 Propositions for Enhancing Lexical Acquisition 44

Chapter 4 Perspectives from Reading Theory 46

4.1 Prescriptions for Current Day Reading Instruction 46

4.2 A Closer Look at the Bottom-Up / Top-Down Dichotomy 47

4.3 Learning Words in Context 50

4.4 Alternatives to Intensive Reading 51

4.5 Automaticity Theory 53

4.5.1 Basic Definition 53

4.5.2 Models of automaticity 54

4.5.3 Prescriptions from automaticity advocates 55

4.5.4 Weaknesses in automaticity theory 56

Chapter 5 Reading As Process: Insights from Information Processing Theory 60

5.1 Basic Information Processing Functions 60

5.2 Processing Lexical Units: A Generalized Model 61

5.3 Taking a Closer Look: From Character to Word 63

5.4 Access Paths to Lexical Data 66

5.4.1 Accessing Character and Word Meaning 66

5.4.2 Accessing Character and Word Reading 67

5.4.3 The Risk of Relying on Isolated Data 68

5.5 Identification of Word and Activation of Associated Meanings 69

5.6 Phrase Level Construction 71

5.7 Higher Order Processing: Interpretation and Storage 73

5.8 Working Memory and the Limits of Processing Capacity 74

5.9 Serial vs. Parallel Processing 77

5.10 Subsumption in Lower Level Processes 78

Chapter 6 Memory Concepts Relevant to Lexical Acquisition 80

6.1 Encoding 80

6.1.2 Strategies to Enhance Encoding 81

6.2 Persistence 82

6.2.1 Lexical Networks 83

6.2.2 Consolidation 88

6.2.3 Negative Factors in Persistence: Decay and Interference 89

6.2.3.1 Memory Trace Degradation *89*
6.2.3.2 Interference Theory *90*
6.3 Retrieval **92**
6.3.1 Vector Nature of Chinese Lexical Retrieval ***93***
6.3.2 Lexical Retrieval in Reading ***94***
6.4 Forgetting **94**
6.4.1 Forgetting Curve Theory ***94***

Chapter 7 Intervention and Assessment **100**
7.1 Contextualizing Experimentation **100**
7.2 Amplifiers and Attenuators **103**
7.2.1 Amplifiers ***103***
7.2.2 Attenuators ***105***
7.3 Need for Establishing Metrics for Testing **106**
7.4 Variability in Memorability of the Target Material **107**
7.5 Assessable Properties of a Target Memory Item **109**
7.6 Applicability of the Psychometric Model **113**
7.7 Expected Effects of Amplifiers on Memory Properties **114**
7.8 Effect of Amplifiers on Longevity and Accessibility Curves **116**
7.9 Redefining Longevity Timeframes: LTM => ULTM and CLTM **119**

Chapter 8 Sample Interventions: Altering Extrinsic Memorability **121**
8.1 Effect of Understanding Compositional Rationale on Character Meaning Recall **121**
8.1.1 Method ***121***
8.1.2 Hypothesis ***122***
8.1.3 Analysis ***123***
8.1.4 Results ***123***
8.1.5 Correlation Between Study Time and Correct Recall ***125***
8.1.6 Conclusions ***126***
8.2 The Value of Phonetic Components in Learning Character Readings **128**
8.2.1 Method ***130***
8.2.2 Test Data ***132***

8.2.3 Hypothesis *133*
8.2.4 Analytical Approaches *133*
8.2.4.1 Testing the Value of Providing Phonetic Information *133*
8.2.4.2 Testing the Effect of Reading Congruence within Sets on Reading Recall *134*
8.2.5 Results *135*
8.2.6 Phonetics and Consistency *136*
8.2.7 Correlation *137*
8.2.8 Qualitative Observations *139*
8.3 Modeling the Effects of Providing Enhanced Data **140**
8.3.1 The Effect of Compositional Rationale Information on Meaning Retrieval *140*
8.3.2 The Effect of Phonetic Element Training on Reading Retrieval *143*

Chapter 9 Implications of the Framework and Avenues for Future Research **147**
9.1 Overall Implications **147**
9.2 Avenues for Future Research **153**

Appendix **159**

Bibliography **165**

List of Tables

Table 1.1 Word lists developed in China and Taiwan....10
Table 1.2 Cumulative coverage achieved by words in Zheng 1997....11
Table 2.1 Qiu classification of character types....17
Table 2.2 Approaches and aims of grapheme classification....18
Table 2.3 Basic character orientations....19
Table 2.4 False parallels at different levels of the orthography....33
Table 3.1 Relationship between exposure activities and level of data reinforcement....37
Table 3.2 Relationship between orthographic entities and associated data....39
Table 3.3 The role of connections at the intracharacter and intraword levels....41
Table 3.4 Mandatory and optional facets of lexical acquisition....44
Table 6.1 Types of consolidation cited in the literature....88
Table 7.1 Jenkins-Roediger tetrahedral model of experimental control variables....100
Table 7.2 Revised model of memory testing parameters....102
Table 7.3 Memory amplifiers and their effects....104
Table 7.4 Measurable properties of a target memory item....111
Table 7.5 Expected effects of common amplifiers on memory properties....115
Table 8.1 Raw data for Experiment 1....123
Table 8.2 Statistical measures: Wilcoxon signed rank / Cohen / r2....124
Table 8.3 Study time: raw vs. enhanced....125
Table 8.4 Correlation between study time and mean correct recall....126
Table 8.5 Typology of phonetic series....129
Table 8.6 Pedagogical usefulness of phonetics as a function of reading congruence....130
Table 8.7 Phonetic congruence of test data....132
Table 8.8 Summary of raw data: percentage of correct responses....134
Table 8.9 Analysis results for effect of providing phonetic information....136
Table 8.10 Summary of raw data: recall as a function of reading consistency....136
Table 8.11 Friedman test calculations....137
Table 8.12 Correlation between consistency and correct answers (phonetics only)....137
Table 8.13 Correlation between consistency and correct answers (including control)....138

List of Figures

Figure 0.1 Key factors that bear on orthographic mastery xvii
Figure 1.1 Resources required to process text 4
Figure 2.1 Sampson taxonomy of writing systems 26
Figure 5.1 General model for scanning entities 61
Figure 5.2 Parsing Chinese words from strings of characters 64
Figure 5.3 Chunking words into phrases 72
Figure 5.4 Meaning construction and background monitoring 74
Figure 6.1 Lexical associations based on textbook materials and semantic schema 85
Figure 6.2 A network organized based on orthographic principles 86
Figure 6.3 Approximate shape of forgetting curves 96
Figure 6.4 Forgetting curve as statistical measure of retrievability 97
Figure 7.1 Longevity and effective memorability 118
Figure 8.1 Correlation between study time and correct response 126
Figure 8.2 Phonetics used to group characters 131
Figure 8.3 Subject scores 135
Figure 8.4 Consistency vs. recall 138
Figure 8.5 Paths from graph to meaning recall: direct and alternative approaches 141
Figure 8.6 Algorithm for character meaning recall: direct and alternative paths 142
Figure 8.7 Paths from graph to reading recall: direct and alternative approaches 144
Figure 8.8 Algorithm for recall for character reading: direct and alternative paths 145

Preface

The languages that the US Foreign Services Institute (FSI) has traditionally identified as most difficult for English speakers to master—Arabic, Mandarin, Cantonese, Japanese, and Korean—represent three sets of languages sharing little in common. Spoken Arabic is a Semitic language characterized by what English speakers, in their kindest estimation, could only describe as a challenging phonemic inventory. Spoken Mandarin and Cantonese are largely isolating and obviously share genetic affinity with each other but otherwise present no extraordinary difficulties beyond tonality. Japanese and Korean are rather synthetic languages that share an unfamiliar word order from the English speaker's standpoint.

Curiously, tonality, word order, and unfamiliar sound inventories are found in a host of languages that do not fall in the FSI level three category. Thai, an isolating tonal language like Chinese, Turkish, a relatively synthetic language in the vein of Japanese and Korean, and Hebrew, a Semitic language related to Arabic, all occupy the FSI's level two category, meaning that they require half the amount of time to acquire advanced proficiency. The above spoken-language challenges notwithstanding, one then cannot help but conclude that the overriding factor must lie in the fact that these five languages are associated with exceedingly difficult orthographies—the immediate result of which is that reading cannot readily serve as an avenue for language input until considerable effort and progress have been made.

In the case of Chinese, orthographic complexity takes the form of thousands of characters that students must memorize to read authentic materials. Whereas students of Romance languages will be well on their way to reading authentic materials by their second or third year of college-level instruction, East Asian language (EAL) students will be struggling to read small quantities of even prepared materials by that time. The first and second years of EAL instruction are invariably taught using textbooks that focus on the basics of grammar and fundamentals of speaking, reading, and writing. The orthography issue looms larger in the third and fourth years of instruction, however, when students leave the confines of prepared texts. Even with the aid of vocabulary lists, the cognitive processing required can make reading one or two pages of material quite

laborious. And if forced early on to tackle unprepared, authentic materials, students are liable to spend far more time perusing a dictionary than actually reading their assignment. This barrier ultimately defeats all but the most talented and determined, as dropout rates from Asian language programs tends to be significantly higher than for easier languages. In his review of Clavis Sinica, a software product for teaching Chinese characters, Everson had this to say: "As students move out of pedagogical materials that carefully control the type and amount of new characters they learn, they often confront authentic materials that inundate them with new characters. Glossaries, dictionaries, flashcards and, regrettably, student attrition are often the byproducts accompanying this process." (Everson 2003) Such attrition can be "shocking", with up to 94 percent dropout for Chinese learners lacking a Chinese family background (Sydney Morning Herald, May 29, 2010).

Various materials are available to help students learn characters. Such materials range from basic character lists with essential readings and meanings and stroke counts (e.g. Choy 1990) to those with character etymologies and/or mnemonics (e.g. Harbaugh 1990). For the most part, however, university curricula tend to use a fairly bare bones approach to teaching the orthography, and tend not to incorporate such materials directly into the curriculum. Students may be given some instruction on the stroke order, meanings, and readings of individual characters in the initial two years, but afterwards such information is provided at the discretion of the instructor, and is generally not provided at all, the implication being that it is the student's responsibility to learn such information if desired. As a result, beyond that point, students tend to learn to write and pronounce characters in the context of another word, without necessarily taking the trouble to learn the meaning of the individual character; or they employ extracurricular tools independently of the formal curriculum.

In a sense, the teaching community faces a dilemma. On the one hand, students clearly struggle with authentic materials, and need extensive annotations to handle vocabulary in their third and fourth years of study. On the other, there is the sense that, given the typical four-year college curriculum, if they do not learn to read newspapers by that time, the curriculum will have failed. And so, ready or not, students are

subjected to authentic reading materials at a stage when reading such materials is a slow and painful process.

I believe that this no frills approach—characterized by a rush to read authentic materials without slowing down for the orthography—may be a mistake. In orthographically and lexically proximal languages, the native English speaker may have to contend with a few phonetic or phonological differences but otherwise enjoys the benefits of working with a largely similar system. The orthography essentially has two levels: space-separated words and a very limited set of letters that compose them. In somewhat more challenging orthographies like Thai or Hebrew, there tends to be a larger inventory of symbols in the script, and there are significant complications—such as writing right to left, the lack of overt word parsing, or extreme redundancy in symbol-to-phoneme mapping—that native English speakers would inevitably view as challenging. Even so, these are evidently overcome with time.

But in the case of Chinese, not only are the words not overtly separated in text, the symbol inventory balloons to—depending on one's final target—some 3500 - 5500 characters comprising roughly 1200 - 1300 recurring elements.[1] Many of these elements look similar enough to the untrained eye to be easily mistaken for one other. Although these elements often have meanings and sometimes even readings of their own, they also embody tremendous amounts of phonological and semantic redundancy. The lexicon has few cognates relative to English, and the few that exist are scarcely recognizable because of their highly altered pronunciation. And despite the presence of phonetic clues in the writing, the characters do not reveal the sound in any obvious fashion.

Because of these tremendous differences, alternative approaches to orthographic proficiency desperately need to be tested. Like his Romance language counterpart, the Chinese language student simply wants to read. But considering the need to parse words from text with no overt delineation and the need to recognize several thousand

---

1 Although educated native speakers are said to recognize from 5000 - 6000 characters, statistical studies indicate that approximately 3100 characters should cover roughly 99.7% of the characters in most texts, leaving the learner with only three unfamiliar characters out of every thousand. 3500 characters is said to account for 99% of words in text. Highly educated native speakers are thought to be able to recognize approximately 5500 characters. (Dew 1999, 35)

individual characters built using a thousand or so elements, it is quite clear that the orthography contains an extra layer of material to contend with. This additional layer in the orthography, and the concomitant additional cognitive overload that it induces, suggest that the orthography ought to be handled explicitly and strategically.

But where does one begin, and what issues should be prioritized if teachers are to reduce the student's burden? Simply asking the question leads directly to a vast body of potentially relevant research. Cognitive psychologists have sought to determine whether lexical meaning is accessed by a phonological or semantic route. Access of the meanings and sounds of words in print is also measured in terms of its speed, which is thought to be a function of both the recency and frequency of past input. Other concepts that bear on the issue of reading competence and lexical acquisition stem from writing system theory, which can be attacked from many angles. A host of memory theories have been offered, with attrition, forgetting curves, and the strength of memory playing key roles in this context. And within the field of second language acquisition, one can explore information processing theories and assess the relative importance of extensive and intensive reading approaches in developing reading proficiency.

While much can be gleaned from the volumes of research conducted to-date, one understandable tendency is for experimental research to focus on individual facets of the greater problem separately, and this can make it difficult to see how the various facets fit together. Though they do shed some light on the issue of how to tackle the orthography, those engaged in unilateral assessments risk drawing conclusions that lack context. Similarly, an overemphasis on language universals can enhance the tendency to overlook important differences that weigh on pedagogical prescription. I believe that such shortcomings derive in part from the lack of a coherent framework for assessing interventions in lexical acquisition.

The aim of this study, then, is to shed as much light as possible on what can be done to minimize the impact of this extra layer in the Chinese orthography, and to provide a rudimentary framework for contextualizing interventions in the field.

This dissertation is loosely divided into three parts. The first identifies unique characteristics of the orthography, establishes tentative lexical targets, explores the role of lexical density, and determines which perspectives on the writing system are relevant

from a pedagogical standpoint. This part closes with an identification of optional and obligatory subcomponents of lexical acquisition, and an assessment of their strategic merits.

The second part looks at reading and cognitive theory to identify concepts relevant to the topic at hand, including lexical access, information processing theory, encoding, forgetting, and automaticity theory. This part ends with an attempt to distill years of memory research to a few critical principles that weigh critically on Chinese reading acquisition, and establishes a crude framework for assessing interventions.

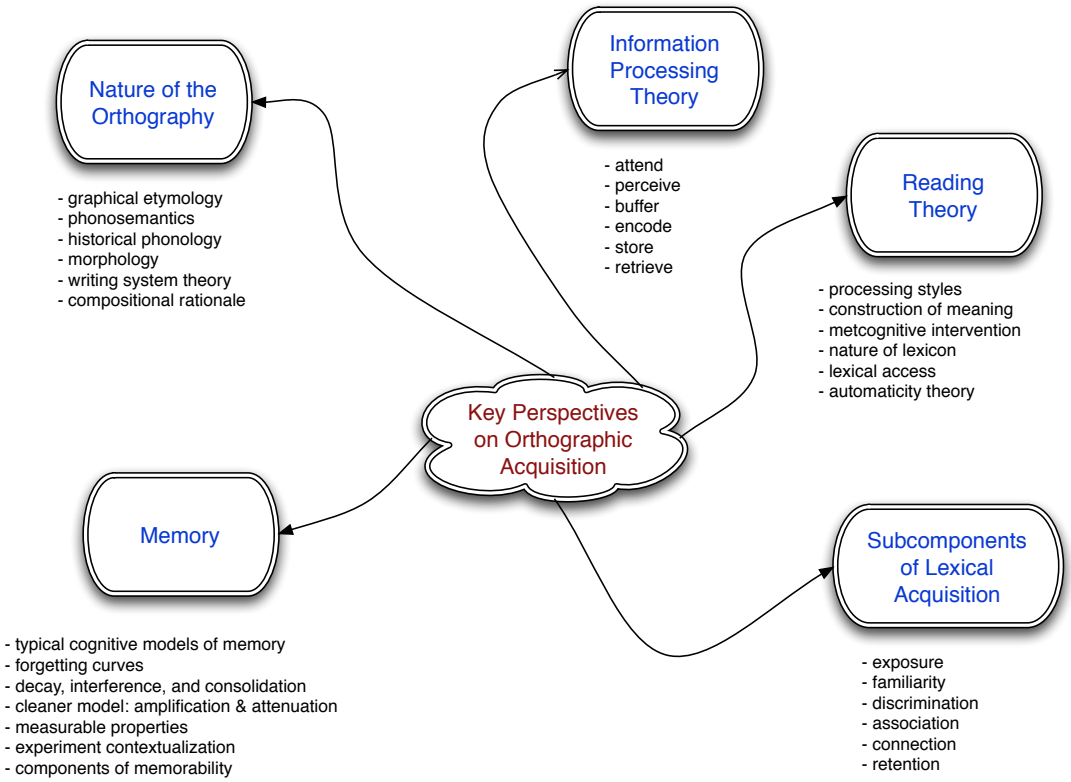


Figure 0.1 Key factors that bear on orthographic mastery

The third part explores the results of experiments conducted as a part of this study to test discrete aspects of the proposed framework. Although these studies are but a small fraction of possible avenues worthy of exploration, they demonstrate the kind of quantitative studies that can be applied once a psychometric model has been established. I hope that they will eventually point to new directions for Chinese as a

Foreign Language (CFL) reading pedagogy, and should underscore the importance of attending to the additional layer identified earlier in the dissertation.

In structuring this dissertation, although I made every attempt to keep distinct domains of discussion separate, previously discussed topics are sometimes revisited in light of newly presented information. Such reexaminations are inevitable given the interplay of the various factors involved.

Getting Chinese language students to learn thousands of characters and even more compounds is a messy and haphazard process that constitutes a perennial challenge to the teaching community. In the standard academic year curriculum, less common vocabulary is often forgotten after extended periods of neglect. During summer intensive classes, words and characters lists can be so overwhelming that many items are forgotten almost as soon as they are learned. The aim of this framework is to help make it possible to predict what levels of retention can be expected within any given curriculum and determine when a given intervention is most appropriate. My hope is that it can eventually help teachers systematize the lexical acquisition process and develop curricula that teach the Chinese writing system more effectively.

Chapter 1
Defining the Problem:
How to Fill a Leaky Bucket?

Although the primary field of inquiry of this dissertation is lexical acquisition, it is important to step back and examine the problem from the higher level of *reading*, since that, and not the mere learning of discrete words, is after all the ultimate aim of reading instruction.

1.1 Reading As an Act of Construction

What do Chinese native speakers do when they read Chinese? At blazing speeds, their eyes scan huge numbers of strokes that have been assembled into meaningful patterns. These patterns, which constitute recurring elements that typically convey semantic or phonetic information, are in turn assembled into larger units that constitute syllabic morphemes. These syllabic morphemes—which in the orthography happen to correspond to characters—are in turn compounded into larger units corresponding loosely to the Western concept of word, and these in turn form part of still larger units known as phrases. This process of continually identifying elements and compounding small units into larger units results in the abstraction of familiar entities such as the sentence, paragraph, and greater discourse structure. While looking at this multitiered assembly of graphical coding for linguistic information, the reader most likely converts certain elements to sound, and in light of the graphical patterns displayed on the page, associates those sounds with specific meanings. In the process, the reader is assembling a parsed representation of the meaningful units of words in his or her mind, even though there is no overt graphical representation in the text to show how to do that. From the accumulation of reconstructed words, the native speaker then attends to higher level meaning represented at the phrasal and sentential levels, eventually cognitively appreciating such upper-level abstractions as the author's train of thought, intent of communication, and the like.

Thanks to years of schooling and practice, to an educated native speaker these processes are almost as automatic as breathing. Most of the levels of element identification, symbolic compounding, parsing, and meaning interpretation are done totally unconsciously. The accomplished native reader will typically spend no time attending to the individual strokes. In fact, he or she will not even spend that much time dwelling on the shape of individual characters, as he or she automatically groups them into words. As a result, what is technically little more than a stream of characters with no overt spacing is nevertheless viewed as if it consisted of separate words.[2] And as proficiency increases, even the words are hardly attended to, as meaning seems to simply leap off the page.

The above description of reading outlines an obligatory process. It is obligatory because it describes what must occur in purely logical terms. Change the strokes at the basic level, and the semantic and phonetic significance of the character elements changes. Change the significance of these recurring elements, and the meaning of the characters changes. Change the meanings of the characters, and the words' meanings will also change. As these effects inexorably percolate up to the higher levels of phrasal, sentential, and discourse-level interpretation, there should be no denying that a continual construction of meaning must take place from the lowest to highest levels.[3]

The above description of reading is also incomplete, because it makes no mention of a host of extremely complicated physiological and biomechanical activities that must take place for reading to transpire successfully. The reader coordinates the extraocular muscles to trace along the text so that the eye skips along at certain intervals, pausing periodically to absorb visual information. Light images converge on the retina, conveying information via the optic nerve through neurochemical impulses to the brain, and in the process, these sensory impressions are somehow interpreted as meaningful

---

2 Because words are not parsed overtly in the orthography, the concept of word is subject to numerous interpretations in Chinese (Packard 2000, 14-20; Chao 1968, 136; Duanmu 1997). Even so, I believe it is safe to assume that on some level, the concept applies to Chinese just as much as in languages writing systems where in which words are parsed overtly in the orthography.

3 The inevitability of the need to construct meaning from the individual elements has been at times disputed by top-down theorists. See below1.2.2.

visual information. Meanwhile, a host of other larger muscles, starting with the head and neck and on down to the rest of the torso, must support the body during the process. And this makes no mention of the more vital physiological functions of the heart, lungs, circulatory system, and so on. Fortunately, most of these processes need not concern us.

Turning back to the notion that this description of reading is obligatory, in the sense that it involves the process of scanning and identifying significant orthographic elements and from those elements constructing meaning at higher and higher levels, in the case of a non-native speaker trying to learn Chinese, achieving even a modicum of competence at each of these levels is extremely difficult. Incorrectly recognizing any of the numerous recurring graphical elements at the base level within the stream of text is likely to result in errors that inevitably propagate further up the chain of identification and meaning construction. The time it takes to learn to recognize the 4000 characters or so needed to attain fluency in reading is in itself nontrivial; automatically associating these characters with meanings, and worse yet, knowing which characters in a text go together to form meaningful words is going to require a huge amount of prior exposure, study, and practice. The learning of the various meanings that an individual word may impart in various contexts is another level requiring the assimilation of an enormous amount of information. And finally, because Chinese is after all not the non-native speaker's mother tongue, there will be a host of twists and turns to be navigated originating from the text's syntax, idiom, cultural assumptions, and even the idiosyncrasies of the author. But irrespective of all these other factors, the crux of this obligatory model is one simple observation: failure to recognize elements will lead to failure to recognize characters, which in turn will lead to failure to recognize words.

Although people learn to read foreign languages all the time, in the case of a language like Chinese, the task is more difficult than usual. One obvious question is whether there is a way to measure the extent of the problem relative to a hypothetical goal of "acceptable lexical competence." Put another way, what does it take to be a skillful reader?

1.2 Critical Resources in Reading

As described above, the very ability to read implies the existence of critical resources, both innate and acquired, that must exist for the process to transpire successfully. It is obvious that skill in reading will require a solid lexical data store together with at the very least a layman's understanding of syntax, phonology, and morphology. Some store of "world knowledge" incorporating the schemata of some language theorists will naturally also exist. Of these components, while it is clear that neither world knowledge nor grammatical prerequisites (syntax, phonology, morphology) are likely to be significantly more difficult in Chinese than for any other language,[4] lexical data acquisition is clearly a significant impediment.

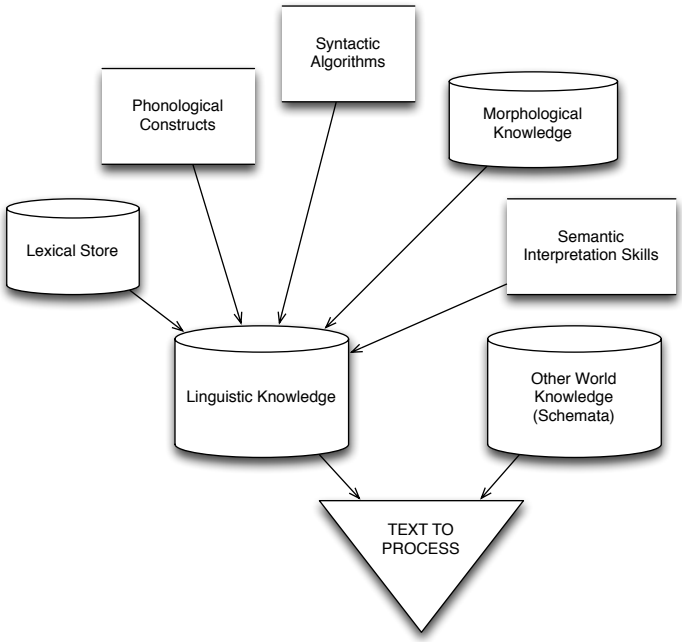


Figure 1.1 Resources required to process text

Starting out with first approximations, if a resource model were to be loosely sketched out, it might look something like the model shown in Figure 1.1. Given that the bulk of formal grammar can be imparted (though not mastered) within a couple of years of college level instruction, it seems clear that the largest impediment will constitute

4 It could even be argued that the relative simplicity of Chinese grammar makes it easier to learn than many languages.

vocabulary. Lest this surprise the reader, consider the rather staggering differences in vocabulary level between what a normal native speaker learns between elementary school and high school and what the typical Chinese student learns in the first two or three years of study. Conservative studies on vocabulary acquisition estimate that five-year-olds know about 2000 words, and that children on average learn approximately 3000 words a year between the ages of 7 and 16. This means that by the time they are in high school or getting ready for college, they will know at least 30,000 words (Gathercole and Baddeley 1993, 41). By some reckonings, advanced high school students may know far more words.[5]

Taking the *Integrated Chinese* (Liu et al. 1997) series of textbooks as a rough measuring stick, if one tallies the vocabulary index at the back of the textbooks, one will find that the two years of instruction introduce approximately 3500 words. Assuming, rather optimistically, that the student remembers every single word he or she was exposed to, then by the third year of college he or she will have the lexical depth of a six year old. True, the actual words known may be quite different, but the overall capacity is approximately the same. Is it any wonder, then, that when faced with authentic reading materials, the student struggles, or that, lacking constructive and coherent approaches to acquiring a larger vocabulary, few ever attain meaningful proficiency in the written language? In the words of Gathercole and Baddeley, second language learning inevitably requires "massive vocabulary growth," almost doubling the phonological forms the individual is familiar with. Given that, in the case of Chinese, it also means an exponential growth in familiarity with individual graphical elements (characters), this one aspect of the resources diagram clearly deserves additional attention.

---

5 Other sources place the high schooler's vocabulary at nearly 80,000 words, which makes the prospect of learning a meaningful level of Chinese that much more daunting (Hirsch 2003).

1.3 Lexical Density and Lexical Targets[6]

While recognizing that a great number of factors come into play with respect to a learner's ability to comprehend a given passage of text, I would assert that assuming the student has taken 2-3 years of college-level coursework, he or she will have sufficient training in the basic syntax and grammar needed to comprehend most materials. At that point, assuming the subject matter itself is not overly technical, or does not draw on background knowledge completely outside the reader's capability even in his or her own native language, then what is left is the lexicon. Accordingly, with relatively proximal languages such as Romance languages, despite several aspects of complexity not found in English, the typical second or third year student will be well on his or her way to reading authentic materials. Naturally, at that point, the issue will be to be able to handle various styles, genres, and content. Given that fundamentally Chinese grammar could be said to be even easier than that of Romance languages, while at the same time noting that there are almost no cognates from the standpoint of vocabulary, the primary impediments arguably derive from the orthography itself and the need to acquire sufficient vocabulary to process text.

For Japanese and Chinese, vocabulary is spoon-fed via the textbooks selected by the instructor. Somewhere in the third or fourth year, however, instructors will begin to use authentic materials or perhaps slightly simplified materials. It is at this stage that the learning curve suddenly becomes much steeper, despite the drop-off in "grammar" instruction per se. Although there is some increase in sentence structure complexity, the key impediment appears to be none other than *lexical density*—the percentage of characters and vocabulary that are either unknown or barely familiar to the learner.

By way of example, a newspaper article presented to third-year students of Japanese or Chinese would almost certainly be extremely challenging from the standpoint of vocabulary, constituting a work of high lexical density. Even when the syntax appears relatively straightforward, because of the high percentage of unfamiliar or scarcely familiar words, student comprehension can be slow and laborious. A fairly

---

6 The term I am defining here has nothing to do with the way "character density" is defined by others, wherein the term simply refers to the number of strokes in a character (Hayes 1987; McEwen 2006).

good measure of the level of vocabulary density in a pedagogical text would be the vocabulary lists associated with the text, and the percentage of new words that occupy the text. Typically, because curricula designers are eager to have students read authentic materials, vocabulary density in advanced textbooks will be as high as 20%, with every fifth or sixth word being essentially new. The immediate result, of course, is a predictable effortful decoding that makes it difficult for the reader to attend to main threads in the discourse.

By contrast, examples of low-density materials would be texts such as those employed in extensive reading series. These purposely minimize complexity at the syntactic and lexemic levels.

A curious philosophy that I have sometimes encountered in educational settings has been the assumption that it is actually good for students to look up words in a dictionary. Obviously, part of learning to use any language involves learning to use the dictionaries of that language. But if a text that native speakers could read in a few minutes ends up requiring 2-3 hours perusing a dictionary and another hour or two decoding, then it becomes painfully clear that the student is not learning to read but is practicing looking up words in a dictionary.[7] For students to avoid being trapped in an overly laborious sensation of decoding, the proper vocabulary density should be established at some reasonable level. One in ten words should probably be considered difficult, and one in twenty fairly manageable. [8]

---

7 Moreover, if my own experience is anything to gauge from, the actual act of looking up a word myself in a dictionary does not in any way increase the likelihood that I will remember the word. I am perfectly aware, however, that for whatever reasons including personal background and upbringing some teachers believe that looking up the word will "force" students to remember the word. Unfortunately, as memory theorists and researchers point out, there is no real program in the brain that one can switch on to remember something. Certain activities enhance memory, but telling oneself to remember something in and of itself is no guarantee of anything.

8 Note that I have failed to mention idiom. The reason for this is that I believe idiom to be nothing more than a special case of vocabulary. Any word in a language is liable to have more than one meaning, and combinations of words are naturally going to convey a more specialized idea or concept or description. The fact that certain words might appear to mean one thing but in combination or in certain contexts mean something else simply suggests that there is an additional level of meaning to the word or words in question. Conceptually, this is no different from the phenomenon of multiple meanings for any given lexeme.

It could be argued that the student can work through this, and that by going over the vocabulary list (if one is provided) and reading and re-reading the passage, he or she will eventually be able to read it fairly fluently, with some ability to attend to higher processing levels such as what is going on with the text. However, what happens when he reads the next article or next chapter? Unless the material was chosen to be very similar—and often, out of a desire to expose students to as many subjects as possible, the very opposite is true—the exact same process will repeat itself. If a list is provided, much mental energy will go to learning the list, and if there is no such list, even more time will be spent in a dictionary. Even if one believes, as some instructors seem to do, that eventually a sufficient number of words will be learned by this process, it is easy to overlook one simple fact. Because so much time will be spent on the new list, items from the old list are bound to be forgotten. This is especially true in a summer intensive course, where students may be expected to learn from 30 to 40 new words a day, this being the equivalent of 10 to 20 new characters, if not more. For this very reason, it is not uncommon for teachers to "give in" and not hold students accountable for previous material past a certain date.

This is the irony. On the one hand, university curricula or language schools abroad often boast about the level of characters and vocabulary the student "will have learned" after completing a given year of the language program. But the white lie no one wants to admit to is the fact that the operative words are not "will have learned" but "will have been exposed to." Forgetting gets in the way of the ability to reliably accumulate the vocabulary required to achieve higher levels of reading proficiency. Given the resulting combination of seemingly constant mental effort, relatively little headway in terms of reading volume, and the dreadful feeling as if one were filling a leaky bucket, it is any wonder that so many students drop out from their programs?

A primary goal of reading pedagogy should be to establish the ability to read independently, meaning the ability to read authentic materials without frequent recourse to dictionaries or outside assistance. For English, estimates suggest that approx. 5000 words would be a minimum to read simple novels or other authentic materials by while being to able to guess most words from context (Nation and Waring 1997; Hirsch and Nation 1992). Relative to any given text, at least 95% text coverage may be needed to be

able to successfully guess the meaning of unknown words from context (Liu and Nation, 1985). At the same time, Nagy contends that to learn a significant amount of vocabulary each year, the foreign learner must read on the order of one million words per year (Nagy 1997). While this may be possible for English speakers learning Spanish, anyone sitting in on a fourth-year Chinese class will clearly appreciate that under the current circumstances such a target is totally unrealistic.

This last estimation, that one should be reading approximately one million words per year, is eye-catching. How much is one million words? And are students getting anywhere close to that level of exposure? I thought it would be interesting to see how this volume stacks up against an "advanced" textbook. *Thought and Society—An Advanced Text for Spoken Chinese* (ICLP NTU, 1997) is a book with a very high vocabulary density and is considered a core textbook at the International Chinese Language Program in Taiwan. While the definition of word is notoriously difficult in Chinese, I estimated that each of its ten articles has approximately 600 words each, or a total of 6,000 words. And this course is given over the course of a summer intensive course, with students expected to take another course using a comparably long textbook. By this rough estimate, students at ICLP in the "advanced" program might read 12,000 or optimistically 15,000 words over the course of a summer. Assuming they did this all year, they might be lucky to see 60,000 words of text. Of course, most college students, cannot afford to study Chinese all day every day of the year, so that even that projection is unrealistic. These numbers alone give one pause. How can the curriculum possibly succeed if text coverage is so low?

Because they are easily counted by an parsing program, the notion that 3000 to 5000 characters will respectively cover 97% to 99.7% of text has been a fairly stable one for some time. By contrast, researchers have had a harder time identifying an appropriate level for words.

Targeting mainly foreign learners while keeping an eye on general Chinese pedagogy, various scholars and institutions have tried to identify core Chinese vocabulary, ranking the identified words by importance. Ke Huawei et al. summarize

some of these early efforts (Table 1.1).

Table 1.1 Word lists developed in China and Taiwan

List	No. of Words	Notes
HSK	8822	Said to cover 95% of all text developed in the early '90s in the PRC
Ye Deming	15262	Based on a 360,000 word corpus of teaching materials, from which 8088 were selected to cover 95% of all text found in an assortment of Chinese language teaching materials parsed programmatically
Zhang Liping (2002)	10,155	Words based mainly on Chinese language teaching materials
Ministry of Education Chinese Language Committee	14469	Between 45,000 and 62,000 words were obtained in each of four separate years, and using this as a basis, a list of 14,469 core words was derived.

Some of the above sources suffered from deficiencies of various sorts. For example, the HSK list was based on a relatively arbitrary criterion, namely, the selection of words appearing 10 times or greater within the Modern Chinese Frequency Dictionary (現代漢語頻率詞典), with some *ad hoc* adjustment made to account for the needs of foreign learners. Zhang 2002 was based exclusively on Chinese language teaching materials. The list from the Taiwanese Ministry of Education used different types of materials in the four years under consideration, making comparisons of relative frequencies somewhat suspect, especially since the size of each year's corpus differed substantially from one year to the next.

Recognizing the deficiencies in the various lists developed so far, Zhang et al sought to develop a list of target vocabulary for pedagogical purposes. Using a weighting system, they graded lexical items based on where they appeared in previously developed lists. For example, something in the "most basic" category of HSK might be awarded 4 points. By comparing the relative gradings for words in the various lists they used, they obtained a list of approx. 8000 words. This list would appear to be more balanced than previous lists from the HSK and other glossaries, as gradings (which varied from one source to the next) were averaged out. They then sought to compensate for the lack of representation by spoken dialog by developing frequency tables for other

transcriptions of spoken material into their formulas. To this end, they assembled various materials that represented spoken dialog and used a program to identify word frequencies for their new material. They then calculated what they term "relative frequencies" (相對頻率) for words in the various lists, this "relative frequency" being the number of occurrences divided by the size of the corpus in question. They then arrived at a final "weighted frequency" adding the respective relative frequencies to the original summation of weights.[9]

Since there was considerable discrepancies among the original word lists as to what was basic, intermediate, or advanced, their attempt to balance the levels should prove helpful. This alone makes their list one of the best starting points for lexical statistics developed to date.

Less helpful is the fact that they rather arbitrarily limited their list to 8000 words, largely on the basis of arguments made by C. C. Cheng that 8000 words is an approximate maximum for active vocabulary by native speakers (Cheng 2004). Compare this notion of 8000 words with the text coverage rates estimated by Zheng Zhaoming cited by Zhang.

Table 1.2 Cumulative coverage achieved by words in Zheng 1997

Grade	No. of Words in Grade	Cumulative Word Count	Rate of Text Coverage
1	2270	2270	50%
2	1440	3710	75%
3	2373	6083	95%
4	4093	10176	99%
5	7636	17812	100%

9 I should point out that mathematics of their approach was rather unusual. When assigning weights to values from three corpuses, Zheng et al. used whole numbers (4, 3, 2, 1). By contrast, relative frequencies were calculated as the number of occurrences divided by the total number of characters in the corpus. Needless to say, the relative frequencies are extremely small fractions (e.g. 0.000224). The result of this approach, in essence, is that a tremendous amount of value is being placed on the derived word list, and very little to the relative frequencies from the spoken language corpuses they had developed independently.

Zhang et al. point out that Zheng's list has pedagogical limitations if viewed as a series of levels to be studied in a given curriculum. Specifically, each set of words (each grade) is sized unevenly, making pedagogical implementation rather awkward. However, this shortcoming is easily overcome by moving the percentage yardsticks, rather than setting them based on "tidy" percentages. Some number crunching could therefore balance out the lists, even if the yardsticks appear random (67%, 82%, etc.). If Zheng's list is at all accurate, Zhang et al.'s list of 8000 words will provide somewhere in the neighborhood of 97% text coverage.

Obtaining well-balanced, graded lists of vocabulary is an extremely important first step in defining what words students should learn and how many. That said, such lists do not directly answer the question of how much is needed to read an average text and to become an independent reader. As mentioned earlier, Liu and Nation argued that 95% text coverage (corresponding 6000 or so words in the table above) might meet the criterion of adequate reading combined with the ability to guess words from context. However, others argue that a higher number is needed for truly independent reading or reading for enjoyment, somewhere about 99% being likely to become truly enjoyable and fluent.

For argument's sake, let us set a threshold of 10,000 words (approx. 99% coverage according to Zheng) as a minimum to read most authentic materials with little recourse to a dictionary and with good ability to successfully guess words from context. Over a four-year college curriculum occupying just over nine months, one can count on approximately 200 days of instruction, or 800 over a four-year period. By that rough calculation, to reach the target of 10,000 words, one would need to learn 12.5 words per day *cumulatively.* That is assuming one never forgets a word over that entire period of time, something which studies have shown is just not the case (Hayden 2003). As for characters, taking the lower figure of 3000 characters, that means 3.75 characters per school day. That value does not seem so intimidating, but when one considers that the rates of re-exposure to individual characters will vary greatly depending on the character's frequency count and the nature of the material to which the students are exposed, natural upkeep is far from assured.

* * * * *

This chapter has defined the problem as follows. A very large number of characters (between 3,000 and 5,000) and a very large number of words (between 8,000 and 10,000) must be learned in a few short years. The degree of familiarity with this lexical material must obviously be good enough that the reading can proceed at a reasonable speed. Other resources, such as implicit or explicit understandings of phonology, morphology, syntax, and semantics go without saying, but are not expected to be greatly different in difficulty compared to other languages. By contrast, the lexicon itself poses an enormous hurdle. Current estimations of the size of materials used in courses today show the amount of text exposure falls far short of that needed for vocabulary to be learned naturally through independent reading.

Meanwhile, subsequent experimentation may some day determine which level of vocabulary density is ideal. Any such ideal would inherently consist of a statistical average for students at a given level, for within a class there will obviously be significant variation in the vocabulary level of individual students. If new pedagogical approaches come to significantly enhance vocabulary acquisition while safeguarding retention (minimizing forgetting), lexical density would also have to be redefined. But the notion is important, and the reason that high-density texts results in decoding will become apparent shortly.

In a nutshell, the problem stems from a basic pedagogical dilemma. Teachers want to expose students to as much new vocabulary as possible so as to achieve some target where reading can begin to take place for pleasure, and thus become a self-motivated activity. On the other hand, that very desire for quick exposure results in unacceptably high lexical density. Reading excessively dense material results in a laborious reading pace, inhibiting one from ever reaching the reading volume needed to consolidate what one has learned. At the same time, the relatively fast pace of many curricula results in forgetting, the "leaky bucket" phenomenon, such that new words are learned, but many are forgotten along the way. This means that vocabulary levels never get anywhere near that needed for independent reading. And until a critical lexical threshold is reached, this situation perpetuates itself indefinitely. Clearly something needs to be done to resolve this issue.

Before addressing this immediate problem directly, one might ask, Is there something about the writing system besides vocabulary that makes things so difficult? In other words, is there something one can glean from an analysis of the writing system itself?

Chapter 2
Perspectives on the Writing System

If the key to cracking the Chinese orthography lies in placing special attention on additional layers in the orthography, then one ought to first clarify the nature of the orthography itself. Imbalanced analyses of the nature of the Chinese and Japanese writing systems can invite erroneous conclusions about the best pedagogical approaches for teaching character-based writing. In the literature on Chinese and Japanese pedagogy, it is not uncommon to find scholars who believe that a focus on characters, graphemes, and morphemes might lead students to suffer an effortful decoding associated with bottom-up processing. Some insist that successful students do not make use of the "logical links" within the elements of characters (Unger 2004, 83). Others cite psycholinguistic literature to assert that characters are processed phonologically, implying that that reading *must* be based on spoken language competence (Brocket 2004, 141).

Such statements are sometimes used to downplay the role that character study plays in lexical acquisition, with the unfortunate consequence that character study ends up being back-burnered in favor of whole-word study. This chapter seeks to provide a more balanced view of the orthography so as to clearly identify what information about characters can and should be taught to students.

Like the proverbial elephant examined by separate blind men, the Chinese writing system can yield to strikingly different descriptions depending on the perspective taken. Some of these perspectives relate not strictly to the writing system per se, but rather to features of the language itself. These too will be included, mainly because their analysis is inevitably conducted with reference to the writing system, they offer helpful insights into the orthography, and their relation to the system as a whole merits clarification. Most importantly, I will look at how the roles these perspectives play can become confused. The multilevel analysis provided towards the end of the chapter offers a tool for avoiding such infelicitous admixtures. After briefly visiting these perspectives, I will identify what they have to offer the student learning to write and read Chinese.

2.1 Graphical Etymology

Graphical etymology is the art of tracing of character forms, deciphering the rationale behind their composition, and tracking their use over time. Known in Chinese as *wenzixue* 文字學, this field of inquiry traces back as far as the *Er-ya* 爾雅 (3rd century BCE), and in more systematic terms, to the *Shuowen jiezi* 說文解字 (first century CE). Graphical etymology may be one of the first things that comes to mind when laymen think of character analysis.

Although a wide variety of radical systems have been used over the centuries—with the number of radicals varying widely from some 550 to the present day number of 214 (Bottero 1996), A systematic classification of Chinese characters known as the *liushu* 六書 or "six principles" was provided in the *Shuowen jiezi* 說文解字 Although this classification system is well known even to this day, the *liushu*—which constitute pictographs, indicative graphs, ideographs, phonetic-semantic compounds, mutually explanatory graphs, and phonetic loans—are clearly outdated, with the mutually explanatory category being viewed as particularly problematic.[10]

Relatively recent advances, namely those made through 18th century philology and those arising with the advent of modern linguistics, eventually led to new interpretations of how Chinese characters should be classified. These more modern interpretations tend to focus on how characters were composed, though there was some mixing in of classifications based on usage and or relation to other characters.

A more modern typology of Chinese characters can be found in Qiu Xigui, who reduces Chinese characters to three types: semantographs (*biaoyizi* 表意字), phonograms (*xingshengzi* 形聲字), and loan graphs (*jiajiezi* 假借字). These three categories delineate Qiu's view on how characters could be composed or selected to represent a given morpheme or word, and are defined in the table below.

Qiu provides numerous examples that delineate subcategories for each of these character types. In the process, he identifies compositional processes such as component substitution, component abbreviation, and other forms of element restructuring. In

[10] Qiu identifies as many as nine possible interpretations for this category (Qiu 2000, 157-161).

Table 2.1 Qiu classification of character types

Description	Process of character composition / assignment to morpheme
semantograph	The character designer chose to use either stylized pictures or abstract symbols either singly or in combination to allude to the meaning of the word to be represented by the graph.
phonogram	Through a variety of processes, the character designer combined one or more semantic elements and one or more phonetic elements to represent the word-morpheme in question.
loangraph	Rather than creating a new character graph, the character designer adopted an existing character and borrowed it strictly for its sound. Semantic implications of the graphical composition employed are usually absent, but are theoretically possible.

addition, he identifies a wide variety of relationships between graphs. For instance, two separate characters having different compositions but used in relation to the same morpheme (a binding of sound and meaning) are referred to as allographs. The evolution of one graph (a "matrigraph") into a later differentiated form is identified as a process for reducing the lexical load of an original graph. Likewise, there is a wealth of analysis on the process of borrowing and its effect on the writing system.

Qiu offers an excellent single-volume accounting of developmental patterns in character composition and usage. However, in tracing the graphical and etymological perspectives of very complex cases, his approach muddies the divide between two separate phenomena: the process of character composition and the way characters are employed in text. That failure to clearly separate compositional rationale from character use can lead to a misunderstanding of the nature of the Chinese writing system. Section 2.8 below provides a cleaner model from both theoretical and practical standpoints.

2.2 Classificatory Analysis

Another avenue for analyzing the orthography is to dissect graphical structures and classify characters accordingly. Stallings identifies two primary main aims for classifying Chinese characters: to provide a classification method for use in search and retrieval, and to analyze the structure in a way that facilitates their manipulation on machines (1975). Machine-based implementations falls into two main categories: input

and recognition. The former includes keyboard input and handwriting input, while the latter includes printed text recognition and handwriting recognition. At the time of Stallings' survey in the mid-seventies, the number of classificatory approaches was quite large and varied, with most being aimed at one or more of the aims described in Table 2.2 below. The table summarizes some of the principle approaches adopted for the classificatory approach undertaken.

Table 2.2 Approaches and aims of grapheme classification

Aims \ Approach	Stroke-Based	Component-Based	Reading-Based
classification for lookup	Y	Y	Y
handwriting recognition	Y	N	N
input methods	Y	Y	Y
optical character recognition	N*	N	N

Depending on the frequency domain, Chinese characters contain an average of between 9 and 11 strokes (Tsai 2005); those numbers drop by approx. 1.5 strokes for simplified Chinese. The number of different types of strokes is rather small, with, according to one classification, 8 basic strokes and 29 combination strokes (Zhonghua minguo qiaowu weiyuanhui 1997, 101-122). In some cases, such as in character recognition technologies, strokes are indeed used as a salient unit of classification.

Although strokes are the most basic unit of Chinese graphs, outside of handwriting recognition technologies, where stroke order and stroke type may be used for identification purposes, the more salient unit of composition is usually the recurring component. From the standpoint of character compositional rationale, recurring

components are used in combination to provide semantic or phonetic clues to the character's meaning or pronunciation, sometimes providing both simultaneously.[11]

Numerous scholars have examined the positional relationships of recurring components, establishing a "grammar" of sorts that divides the structure into several combinatorial possibilities. One earlier example developed by Rankin divided characters into three types: north-south, east-west, border-character (Stallings 1975). For comparison, Table 2.3 shows a more complex classification as identified by the Unicode Consortium. One configuration that is obviously missing in the Table is the unibody entity, wherein there are no divisions or enclosures at all.

Table 2.3 Basic character orientations

Configuration	Unicode Description Character
east-west divisions	
north-south divisions	
two-sided enclosures	
three-sided enclosures	
four-sided enclosure	
overlapping elements	

11 The nature of the recurring elements themselves seems to cause some confusion in terms of terminology. Many cognitive psychologists speak of "semantic and phonetic radicals," though such a terminology is ill-advised for two reasons. First, the term radical has been traditionally used to designate a conventional set of elements by which characters are classified in any given dictionary. Currently, a "traditional set" (meaning that mirroring the Kang Xi dictionary of 1685) is considered standard, though historically, different radical sets have existed. Secondly, the term radical itself is misleading in that it suggests that the component in question is somehow the "root" of the character, when in fact, nothing could be further from the case. A radical is simply an arbitrary classificatory element, usually but not necessarily of semantic import, and is frequently not at all closely related to the character's meaning.

Although the above orientations suggest that there are only twelve basic graphical structures, in fact, many of the element orientations are subject to subdivision. For example, a left-hand or a right-hand component in an east-west division could consist of an element that is in itself a north-south component. Conversely, either component in a north-south division could itself consist of two or more elements in an east-west orientation. Enclosures tend to consist of a single entity, but the enclosed element may also have a more complicated orientation, usually east-west or north-south.

As Rankin and later scholars have pointed out, many recurring components are restricted to certain positional relationships within a given graph. Thus, for example, some components are found exclusively in one position; others are somewhat free-floating, but have a positional relationship defined by the fact that they are never found in a particular position. This tendency to restrict orientations helps provide a certain element of systematicity in graphical structure that over time can be appreciated by the learner. Such systematicity may contribute weakly to recall in much the same way that phonotactics is something intuitively appreciated over time by learners of alphabetic scripts.

A more complex issue with respect to graphemic classification lies in the determination of what constitutes a recurring component to begin with. Due to complex stylization processes and the addition, deletion, and substitution of components as the script evolved, many modern recurring elements have two analytical structures: a superficial one based on well-known "apparent" components (as viewed by someone with a "naive" understanding of the system, and a deeper structure based on what is known of the element's purported significance. Reflecting terminology in vogue at the time for syntax, Burkart makes this distinction in terms of *surface* and *underlying* structures. However, such a dichotomy may not be so straightforward as all that, and it may not be possible to define such concepts in absolute terms for the simple reason that what constitutes a recurring component depends on one's aims as much as anything else.

A couple of examples will illustrate the difference. The character 莫 appears to contain three elements, plant/vegetation ⺾ at top, sun 日 in the middle, and large 大 at

the bottom. This combination of elements in this exact configuration is used in a handful of characters to provide a phonetic value; arguably, in some cases, 莫 may play a semantic role as well. Now since 莫 is a recurring element—that is, a meaningful unit in its own right—the question to ask is: how many elements does 墓 have? The bottom character is simply earth 土 and so the answer depends on how one analyzes the original element—whether as a single entity 莫 or as three separate elements (艹, 日, and 大). Burkart might have divided these into three, whereas I would argue that they are not divisible in this manner given the purported etymology. In short, if one is simply interested in analyzing graphs into their smallest unit above the stroke level, one can divide components willy nilly and state that they are legitimate elements. However, if one accounts for the component's significance—meaning, the role it plays from the standpoint of compositional rationale—such subdivisions are not always possible. At this stage, Burkart's distinction between "surface" and "underlying" structure become rather tenuous. No matter which approach one wants to take (ignoring or accounting for compositional rationale), one will have to establish to perform in real-life cases, and will further depend on one's understanding of the significance of recurring elements and composite components. The answer is certainly not answered by "native-speaker intuition" (a common litmus test employed in other areas of linguistics at the time), as very few native speakers have the expertise to pass such a judgment. The issue is further complicated by the fact that there are many characters where, quite frankly, the best information about its etymology is highly speculative and certainly not a grounds for a clear definition of layer.

As much as the structural systematicity of Chinese characters no doubt contributes to their learnability, as an avenue of explicit instruction, beyond, perhaps, a very basic notion that recurring elements and composite components can have orientations like those shown in Figure X.2, there is little to teach. Occasionally, it may be helpful for them to distinguish between semantically unrelated elements (e.g. right-side 阝 and left-side 阝), as such a distinction is important from the standpoint of lookup using the radical system. But in most cases, there is nothing about relative positions to "study" per se. After all, thousands of characters have a left-right orientation, thousands others are "north-south" characters, but the fact that they have a given orientation tells

rather little about their meaning or reading, nor will it make significant contributions to their pattern recognition skills (compare Chapters 7 and 8 in this regard for approaches that do work on pattern recognition).

In closing this section, I should point out that most of the "mechanistic" avenues of analysis identified by Stallings are now largely obsolete thanks to the current adoption of phonologically based input methods and other computer technologies for handwriting recognition and optical character recognition. It is true that better character lookup methods are still needed, but overall, many of the classification approaches aimed at input are now obsolete.[12] Needless to say, these classifications do not bear directly on instruction, but only on facilitation of lookup, and so are beyond the scope of analysis intended here.

2.3 Phonosemantics

Phonosemantics is a field that purports to identify semantic connotations within words that share some form of phonetic similarity. Among the most common examples in English is the idea that words starting with *sli-* somehow connote an ease of relative displacement between two objects in contact: unequivocal examples hence include slide, slippery, and slither. Although phonosemantics is a fascinating avenue of inquiry, one of the difficulties in grouping words together is determining just how similar in meaning words have to be to be considered part of the same word family. Extending the example just given, the question becomes whether one can include in the *sli-* group words like *slime* (a dirty viscous liquid that can slide over another surface or that facilitates sliding between two surfaces), or *sliver* (something thin, like the space between two sliding objects). The flip side of the equation, of course, is the question of just how similar-sounding groups of words must be to be treated as part of the same word family.

Three attempts to systematically identify word families in Chinese include Karlgren (1934) Todo (1965) and Wang (1982). I once examined these three systems to see whether they present a similar view of Chinese word families. My approach was to compare and contrast their results for a rhyme group that would be considered

12 See Halpern 1999 and Child 2007 for examples of innovations in this field.

relatively uncontroversial. As it turns out, these three scholars applied widely different solutions to the issue of semantic and phonological proximity, suggesting that no one has clearly spelled out the underlying semantic and phonological criteria to be applied when identifying cognates and word families in Chinese. If that is the case, then caution should be applied to the process of calling words "cognates" or members of a word family.

Phonosemantics is not a field that is well-known to students of the language. And this is probably for good reason. Recognizing that sets of words once had similar pronunciations with some degree of semantic affinity may well be helpful in historical phonological reconstructions. However, knowing that certain words at one time had similar pronunciations and some semantic affinity is not likely to help students all that much today, now that phonetic pronunciations of most of them are likely to have drifted apart.

2.4 Historical Phonology

Undoubtedly the most technically challenging of the analytical perspectives applied to the analysis of Chinese is that of historical phonology. Scholars typically divide Chinese reconstructions into two major periods—sometimes referred to as Old Chinese (corresponding to early and mid-Zhou dynasty), and Middle Chinese. Middle Chinese is nowadays typically subdivided into Early (corresponding to the language as described in the Qieyun of 601 CE) and Late periods (corresponding to the late Tang dynasty) (Baxter 1992: 14).

Major approaches for reconstructing Middle Chinese include triangulating from modern dialects to Middle Chinese; triangulating from contemporary neighboring languages such as Korean, Japanese, and Vietnamese based on their phonemic renderings of words borrowed from Chinese, and interpreting the likely meaning of Middle Chinese rhyme dictionaries and tables.

Reconstructions of Old Chinese, in turn, tend to constitute extrapolations from whatever reconstruction of Middle Chinese one happens to be using, partial triangulation from the Min dialect, and identification of rhyme groups (groups of characters felt to rhyme based on their usage in poetry). A more recent trend is to

include intermediate data between what has traditionally been considered Middle and what is traditionally considered Old.

Sagart and Baxter have recently attempted to blend the techniques of historical phonology with those of phonosemantics (Baxter and Sagart 1997), and have since posited reconstructions for 4000 Old Chinese words and 9000 Middle Chinese words (http://lodel.ehess.fr/crlao/document.php?id=1217, 2011).

Details on the various theories of historical Chinese phonology are almost entirely outside the scope of interest of applied language acquisition. Technically, its purpose is to reconstruct sound systems, and in the process, divine how words (as spoken) may have sounded (at least, to a phonemic approximation) hundreds, if not thousands, of years ago. It is hard to see how students would need to know such information. The only possible issue of relevance would be whether a particular character element is of semantic or phonetic import in any given case.

Because Chinese words were of course represented by characters, the field of historical phonology breeds a completely different meaning of the word etymology from that employed in graphical analysis. Thus, whereas a historical phonologist might describe the etymology of 田 tian2 'field' as 田: tian2 <den< *din,[13] *wenzixue* scholars would observe that the character evidently depicts square rice paddies with foot paths between each field. Both descriptions may be perfectly accurate, and there is no need for one etymology to be considered to be superior to the other. They simply describe two very different phenomena. Unfortunately, some scholars seek to employ historical phonological interpretations of word histories as proof that most graphical etymology is little more than a folkloric game.[14] Certainly, from the standpoint of source materials, neither field is without its serious shortcomings, and as a result, all opinions in these fields should be treated with due circumspection. But from a theoretical standpoint, asserting the superiority of one approach over the other is absolutely senseless: they analyze completely separate phenomena.

---

13 This example happens to reflect Baxter's analysis (1992).

14 Unger and DeFrancis are particularly dismissive of graphical etymology.

2.5 Morphology

On account of its relatively isolating nature, Chinese morphology does not typically receive much attention. However, as Li and Thompson (1981), Packard (2000), and others demonstrate, the language does lend itself to morpheme classification. Thus, rather unsurprisingly, Chinese distinguishes between free and bound morphemes, as well as between content and function morphemes. The intersection of these two categories yields four morpheme types: function words, affixes, free roots, and bound roots.

The compounding of morphemes to form words can also be classified into a set number patterns, such as reduplication, affixation, and compounding. However, the precise significance of these processes involved depends on the theoretical framework from which they are viewed. Analyzing the internal structure of words will therefore depend on how one describes the word's components., which can be defined relationally, semantically, syntactically, or based on their form class (Packard 2001, 21 - 36).

Most of these theoretical subtleties will be lost on the typical Chinese language student. The question that matters for students is whether intraword relations are useful to the student. If studies such as those cited by Koda (2004, 71 - 94) are any indication, appreciating the rationale behind a word's internal structure and its lexical meaning can indeed aid in understanding. If 60% of English words are thought to be morphologically transparent (Koda 2004, 77), then in a language like Chinese, where characters clearly play a strong semantic role, the value of understanding the meanings and roles of individual morphemes should prove extremely helpful for lexical acquisition. And since the individual morpheme is almost inevitably represented by the character, then the importance characters as a target of study cannot be underestimated.

2.6 Writing System Taxonomy

Numerous scholars have sought to characterize writing systems as a whole in the context of writing system theory. General comparative approaches can be found in Coulmas (1989), Sampson (1985), as well as DeFrancis (1989). Analyses that focus

primarily on Chinese and/or other East Asian scripts include DeFrancis (1984), Unger (2001), Boltz (1994) and to some extent, Hanas (1997).

If I may be permitted to distill several books into a couple of paragraphs, the first question is whether all writing systems inevitably mirror speech. Sampson describes this dichotomy as one between glottographic systems, which mirror speech, and semasiographic systems, which are based purely on signs devoid of the spoken language. Arguments can be made either way as to the theoretical possibility of semasiographic systems. However, as no cases of semasiographic writing systems are known to exist, the next step comes down to categorizing glottographic writing systems —that is, those that mirror speech. Here, classifications begin to differ markedly depending on the scholar.

Sampson identifies an ensuing dichotomy, this time between logographic systems, wherein the writing units constitute words or morphemes, and phonographic systems. These are divided as shown in the figure below.

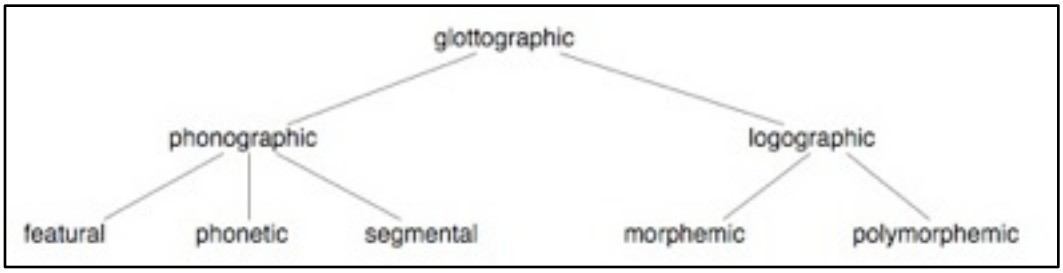


Figure 2.1 Sampson taxonomy of writing systems

By contrast, some scholars only recognize the possibility of phonographic systems (the left branch Figure 2.1) on the assumption that writing systems must mirror speech and therefore convey sounds.

In my opinion, the argument that writing must mirror speech does not necessarily imply that writing explicitly represents the sounds conveyed. Rather, the degree to which sounds are represented can lie anywhere along a continuum. A purely logographic system with no phonological system would inevitably require an unwieldy morphemic inventory that is difficult to learn. On the other side of the spectrum, an ideal phonographic system would need only a few dozen symbols to convey everything that could be expressed in the language. Unger takes exception with Sampson's

typology, wherein languages like Finnish and Spanish lie towards the phonographic end of the continuum, and Chinese, Japanese, and Korean lie towards the logographic end. Instead, he argues that most of these systems all lie fairly close together in a central position along the continuum (Unger 2004, 31 - 32). However, while it is true that in practice, writing systems mirror speech, it does not necessarily follow that the sounds are explicitly represented. In other words, the *original compositional rationale of a character must be distinguished from its explicit current-value representation*.

To see the difference, one need simply compare Hoosain's accounting of meaningful phonetics—those where the phonetic and phonetic compound have the same reading—with the frequently cited claim that by modern times, 97% of all characters are phonetic compounds. Hoosain notes that phonetic-to-character-reading consistency is 26.3% for modern usage characters and only 18.5% if one takes frequency into account[15] (Hoosain 1991, 10). By contrast, DeFrancis asserts that the number of phonetic compounds rose steadily over the centuries from 34% in the oracle bone period to 97% in the 18th century (DeFrancis 1984, 84). The implication is that in present-day Chinese, nearly everything is phonetically represented.

There are two problems with the "97%" argument. First, the 18th century figures are based on the Kangxi Dictionary, which contains tens of thousands of short-lived characters that historically were rarely used. From the standpoint of generative potential, phonetic compounds should expectedly become far more prolific in later times, especially if designed for proper names, where the motivation to employ semantic strategies would be practically nonexistent. But if instead one searches characters actually in use, one finds that the ratio of phonetic compounds is far lower than the 97% typically claimed. And, as the Hoosain statistics indicate, if one is judging phonetics in terms of their transparency, the true number is around one in four.

The second problem is that the argument that nearly every character is a phonetic compound employs a pessimistic assessment of possible semantic implications in character elements. Karlgren himself, who after all brought Chinese historical phonology to the modern era, and should therefore have a healthy dose of skepticism of

---

15 That is, the percentage one would encounter if reading a typical text.

any "youwen" interpretation of characters, admits that many phonetics played a dual role (Karlgren 1923, 17), though he makes no attempt to explicitly label them as such, saying that such roles are "obvious."

Another factor determining the "true" degree of phonetic representation in Chinese characters lies in the manner of analyzing the roles of recurring elements in characters. Some approaches inherently guarantee an over-assessment of the percentage of phonetic compounds. To see why this is so, consider the common case of a polysemous character, one having two meanings associated with it. Assume further that the original character is a basic semantogram (pictograph or ideograph). If, to lessen ambiguity in the writing system, one differentiates the character's two meanings by adding a semantic element, the result will look like a semantic-phonetic compound (SP, where S represents a semantically motivated element, and P, a phonetically motivated element). In fact, provided that the original element was semantically motivated, then addition of a semantic element results in an SS character. But because the original element also carries the same reading as before, it will look structurally identical to a true SP character, wherein the second element provides only phonetic information. In other words, a significant percentage of so-called "phonetic-compounds" actually contain a primary element whose role is to provide semantic information, with phonetic representation being a secondary byproduct of the fact that the original semantic graphic had that reading. Unfortunately, without significant research into the graphical etymology, there is no way to distinguish a phonetically driven SP from a semantically driven SS(P), as compositionally they look identical. And because of the paucity of data such research may lead to frequent dead ends.

There is a fourth reason why the notion of huge percentages of phonetic compounds may be something of an overestimate. It is fairly obvious that even when an element was chosen as a phonetic, of the many possible phonetic candidates containing that reading, there was a strong tendency to choose something that could provide a semantic hint as well. Such elements could be considered "dual-role" elements (D). In such cases, the exactness of phonetic matching might be compromised in favor of something that also provides semantic hints. This process would especially true if a

character were created "from scratch" instead of by adding a determinative to an existing character.

Finally, as Creel rightly points out, employing the approach taken by Boodberg and Boltz, any element whose original pictorial significance is unknown can be easily mislabeled as a phonetic.[16] In short, there is a risk of circular reasoning in historical reconstructions if information about the graphical etymology is insufficient.

By virtue of the fact that the symbols are used to present words on a page, they also represent sounds. But that does not mean that the sounds are represented overtly. Whatever phonetic representation does exist does not typically break down to individual phonemic segments,[17] but at best indicates varying degrees of affinity between the sound of the element used as the phonetic and that of the character in which it is used. Although psycholinguistic studies have shown that intracharacter elements are sometimes used to identify characters, the sounds of characters cannot be clearly identified from the constituent elements any more certainly than their meanings can.

Actually, even the argument that writing must mirror speech is questionable when looking at classical Chinese. While writing system analysts have tended to take this fact as a given, when one views the telegraphic style of older classical Chinese texts, one has to wonder whether a not-quite-spoken approach was taken when putting thoughts to paper.

If one looks at the most obvious distinction between a system like Chinese and more obviously "phonographic" systems like English, it could be argued that Chinese characters have a somewhat closer relationship to meaning than the phonic units used in English or other alphabetic scripts. This is because the graphs represent not just sound but the intersection of a particular sound and meaning. In fact, a given character can

---

16 The argument between Boodberg and Creel transpires in a series of papers in T'oung Pao between 1936 and 1940. DeFrancis characterizes the interaction as having been "won" by Boodberg (DeFrancis 1984, 85-87). However, in my opinion Hanas provides a more accurate interpretation of Creel's arguments, which admittedly were rather poorly expressed (see Hanas 1997).

17 Phonetics representing a single vowel would be obvious exceptions here.

represent several sounds and several meanings as well, and this multivalence is quite a different phenomenon than the simple multivalence of phonetic realization of phonic units used in "imperfect" spelling systems like English.

By contrast—barring cases of single-letter words—letters simply represent sound. Still, one should not overlook the ability to analyze a language like English morphologically. In that case, of course, both sound and meaning are represented, a phenomenon that *does* constitute a universality in full-fledged writing systems.

In examining the above taxonomies, I neglected to point out that many authors look at writing systems from a developmental perspective. Boltz, DeFrancis, and scholars of that mind set take the primary stages of development as a transition from pictorial representation to rebus representation followed by disambiguation of homophonous or polysemous graphs through the employment of semantic or phonetic determinatives. For Chinese, Woon (1987) identifies three driving principles: phonetic loans, semantic representation, and semantic extension. If students are to be taught something about why characters mean what they do, their compositional rationale, and the evolution of their meanings, then indeed, some of this information could be useful if only presented in the broadest possible strokes. Understanding the notion that meanings can be extended, or that characters can be borrowed for their sound, is naturally helpful in explaining why there is so much semantic and phonetic multivalence in the writing system.

2.7 Differentiating Character Usage and Compositional Rationale

Based on research I conducted into the formation and compositional rationale of *kokuji* (Chinese-like characters of Japanese origin), I proposed a tentative typology for compositional rationale and character use (Child 2004). What follows is a new presentation of that typology with slight modifications to reflect the more general case of Chinese-origin characters.

This model (see outline on following page) draws a clear distinction between the motivation for composing a character, the manner in which it came into being, the technique or rationale underlying composition, the types of elements employed, the processes of composition, and the way the character is employed.

Motivations for Composing a New Character

- to represent a new word in the language
- to represent an existing word in the language
- to represent a proper name or title
- to provide a shortcut to writing or transcription
- none—meaning that creation was accidental

How the Character Came into Being

- by design
- by accident

Rationales for Character Composition[17]

- to provide phonetic clues to its pronunciation
- to provide semantic clues to its meaning
- to provide both phonetic and semantic clues simultaneously through one element
- none—meaning the character is an apparently arbitrary sign

Types of Elements

- phonetic indicators (reading-suggestive)
- semantic indicators (meaning-connotative)
- dual indicators (suggesting both meaning and reading simultaneously)
- abstract symbols with no appreciable semantic or phonetic value

Processes of Composition

- use of a single element corresponding to one of the above types
- compounding of any of the above element types
- reduplication of a character or element
- deletion of one or more elements from an existing character
- addition of one or more elements to an existing character
- substitution of one element for another
- creation of abstract symbols
- stylistic alteration employing elemental correlates between one style of writing and another
- stylistic alteration combined with any of the above forms of modification of constituent elemental composition

How the Character Is or Was Used

- to represent a proper name or part of a proper name
- to represent a word or morpheme in the general lexicon
- to represent the sound portion of a word without providing semantic connotation
- to represents the sound portion of a word while also hinting at the meaning of the word

If one looks closely at the list, one will quickly see that what might easily be construed as a simple phenomenon of graphical etymological development actually has

many tiers, with marked differences, say, between character use (a relatively simple affair), processes of composition (with at least nine subtypes), and compositional rationale (also a simple affair). This last statement, and in particular, my assertion that character use is a simple affair, may strike some as odd given that any given character can be associated with a wide variety of meanings and morphemes. What I mean is that the principle of character use is simple: in every single case, a character represents a morpheme (the combination of sound and meaning), or a sound devoid of meaning (as in some proper names or foreign words), or a sound whose meaning is hinted at (as in some clever renditions of foreign words). The fact that any given character may be applied to a variety of meanings does not belie this truth. Besides, as Qiu points out, many of the meanings and readings of any given character are not employed synchronically, but rather simply result from the tendency of dictionaries to list every and all meanings and readings of a character in its long history. So at any given point in time, character usage is not as complex as some dictionaries would have one believe (Qiu 2000, 369).

In looking at the above characterizations and levels of analysis, one must also be aware of false parallels between the levels. Other than between element types and element usage, the various levels of analysis do not have one-to-one correspondences, despite appearances to the contrary. In particular, I would like to draw attention to the difference between the last category, how a character is used, and compositional rationale.

Compositional rationale can be none, phonetic, semantic, or both phonetic and semantic simultaneously. Essentially, this means there can be semantic, phonetic, or phonosemantic characters, with the odd purely abstract composition tossed in.[18] As is clear in the table, the types of character elements follow an exact parallel. And so, one could hardly fault people for expecting character use to adhere to these four categories.

As it turns out, character usage can really only play one of two roles. Namely, a character can represent a morpheme (a marriage of sound and meaning), or it can provide purely a phonetic role in the larger context of a polysyllabic word (as is typical

18 The only such wholly abstract characters I have found were for little-known and no-longer-used characters for dollar or other units of measure.

Table 2.4 False parallels at different levels of the orthography

Compositional Rationale	Character Elements	Character Use
provide semantic clues to original meaning	semantic (meaning-suggestive)	N/A
provide phonetic clues to pronunciation	phonetic (reading-suggestive)	represent sound (only) in a (usually foreign/proper) word
provide both semantic and phonetic clues simultaneously	dual role indicators (suggesting both reading and meaning simultaneously)	represent morpheme (marriage of sound and meaning) OR represent sound and hint at meaning (as in some foreign words)
none (purely abstract composition)	purely abstract	N/A

in foreign words). But there is no guaranteed or necessary correlation between the semantic and phonetic clues provided in compositional rationale with the morpheme (sound and meaning) of use. This distinction is crucial.

Another thing to notice is that Qiu's loangraph category is simply not possible as a "character type" in the same sense that semantographs and phonograms are character types. Borrowing a character does not result in the composition of a new graph; it is simply a reassignment of the link between a graph and one of the morphemes with which it is associated. Therefore, Qiu's "three principles" in fact incorrectly mix levels of analysis. His first two categories (semantographs and phonograms) refer to compositional rationale, while the last (loangraph) is a category of usage. To put it another way, a character may be borrowed to represent a totally unrelated morpheme, but it cannot be composed for the purpose of loaning its form. And once characters are composed, there has never been any stopping people from using them in a variety of unforeseen ways, none of which relate to the mnemonic values their constituent elements.

A corollary to this realization is that there is usually considerable distance between the process of composition and the strategies of use. Therefore, when scholars employ a line of Chinese text to "prove" that there is no correlation between the writing system and what the characters depict or mean on their own, they are missing the point.

There is not supposed to be a strong connection between character use and compositional rationale in the first place. In fact, depending on the character in question, there may well be no connection at all. The assembly of semantic and phonetic elements in a character can at most hint to the character's original meaning and reading, and says nothing about future use. Whereas the role of character composition is primarily mnemonic—that is, to avoid having to remember thousands of wholly abstract elements—usage can result from a variety of complex processes, including semantic extension and phonetic borrowing, not to mention morphological relations between characters

What does this mean for the student? It means that, barring accidental creation and wholly abstract characters, nearly every character must have some compositional rationale whether present-day scholars can deduce it or not. A character's compositional rationale may be helpful from a mnemonic standpoint, but learning the character's use will require a separate effort. One of the questions explored later in this dissertation is whether the effort of learning the compositional rationale is worthwhile.

Finally, careful study of phonetics may be extremely useful. The reason is not that the phonetics are reliable predictors of character pronunciation, but because they serve as the greatest common denominator in most Chinese characters. And, in perhaps 30% - 35% of all such characters, there is likely to be some semantic contribution. From the standpoint of character discrimination, by far the most helpful component is likely to be the so-called phonetic, or non-radical, component.

* * * * *

Having examined our proverbial elephant from various perspectives, the obvious question becomes: What information is of use to the student? What can and should be taught? And what aspects of information deserve testing to see if they benefit the student in any way?

Immediately, historical phonology and phonosemantics can be ruled out for the reasons cited earlier. Except for those aspiring to study ancient texts, and in particular, the rhyming schemes of old poetry, the readings of words centuries ago is unlikely to help the student of Chinese today.

Writing system analysis, and particularly, a quick overview of the development of the Chinese writing system and basic principles of character composition, seem potentially helpful. Such information could, in principle, be distilled to a few short pages of introductory material.

Finally, although recurring character components are subject to widely differing interpretations, analysis of semantic and phonetic elements could potentially help learners develop intracharacter recognition and discrimination skills. Morphological analysis might, for parallel reasons, prove beneficial in encouraging intraword analysis, which should help in word meaning retention, reading identification, and the ability to guess the meanings and readings of unfamiliar words.[19] These suppositions are tested in later chapters.

---

19 Psycholinguistic studies have been cited to suggest that native and non-native speakers use different strategies in activating the sounds and meanings of Chinese characters. Such differences have been overplayed, and they also fail to note one very critical observation: the psycholinguistic realization of phonological or semantic activation is very likely to depend on what and how the person learned about characters. Thus, people who know nothing about a character's makeup will not "see" structure that exists, whereas those who have been instructed in graphemic makeup will see characters differently. Chapter 5 discusses this issue at length.

Chapter 3
Mandatory and Potential
Facets of Lexical Acquisition

Chapter 1 identified the lexicon as a primary issue impeding mastery of Chinese orthography, together with minimum targets for character and word acquisition. Chapter 2 explored analytical perspectives on the Chinese writing system itself to determine which perspectives are pedagogically relevant. This chapter takes a closer look at the data that is or can be associated with the lexicon, and what approaches can be taken to utilizing that data.

Reading entails a host of processes taking place at successively higher levels, including the recursive construction of meaning from words to phrases to sentences and on to higher levels of discourse. As this chapter reveals, lexical acquisition invokes a suite of subcomponents, including exposure, familiarity, discrimination, association, connection, and retention. These subcomponents operate at multiple levels—specifically, at the recurring element, character, and word levels. The subcomponents explored in this chapter do not constitute activities that would be normally attended to by a native speaker during the act of reading, but rather are potential facets of the encoding process. In other words, they speak directly to the nature of the lexical data that can or could be presented to the learner.

The facets of lexical acquisition discussed in this chapter are defined as follows:

- **exposure** is the act of being exposed to some form of information
- **familiarity** is the sense of having been previously exposed to something
- **discrimination** is the ability to distinguish patterns that might otherwise appear similar
- **association** is the act of tying information to a cue; by definition it is a vector relation
- **connection** is the act of tying separate portions of an entity together in a meaningful way
- **retention** is the act of storing information

3.1 Exposure

It goes without saying, but the first element of lexical data acquisition is exposure. If there is no exposure, there is no learning of any kind. One must be exposed to the data. The act of exposure itself is of little interest, but the manner of exposure is extremely interesting when considering the three-tiered structure of the Chinese orthography.

Elements compose characters, and characters compose words. When one is reading text, it might appear that one is exposed to all three tiers of the orthography simultaneously. In fact, depending on the nature of the exposure, meaningful exposure usually does *not* occur at all three levels. More often than not, that exposure to associated data only occurs at the word level.

How is this possible? Quite simple. If one does not perceive a type of data, one cannot very well be said to have been exposed to it. This is attested to by the fact that beyond, perhaps, first and second year textbooks, most books provide only word lists, not character or character element lists. If the only data provided is at the highest level, that of the word, then from the standpoint of meaningful exposure, little data exposure is likely to take place at the element and character levels. The effects of different types of exposure differ very much on the activity involved and the manner of presentation, as shown in Table 3.1.

Table 3.1 Relationship between exposure activities and level of data reinforcement

Activity / Exposed Level	Element	Character	Word
Reading graph etymologies	Yes	Yes	No
Reviewing standard textbook character data	No	Yes	Maybe
Standard textbook word list	No	No	Yes
Reading	No	No	Yes

Some instructors have expressed a concern that an excessive "focus on characters" may be unhelpful. This concern may reflect the fact that reading graph etymologies does indeed do nothing for word learning *directly*. To get through a text, one does not read Chinese character elements, or even Chinese characters; one reads words. But does that mean one should not learn something about how or why characters are put together? This question is addressed at length in Chapter 8.

For now, it is enough to note that exposure to lexical data takes place in context or out of context. And the grouping of lexical data can be "organic," meaning that it arises naturally based on some text that contains the lexemes in question, or "inorganic" meaning that it has been put together using some artificial rationale.[20]

3.2 Familiarity

The term *familiarity* has been used previously in literature on lexical acquisition, but it is rarely defined (Melka 1997, 85). Melka suggests that *familiarity* could be viewed as an extent of knowledge that runs the gamut from uncertain levels of reception to comprehensive knowledge sufficient to permit accurate and appropriate production in speech or writing.

In this chapter, *familiarity* is defined a little differently, as it is taken to mean the most basic, if uncertain, sense of having seen something before. In and of itself, familiarity does not constitute sufficient information for actual reading, but it is the most essential aspect of recognition of something previously seen or learned. As it turns out, this definition of familiarity happens to have a basis in cognitive neuropsychology, as it has been discovered that recognition that something is vaguely familiar operates via a different mechanism than that when recalling information about something (Smith and Kosslyn 2009, 218-219; Yonelinas et al., 3002-3008).

Both definitions of familiarity are helpful from the standpoint of lexical acquisition, though it is good to realize that they refer to rather different phenomena. The former is simply a sense of having seen something before. The later describes the level of lexical data associated with a given target lexeme.

---

20 An experiment of character grouping by graphical similarity would be an example of an inorganic presentation (e.g. Child 2004).

3.3 Association

Association is defined here as the association of a cue with a target. In the case of the Chinese lexicon, it refers to the possible interrelations between graph, sound, and meaning at the three levels of the orthography: the element, the character, and the word. All three—graph, sound, and meaning—can be found at all three levels: recurring element, character, word, though arguably many recurring elements do not officially have a sound, and regardless of their status thousands of years ago, many elements do not have an unambiguous meaning either. These relations are expressed in Table 3.2.

Table 3.2 Relationship between orthographic entities and associated data

	Element	**Character**	**Word**
Graph	Yes	Yes	Yes
Sound	Maybe	Yes	Yes
Meaning	Maybe	Yes	Yes

Semantic associations can be theoretically be found at all three levels, and characters and words can have sound and meaning associations. Recurring elements may or may not have an associated sound or meaning.

Typically, native speakers of Chinese know very little about the original meanings of the many elements used to compose characters, and at best may be aware of the meanings of a few dozen common radicals. Likewise, some studies suggest that there are many characters for which even native speakers do not have a clear semantic association.[21]

Given the previous section on exposure, the significance of the relationships becomes a little clearer. Any particular activity will trigger different associations depending on the direction between cue (i.e., input modality) and target invoked during the exposing activity. For example, writing requires one to think of a meaning and then produce a character or word. By contrast, reading involves presenting graphs and will

21 In one such case, Chinese native speakers struggled to come up with a meaning for the all so familiar character 巴 ba.

trigger meaning and sound responses, with perhaps some reinforcement of the graph. The vector-like nature of these associations is explored in the chapter on memory.

3.4 Discrimination

Discrimination is defined here as the ability to clearly distinguish target entities. As the reader might have guessed, in the context of the orthography, this activity too can take place at three levels: the recurring element, the character, and the word. In the literature, a positive correlation has been found between learners' performance on Chinese vocabulary production achievement tests and the study of character components (McEwen 2006, 54). Even so, one finds occasional concern that a focus on graphemes would be counterproductive. Despite such concerns, one thing is clear: regardless of the degree to which a reader may be aware of the intracharacter composition of characters, by definition, to recognize Chinese words without error requires the ability to clearly discriminate the graphs.[22] And this, in turn, means the ability to clearly discriminate constituent elements. If this were not the case, then one would be able to substitute a host of similar elements and still have reading proceed as normal. This is clearly not the case.

There is no real need to reproduce the table like that shown under the section on familiarity. The simple fact is that intracharacter and intraword awareness will largely reflect the training and learning process that the subject underwent. Thus, it is quite conceivable to have a foreign student develop a highly acute ability to discriminate lookalike elements within characters, and to be sensitive to the intraword connections found in most compounds. Needless to say, without training in that regard, these skills will be wholly lacking. However, regardless of the training, on some level, whether conscious or unconscious, the ability to discriminate elements within characters, and characters within words, is an absolute prerequisite, as failure to discriminate at these levels inevitably results in recognition failure.

---

22 Using top-down processes, students trying to recall or recognize a compound can fudge the process of recognizing a first character while having a vague sense of what a second character should look like. This will not always work, and will eventually lead to error as more difficult materials are encountered.

3.5 Connection

In the context of this dissertation, connection refers to the act of drawing connections between constituent elements of an entity. In principle, the notion of connection naturally applies to both the intracharacter and intraword levels. Someone who has studied graphical etymology may see the connections between or among the constituent elements of a given character. Likewise, someone who for whatever reason pays attention to intraword relations will likewise form a sense of the morphological relations within words.

By definition, connection implies prior exposure to and familiarity with an entity's constituent elements, together with the existence of meaningful associations for those elements. At the character level, this means knowledge about the elements used in the character. At the word level, it means knowledge of the meanings and readings of constituent characters within the word.

Table 3.3 The role of connections at the intracharacter and intraword levels

Intracharacter Connections	Intraword Connections
What are the contributions of the elements to the character a whole, and is there an interrelation between them?	What are the morphological roles of the characters? What sounds do the characters contribute to the word?

Course curricula typically give little attention to intraword and intracharacter connections. Therefore, it is not uncommon for students, when faced with word lists, to learn words as whole units without analyzing the connection of constituent characters. Needless to say, most students probably know even less about intracharacter elements.

Which is not to say that there is no interest in such a thing. In my study on whether character compositional rationale and character grouping assisted with retention (Child 2006), I conducted a follow-up survey to identify each subject's study habits and reaction to the materials. As it turned out, nine out of ten students were familiar with a book entitled *Genealogy of Chinese Characters*. This book makes no pretense of offering serious explanations about character elements. But it does give

students *something*, and it seems they appreciate that ability to identify character components as something other than a random collection of strokes.

Perhaps the most obvious observation to make about drawing connections during lexical acquisition is the fact that connections made during lexical encoding at the two possible levels—intracharacter and intraword—will depend on the information available to the student and on the student's own tendency to either utilize or ignore that information.

3.6 Retention

The last subcomponent, retention, is defined here in a perfectly natural and expected way—namely, the act of retaining lexical data. In principle, retention can apply at three levels in a language like Chinese: the intracharacter level, the character level, and the word level. Given that it is a well-known phenomenon for students to cram information for an exam and to then promptly forget it some days, weeks, or at best months later, and given that to be able to read any given text the student will need to know some very large number of words, anything that can be done to enhance retention would be extremely helpful.[23]

3.7 Interrelations Among the Facets of Lexical Acquisition

The above-described facets of lexical acquisition interact in curious ways with the three different levels of the orthography in one respect: some subcomponents are mandatory, and some are strictly optional. Let us consider some examples.

One cannot read fluently without retaining meanings; else, too much time is spend in a dictionary, and the end result is an effortful decoding of the material at hand. Associations of meanings and sounds with words is also required. At the character level —unless the character itself constitutes a word—meaning and reading are more optional: one can recognize a word without being directly cognizant of individual meanings and sounds of the constituent characters. In fact, because the ultimate goal is word recognition, some scholars actually believe it is a bad thing to study below that

23 Chapters 6 and 7 explore these issues in detail.

level. However, failure to have solid associations at the character level leaves one helpless when encountering new words that contain familiar characters. One may have pronounced the character countless times in the context of another word, but if the cue-target association is nonexistent, then that sound will not be activated. And even if the morphological connections between the constituent characters is totally transparent, if one does not have semantic associations with the individual characters, again, the reader will have no recourse but to look the word up.

By contrast, on some level—whether conscious or unconscious—discrimination is mandatory at all three levels: the element, the character, and the word. This is because pattern recognition failure at any one level will inevitably percolate higher up the orthographic levels. Technically, it is possible for the mind to "fudge", meaning that students sometimes get away with not being terribly clear on the details of a character's parts and structure. Such fudging may work at the lower levels, where the overall shape of the character may be unique enough to distinguish it from anything else in his or her knowledge domain. But as the breadth and authenticity of material increases, the likelihood of encountering another character or word with that general contour grows, and fudging will inevitably lead to error. For example, a word that overall has a fairly unique shape may be distinctive enough to not require perfect appreciation of the elements within the constituent characters for certain levels of reading. However, if discrimination is lacking and one later encounters something that seems "the same," then a recognition error will occur.

Most curiously, of all the subcomponents of lexical acquisition, the least mandatory is connections. Connections are not strictly mandatory at any level. A student can, at least in principle, learn character shapes as entire entities, or associate meanings with entire words rather than the constituent character-morphemes. Whether they are helpful for aiding retention is a separate question addressed later. Figure 3.4 shows the rather haphazard relations among the subcomponents and the levels of the orthography.

In a nutshell, nearly all acquisition activities are optional at the level of recurring elements. At the character level, strictly speaking all activities are optional. This may come as a surprise to some readers. But the fact of the matter is that, unless the character itself represents a monosyllabic word (in which case one should look at the last column),

all activities are optional. One can get away with recognizing the lexicon logographically, without any conscious awareness of the sub-word elements. Finally, at the word level, nearly all activities are mandatory, except for connections, which are left to the reader's discretion.[24]

Table 3.4 Mandatory (M) and optional (O) facets of lexical acquisition at different orthographic levels

	Elements	Characters	Words
Exposure	O	O/M (if word)	M
Familiarity	O	O/M (if word)	M
Association	O	O/M (if word)	M
Discrimination	M	O/M (if word)	M
Connection	O	O	O
Retention	O	O/M (if word)	M

3.8 Propositions for Enhancing Lexical Acquisition

Interestingly, this tendency to increase optionality as one goes down the hierarchy from word to element is strongly reflected by the emphasis implicit in university and other Chinese language curricula, where students must learn words all the time, the study of characters is largely relegated to the early stages, and the study of elements is left entirely to the student's discretion. The reasoning is somewhat understandable given the dilemma described in Chapter 1, wherein teachers want to reach as high a vocabulary level as soon as possible in a very limited amount of time. Many instructors simply feel that the few resources that exist must go directly to the aim of learning words, the minimal unit required for reading passages. Some teachers also protest that knowledge of individual characters will not in and of itself enable a learner to read. And in a sense, they are absolutely correct. During reading, one’s attention should be at a higher level.

---

24 Note that element connects result in a character, character connections result in a word, and word connections result in complex lexemes that are composed of more than word (e.g. State Department is a single entity, despite the fact that orthographically it contains two words).

That said, I would like to propose the possibility that the following relations exist at different levels of the orthography:

(1) U α 1/VD
(2) VD α 1/VR
(3) VR α IWA
(4) IWA α ICA

where U is understanding, VD is vocabulary density, VR is vocabulary retention, IWA is intraword awareness, and ICA is intracharacter awareness. In other words—working backwards from (4) to (1) above—if knowledge about characters can aid in distinguishing lookalikes, remembering their meanings and readings, and overall remembering them; and if knowing characters better helps one to remember word meanings and readings, while enabling one to decipher unfamiliar words with morphological transparency, then the resulting enhanced retention of words should diminish vocabulary density and thereby increase understanding.

In essence, the above four propositions constitute a hypothesis that could very well be tested. Briefly stated, if learning vocabulary is principle bottleneck to moving beyond the intermediate level, then learning to recognize intraword information, and providing the knowledge and skills to be aware of intraword information may well help with vocabulary acquisition. Further, in light of the additional layer in the Chinese orthography that does not exist in alphabetic writing systems, knowledge and training in recognizing intracharacter information may also improve acquisition and retention of characters.

* * * * *

Facets of lexical acquisition largely concern the data, the raw material that the learner works with. However, this is only one side of the interaction between reader and material. The other is the thought processes and cognitive skills that the reader brings to bear on the material. The next two chapters look at these aspects from different perspectives.

Chapter 4
Perspectives from Reading Theory

If one explores the literature on reading generally and how to facilitate reading in a foreign language in particular, one will encounter a set of tenets not usually discussed under the same breath, but which ought to be viewed as part of a general philosophy of reading. This chapter looks at these perspectives with an aim to assessing whether they are actually helpful in the context of Chinese lexical acquisition.

4.1 Prescriptions for Current Day Reading Instruction

Consider a typical "modern" approach to reading theory. A TESL website by the British Council explains that reading theory has progressed through three stages, which the website author terms "traditional," "cognitive," and "metacognitive" (Vaezi 2006). By *traditional* approach, one is meant to imagine a bottom-up undertaking wherein the reader puts linguistic units together. By *cognitive* approach, Vaezi refers to the top-down theory of reading, wherein the reader activates schemata, anticipates what comes next, and engages in contextual guessing. By *metacognitive* approach, one is to understand that the reader is self-aware of the reading process as he reads while engaging in a blend of top-down and bottom-up processing.

Based on the "understanding" that the modern metacognitive view is now widely recognized as a truer description of reading, the British Council site provides a number of instructional prescriptions. These prescriptions are divided into *prereading*, *during-reading*, and *post-reading* activities. Prereading activities include such things as explaining key vocabulary, deciding what style of reading will be adopted (intensive vs. extensive; skimming vs. scanning), and examining the text's format to see where useful information is likely to be. During-reading activities includes what technically appears to be a hodgepodge of strategies, including: predicting what comes next, activating schemata and prior knowledge, skipping unimportant sections, rereading portions that give difficulty, employing contextual guessing, reading groups of words together ("chunking"), breaking words apart to understand unfamiliar words ("decomposing"), taking breaks to absorb the material, sub-vocally recapitulating the material, and

checking one's understanding as the reading progresses. Finally, various post-reading activities are described. These include the kind of follow-up activities one often sees in classrooms, including answering questions, writing about the text, and others. They do not relate to actual reading, however, and so I will not go into detail here.

This website is interesting not for its depth but for its typicality. Of special note is the fact that in this so-called "metacognitive" approach—with the exception of decomposing, chunking, and vocabulary preparation—nearly everything consists of what would be termed a top-down skill. They all involve thinking about the nature of the text, or about one's goals and metacognitive strategies for tackling the material. To understand this heavy bias toward top-down skills, one needs to examine the historical context of the dichotomy between top-down and bottom-up processing.

4.2 A Closer Look at the Bottom-Up / Top-Down Dichotomy

One of the great shrines of reading theory over the past few decades has been the notion that reading takes place in two essential manners: from the bottom up and from the top down. Bottom-up processing is described as a process of scanning in small units of text and constructing meaning from them. Top-down processing, by contrast, is described as the process of taking the reader's knowledge and mapping that understanding onto the text, such that the text is considered as merely a guide to the construction of meaning. Until only recently, many researchers have argued that top-down processing is a significant factor in fluent reading, and much of the literature on pedagogy was informed by this opinion. Thus, one finds well-read SLA experts alerting their readers to the dangers of "promoting decoding," as if it were some kind of retarding influence on reading development. Exhortations to guard against an excess attention on kanji in the instruction process are similarly invoked in other sources.

On careful reflection, however, it should be fairly easy to demonstrate that top-down processing is in no way more important than bottom-up processing, nor do the major assumptions made by top-down theorists hold water. Some of the most important assumptions of top-down theorists are that the reader can (and should!) guess what comes next, and that the reader does not need to see all the words to comprehend the text, or that many words can be guessed from context.

Let us consider each of these assumptions in turn. If the first were true, then we should be able to cut passages short and have the reader guess the rest. Obviously, it is possible in some well-constructed texts to anticipate some information. But it is equally easy to prove that there are limits to this ability. More importantly, if asked to "fill in" the rest of a passage, a considerable amount of time would be needed for the reader to do this compared to the amount of time it would take to actually read the completion of the text. SAT and GRE reading passage comprehension sections actually test this kind of reasoning, but the answers are hardly obvious to the person being tested.

Now consider the second assumption, that the reader does not need to see or know all the words, and can fill in for what is not seen by employing top-down reasoning. There are two ways to interpret this. The first is that language learners can guess the meaning of unfamiliar words from the context. The second is that skilled readers need only scan salient words and "fill in the rest", somehow glossing over details provided by less critical words. With respect to the first assumption, it turns out that studies show just the opposite: that lack of skill at the decoding level forces language learners to compensate by guessing, both slowing down speed and impairing comprehension (Stanovich 1987). In other words, bottom-up processing or "decoding" is effortful for poor readers and actually very good in skilled readers. Now consider the second assumption—the notion that good readers only attend to salient words. While it has been shown that depending on context, some words may be attended to for a shorter fixation, nonetheless, even without fancy laboratory equipment, it is easily shown that the "less salient" words still require some attention, enough to absorb their role in the text.[25] Otherwise, the following passage would be just as easy to read as the original text from which it was taken.

> Finally, provide .... classification system ... avoid .... pitfalls ...
> approaches ... so far. ... get there... however.... review .....literature ....
> writing system .... nature ... characters .... Chinese .... particular.

---

25 I am roughly taking salience as a function of word length and frequency, such that short, highly familiar words are expected to require less processing during lexical access.

True, a skilled reader can construct a plausible interpretation of the above word sequence, but he or she would clearly have to exert considerable time in doing so, and the accuracy of the results would be uncertain. The fact that some contexts permit less attention on certain words does not mean they need no attention at all. And so—whether one posits a hypothetical native zoning in on salient words, or a typical foreign student only familiar with more common words and guessing the meaning of what is unfamiliar—any notion that meaning could be easily abstracted via guesswork is clearly wrong. Quite the contrary, such guesswork would prove slow and mentally exhausting. The assumption that many words escape attention falls quite flat, on closer inspection, despite what interpreters of saccadic movement studies might wish to suggest.[26]

Fortunately, more recent studies have shed a more critical eye on the metaphors of bottom-up and top-down processing. In his excellent summary of the relationship between the two, Amos Paran[27] argues that less skillful readers have no choice but to focus on the lower-level decoding processes precisely because they have not achieved automaticity in the lower-level arena. Moreover, their frequent use of top-level skills is largely a compensatory strategy, not one to be emulated. Along similar lines, in advocating an "interactive-compensatory model" for reading fluency, Stanovich argues that "general comprehension strategies and rapid context-free word recognition appear to be the processes that most clearly distinguish good from poor readers" (Stanovich 1980, 32).

It is perhaps understandable why teachers may not find focusing on the rudimentary skills an exciting proposition. After all, drills on characters and vocabulary can hardly seem inspiring instructional material. But if these recent insights on the

---

26 Eye tracking studies distinguish between saccades (skips) and fixations (stopping points), with studies showing that only a limited area is seen clearly to the right and left of the point of central fixation. Contrary to what one might expect, superior music score readers employ more fixations with smaller saccades, suggesting that they actually do their best to see as much as possible clearly. If the same is true of good readers, then the argument of not needing "unimportant" words is further diminished. For a landmark article discussing saccadic suppression see Matin (1974).

27 Originally found at http://www.rdg.ac.uk/AcaDepts/ll/app_ling/buptdown.htm though now removed from that website.

failure of top-down theory to account for the facts are true, it suggests that, now more than ever before, greater attention is needed on finding ways to make the dreary task of mastering lower levels of the Chinese orthography more palatable, if not enjoyable. In sum, the metaphors of top-down and bottom-up processing do seem to describe viable strategies for abstracting meaning from text. However, while top-down processing is useful in some circumstances, fluid bottom-up processing is the foundation upon which competent reading is based.[28] Chapter 5 will revisit the issue of types of processing, and will present arguments as to why this would be the case.

4.3 Learning Words in Context

Another common prescription with respect to lexical acquisition is the notion that words should always be learned in context. Intuitively, there is a certain appeal to this idea, and no doubt many students can, with some dread, envision long lists of vocabulary that they had to learn at some point in their lives, whether in their own language or in a foreign language. The argument is straightforward enough: without seeing how words are used, knowing their meaning in a declarative-memory sense does not ensure that they will be used correctly. For some words, connotation and referent are key components of the word's use, and to not experience such usage in context makes the learner prone to error. So far, nothing in this argument seems amiss.

The problem is that the notion of learning words in context, when applied to foreign language acquisition, often accompanies a presumed corollary, namely the notion of "guessing words from context." The above-cited British Council website is hardly outside the norm, as numerous well-known sources strongly advocate this approach.[29] But how viable is this strategy, and does it really help students learn words?

Noting that many of today's most influential scholars were strongly influenced by the "naturalistic" approaches of the 1970s and 1980s, where the emphasis was on "implicit" learning of vocabulary without glossing words in text, Sokmen points out that this naturalistic approach has shortcomings. For one, there is no guarantee that the

---

28 For a review of the extent to which decoding determines reading skill in the third and fourth grades, see Hirsch (2003).

29 For a strikingly similar list, see Teaching Language in Context (Omaggio Hadley 2001, 206).

reader's guess will be correct; in fact, studies have shown that the reader is actually unlikely to guess the meaning correctly (Sokmen 1997, 238). Citing another study by Nation and Liu, Nation and Waring state that the percentage of words one must know to guess correctly from context is above 95% (Nation and Waring 1997, 11). However, their estimate (for English) of needing only 3000 words to accomplish this is clearly off. If one is interested in a wide variety of subjects, the numbers tend to be much higher, with estimates as high as 15,000 for 97.8% coverage in English. One would expect at least that number for Chinese.

There is an even bigger problem with trying to "get by" on a low vocabulary level in Chinese. Unless the learner has a lot of training in character components and the phonological reflexes of various phonetic elements—and that is obviously not the case! —unknown words guessed from context will often not be pronounceable to the learner. Whereas the pronunciation of unknown words in an alphabetic script can, with varying grades of success, be worked out by piecing together the sounds of the represented phonemes, unless a Chinese word happens to contain known characters (it often does not), the guessed word will have to be treated using some "other language" while reading text containing that word.

Another significant problem with contextual guessing is that even if the student guesses correctly, there is no guarantee that the word will be retained, even when the material has been altered to make the context "rich" enough to facilitate correct guessing.[30] As it turns out, explicit vocabulary training—the use of word lists or computer-based vocabulary training—has been shown to facilitate reading proficiency (Sokmen 1997, 239). Learning the most frequently encountered words is a rational first step, but obviously at some point, difficult words need to be learned as well, and relying on context may not be the best strategy after all.

4.4 Alternatives to Intensive Reading

Another reaction against reading approaches alleged to derive from the so-called traditional approach to language learning has resulted in a rebellion of sorts against overemphasis on "intensive" reading—reading where one closely examines a limited text

30 The reason such guessed words might fail to be retained should become clear in Chapter 6.

for detailed understanding and analysis. Since intensive reading is what characterizes almost all textbooks for Chinese learners, there is hardly need to define the concept more closely. This is the reading of passages accompanied with heavy glosses, if not in the margins, then in an accompanying vocabulary list. Typically, relative to the student's current level, the lexical density tends to be quite high, with many new words introduced in each lesson.

As mentioned in 4.1, alternative reading styles including *scanning* for specific information, *skimming* for the gist, and *extensive reading*, the reading of materials that are relatively easy and for pleasure (Day and Bamford 1998). In a twist on Krashen's famous comprehensible input hypothesis, wherein suitable input should be *i plus 1*, with *i* representing the student's current level, and *1* indicating some unspecified incremental amount of difficulty above that level, Day and Bamford describe extensive reading as *i minus 1*, in other words, material where lexical density and grammatical complexity is well within the student's reach (Day and Bamford 1998, 17).

Bamford and Day report that extensive reading has shown significant benefits in terms of students' attitude toward reading (Day and Bamford 1998, 35-36). This is not surprising, given that this approach has numerous favorable goals such as enabling students to read subject matter that interests them and reading as much as possible for pleasure. One would also expect significant gains in reading speed for materials at that level (*i minus 1*), since extensive reading incorporates the aim of reading large amounts of material at a fast speed.

But does extensive reading help with vocabulary acquisition? Here, it turns out that the results are not so promising, as even Day and Bamford admit. Studies have actually shown that while extensive reading may help to consolidate vocabulary, the gains for new words are mixed (Day and Bamford 1998, 36). This should come as no surprise. Since by definition the material consists of *i minus 1* level text, neither the vocabulary level nor grammatical complexity is meant to greatly challenge the student. Most of the words are supposed to be known.

4.5 Automaticity Theory

Numerous reading theorists assert that for fluent reading to transpire, words must be recognized rapidly, and their meanings retrieved immediately. This notion is generally referred to as *automaticity.* Automaticity is less a model of learning than a description of presumed fact: poor readers struggle with the various subskills of reading, and fluent readers perform lower level tasks automatically. In particular, researchers of reading have shown that a major difference between poor and good readers is that the latter have almost full command of lower level skills such as word recognition. Since it is precisely this area where Chinese learners are apt to struggle the most, the literature in this field to-date deserves a more careful review.

4.5.1 Basic Definition

Stanovich (1990) characterizes automaticity as having the following qualities: 1) the skill is performed extremely quickly, almost instantaneously; 2) the skill requires no conscious attention; and 3) the skill's performance, upon reception of the stimulus, is obligatory, meaning that the performer "can't help" performing the skill. Logan identifies a slightly different set of criteria: 1) speed, 2) effortlessness, 3) autonomy,[31] and 4) degree of consciousness (Logan 1997). Psychologists have further asserted that automaticity of lower-level skills is a prerequisite to skillful performance in complex tasks. In the case of reading, this means that word recognition must be automatic. Yoshimura points out that word recognition is not the only automatized subskill found in reading. For Japanese, Yoshimura identifies three components of text subject to potential automatization: word recognition, phonemic/graphemic decoding, and syntactic feature recognition (Yoshimura 2000, 5 - 6). One can imagine that the set is similar—with perhaps slight variations of emphasis—for other languages regardless of typology, as automaticity in grammar should enable the student reader to instantly recognize word classes, verb endings, and the like.

---

31 As used here, autonomy is the notion that a process runs from beginning to end without intention. It is thus akin to Stanovich's "obligatory" criterion.

4.5.2 Models of automaticity

There are essentially three primary schools of thought with respect to how automaticity is achieved. The first is *strength theory*, whereby repetition of attending to a task is thought to strengthen a connection between stimulus and response (Logan 1997, 128). A second theory is characterized by different terminology depending on the proponent, but the general idea is that a task initially perceived as a sequence of individual components is gradually, through practice, "chunked" into fewer, bigger tasks. A typical example of this process would be how learners initially view a word as separate letters, then as a combination of syllables, and finally as a single entity, the word itself. While other terms have been used to describe this process—including *proceduralization*, *encapsulation*, and *subsumption*—the basic notion is the same.

The third theory is known as the *instance theory* of automaticity (Logan 1988). According to instance theory, each exposure to a stimulus performed with attention lays down a trace in memory which constitutes an instance of the stimulus; and it is this trace that is allegedly accessed during the act of recall. Instance theory assumes that attention alone is sufficient to cause the creation of memory traces, and that attention also automatically causes associated memories to come to consciousness. Most importantly, each instance of the stimulus is stored and received separately in episodic memory, meaning that instances are similar to the memory people have for events in their lives. To account for the fact that repetition strengthens the characteristics of automaticity (speed, effortlessness, and the like), Logan and others claim that the more traces laid down in memory for given entity, the more likely it is that the brain will retrieve one of those traces quickly.

Within the branches of automaticity theory, explanations of exactly how the memories are created and strengthened are fuzzy. Samuels and Flor simply state that repetition eventually creates a highly accurate representation in memory (Samuels and Flor 1997, 111). Logan, a proponent of instance theory, notes that as there appears to be no write-to-disk command in the mind, memory is said to be a by-product of attending to a task (this is known as the *obligatory attention* principle*).* Accordingly, the number of practices is felt to be key.

4.5.3 Prescriptions from automaticity advocates

Initially, researchers felt that automaticity was essentially dichotomous, such that its various characteristics were either present or absent. Eventually, however, it was proven that certain characteristics could be partially or wholly absent, leading to a "mixed states" effect where certain criteria were viewed as automatic and others, not quite so. This gave way to the notion of what should have been obvious in the first place, namely, that given its inherent nature, something like automaticity can naturally be measured along a more or less smooth continuum.

From a practical standpoint, Samuels and Flor (1997) distinguish two significant thresholds within the continuum: accuracy and automaticity. The first, accuracy, is said to be the level at which the skill has been acquired to the extent that it can be performed with accuracy. This level, however, is described as being insecure, in that it is subject to ready degradation if the item in question is not continually practiced. By contrast, automaticity is the level where the skill will not be easily forgotten, and where the above-described qualities (speed, lack of attention, obligatory response to trigger) have been fully achieved.

One obvious implication of this distinction comes in the context of institutionalized learning, where students of a foreign language learn words and grammar using structured materials that tend to also serve as the basis for testing. Certainly, by test time, most students will have acquired some degree of accuracy for the material; but leaving the material at that state is perilous if follow-up practice is not employed. In addition, studies have been conducted to test whether there are significant differences in the amount of practice needed to achieve these two levels as a function of intelligence. Here, the verdict is mixed, with some studies suggesting that there is greater variance in the time to reach accuracy, and others saying that there is greater variation in the time to reach automaticity. Others have found that employing a distracter task once accuracy had been achieved during practice will greatly enhance the degree of automaticity achieved (Samuels and Flor 1997, 114). Flor describes the role of the distracter task as forcing the learner to "take things to a new level" while preventing boredom and lack of attention from setting in.

Descriptions of the process by which a skill such as word recognition become automatic are less compelling. Samuels and Flor suggests that factors in the development of automaticity include: instruction that facilitates learning the task, opportunities to practice so that the student moves beyond accuracy, and motivation to staying on task (Samuels and Flor 1997, 111-112). One can immediately envision fundamental problems with the kind of instructional approaches one might find in a Chinese curriculum. Not only are there breaks in the school year, but because of the focus on the content of the instructional material, there is not likely to be enough time to reach automaticity for much of the material; and there is also going to be ample time for decay of task skill (word and character recognition). And once a critical exam or the course itself is completed, in the frequent case where the instructor chooses not to make vocabulary responsibilities cumulative, motivation to review previous material is bound to decline in most students.

This leads to three immediate questions in the context of lexical acquisition. First, is it possible to speed the process of acquiring accuracy? If so, how? Second, how does one move lexical competence beyond accuracy and into true automaticity? And on a clearly related note, how does one get students to "stay on task" beyond immediate course requirements, or is that even a realistic goal?[32]

4.5.4 Weaknesses in automaticity theory

Although automaticity theory offers a helpful paradigm in the field of reading, it is nonetheless deficient in critical areas. To begin with, the reader is treated like some sort of black box. Attention "happens," instances are stored, memories are strengthened. Each "instance" appears to be treated the same, even though it is obvious that different types of exposure lead to different levels of memory strength (see Chapter 7).

Relative to other theories of automaticity there is something inherently different about the chunking model of automaticity in that it describes a process that could in principle be merged with either instance or strength models of automaticity. It seems to

---

32 This issue obviously bounds on the realm of teaching philosophy. The upshot seems to be that serious students must be warned that four college years of instruction will not automatically result in language mastery.

be relevant to types of information that are subject to organizational processes. As it turns out, encoding is thought to improve with greater organization (Schunk 2004, 291). However, chunking is more likely to take place very early on, when attempting to attain "accuracy" of data recall. After all, once the ability to chunk a set of data has been recognized, it is not as if one needs to practice that act of chunking numerous times. By contrast, attaining a very rapid response to some stimulus clearly takes many trials. In this sense, the various models of automaticity seem to describe different phenomena.

Second, the notion adopted in instance theory that memory traces are invariably episodic is also puzzling. Initially, the first few times one is exposed to a word, that might be true. But if it were always true, imagine how many "competing episodic traces" a native English speaker might have for the word *the*, or a Chinese speaker for the word 不 (bù 'not'). Although one does sometimes remember exposures to unusual words or phrases episodically, the notion that each and every word is recalled by examining competing episodic traces seems very unlikely. Cognitive psychologists typically assert that thoroughly assimilated words fall into the category of *semantic* memory, which is a context-independent subcategory of declarative memory, in contrast to episodic memory, which is context-dependent. Since semantic memory is not tied to the spatiotemporal context in which it was acquired (that is, it is not tied to personal experience), then clearly something important happens between the phase where words are stored in an episodic manner to when they become fully assimilated. Contradicting the notion of an episodic-memory basis for automaticity, one source describes this transition, and the relation between semantic and episodic memory, as follows:

> The semantic memory is generally derived from the episodic memory, in that we learn new facts or concepts from our experiences, and the episodic memory is considered to support and underpin semantic memory. A gradual transition from episodic to semantic memory can take place, in which episodic memory reduces its sensitivity and association to particular events, so that the information can be generalized as semantic memory. (Mastin)

Finally, automaticity theory also falls short as a model for long-term retention. The notion that all instances of attention to a task are equal is clearly false, as the manner of encoding can definitely make a difference. Besides, as anyone can attest, there are times when one is simply inattentive and fails to take in the information presented. Given that reading with insufficient lexical resources stored in memory is essentially impossible, and that learning characters and words is a huge stumbling block in East Asian languages, automaticity theory's failure to adequately address the issue of long-term lexical retention means one must look elsewhere for answers.

On a more positive note, the notion of automaticity does indeed describe a critical prerequisite to fluent reading. In fact, automaticity appears to be one of two critical aspects of memory strength—the other element being longevity. The argument that would-be readers must achieve automaticity in lower level skills is exactly correct, and coincides with the sentiments of critics of top-down theorists who see reading as an exercise in anticipation and schemata activation. At the same time, because vocabulary (and character and grapheme) memorization is greatly hampered by the process of attrition in both formal and informal instructional contexts, if students are to achieve meaningful competence at anything near a highly advanced (ILR 3.5+) to superior (ILR 4.0) level of reading competence, the issue of attrition (or its converse, longevity) must be addressed. Without providing a viable approach to preventing attrition, formal instructional contexts are apt to perpetuate a rather fragile level of reading that is prone to rapid decay.

To perhaps oversimplify, the guts of automaticity theory leads to one prescription: practice, practice, and practice. In this respect, extensive reading should be expected to play a constructive role, not just because it provides such practice, but because it also has positive motivational effects.

But do all forms of practice count the same, or can certain approaches have a deeper impact on enhancing memory traces than others? Intuitively anyone will readily admit that some experiences or facts are more memorable than others, and it seems only natural that instructors should find ways to exploit this obvious fact.

* * * * *

A brief overview of reading theory leads to the following conclusions. Although there is still a tendency to incorporate top-down strategies as part of the curriculum, in fact, automaticity theorists see top-down strategies as largely a compensatory strategy employed by poor readers. Whereas *fluent* reading might appear to entail an elegant dance between top-down and bottom-up processing, insufficient bottom-up skills will stop reading in its tracks. Metacognition has valid roles—specifically, the ability to think about a passage and to identify problems in comprehension. But using metacognition as a form of compensation is not something to encourage, but rather to avoid, to use when there is no other recourse. Issues with higher level processing are a symptom, not a cause, of poor reading. The root cause is inadequate skill in decoding.

On a separate note, although it is through context that one sees subtle aspects of word usage, this fact should not be taken to mean that extra-context studies are not warranted. Quite the contrary, it is possible to become overly context-dependent in the recognition of characters and words. Studies also show that extrinsic reading may help for vocabulary consolidation, but is probably not the best course for acquisition. This further suggests the need for a separate course of action for vocabulary study.

Finally, although there seems to be quite a bit of variance in the models of automaticity, the basic assessment—that lexical processing must be swift and automatic—sets the stage for positive intervention by establishing a clear goal, one that can help point learners away from the effortful decoding so many commentators decry.

Chapter 5
Reading As Process: Insights from Information Processing Theory

Reading is obviously an incredibly complicated activity. If the goal of improving lexical acquisition is to determine what sorts of interventions can help, then some understanding of the information processing activities that take place during reading is essential. This chapter seeks to identify basic functions at play in reading and to explore existing paradigms in information processing theory. This exploration will be aimed at determining which parts of the process are most ripe for intervention geared to enhancing lexical acquisition.

Given that so many subcomponents are involved in the reading process, it should come as no surprise that many theories have been advanced on the cognitive activities that transpire during language processing. The following discussion looks mainly at the processes involved in reading.

5.1 Basic Information Processing Functions

Information processing models divide cognition into a set of functions. While models of cognition may differ somewhat, some of the most commonly accepted functions include: *attention*, *perception*, *sensory registry*, *encoding*, *storage*, and *retrieval* (Smith and Kosslyn 2009; see also Schunk 2004). Attention selects a subset of the sensory stimuli that are available, and sets limits on what will be processed. Perception involves pattern recognition and attaching meaning to those patterns, operating on the basis of intrinsic organizing principles. A given sensory registry necessarily corresponds to one of the senses, though for the purposes of language, the registries of interest are either iconic (visual) or echoic (auditory). The information in these registries decays very rapidly. Encoding is the process used to store memories. Storage is, rather obviously, the act of retaining information. Finally, retrieval refers to the ability to obtain information based on a cue. In reading, one would expect all of these functions to play a role.

5.2 Processing Lexical Units: A Generalized Model

The description of the obligatory aspects of the reading process in Chapter 1 presupposed a sequence of activities that includes a recursive process of scanning, identification, composition, and interpretation. Figure 1 depicts a rather simplistic algorithm that shows a possible way for elemental units to be processed.

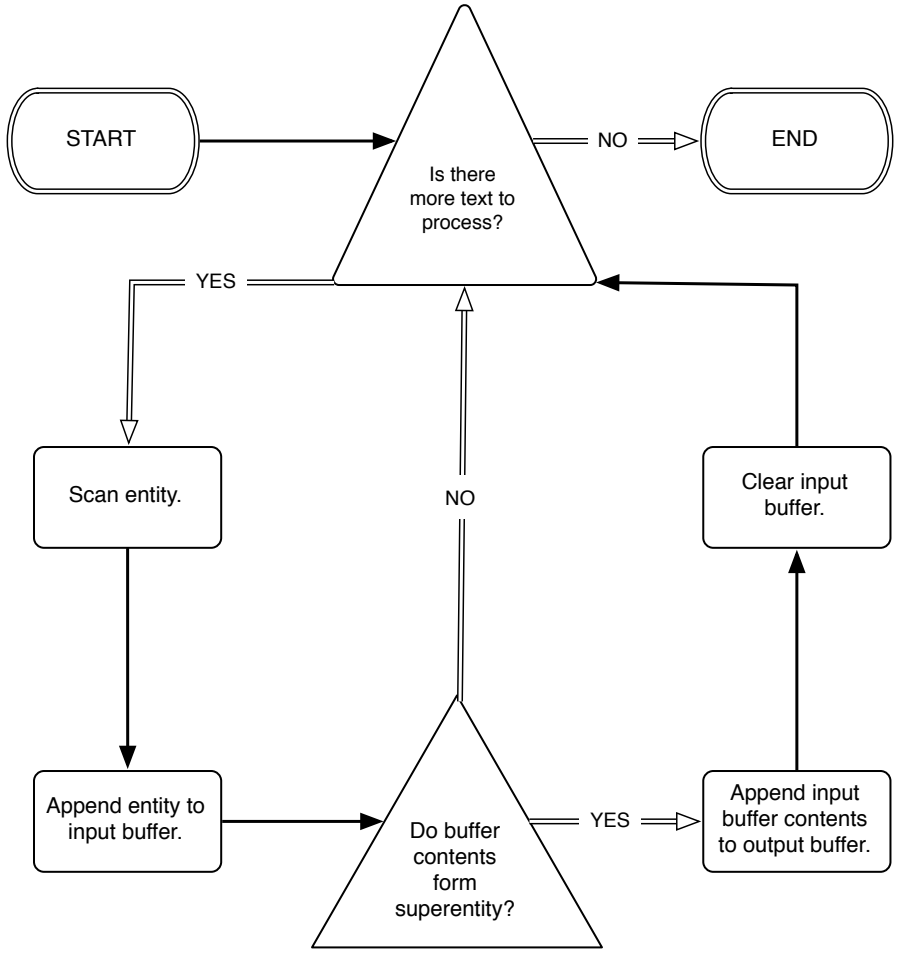


Figure 5.1 General model for scanning entities

As can be seen in the figure, the reader first scans an entity and places it into an input buffer. A determination is made as to whether the contents of the input constitute a super entity, that is, an entity at the next level in the linguistic hierarchy. If not, the reader scans in another entity and append it to the buffer, repeating this process until a

super entity is formed in the input buffer. If the input buffer does constitute a super entity, then the input buffer contents are passed up to the output buffer, where it will be processed as if it were an entity. In other words, this cycle would then repeat at the next level in the linguistic hierarchy, only at that higher level, super entities from the first level become entities at the next level. Meanwhile, back on the lower level, the input buffer is cleared, and the cycle of scanning for an entity resumes.

There are a number of things to note about this preliminary algorithmic description. First, this is a generic representation of a process that happens at a variety of levels. In the case of the Chinese orthography, the process of transformation applies at several levels—from element to character, from character to word, from word to phrase, and from phrase to sentence.[33]

Secondly, this algorithm is recursive. The output buffer of a lower-level orthographic entity constitutes the input buffer of the next highest orthographic entity. The same function is called repeatedly until a critical state is reached—namely, the state where a full super entity has been identified—at which point buffer contents are flushed so that the process can start all over again.

Third, this algorithm presupposes perfect linearity in the reading process, something which most certainly does not fit with the facts of eye scanning experiments, where readers eyes skip back and forth in very complicated ways. However, the idiosyncrasies of individual readers need not concern us at the moment. Rather, this algorithm can represent the "ideal" state where regression is unnecessary.

Fourth, in describing a process of scanning input and placing parsed input into discrete separate units in an output buffer, this basic model glosses over fundamental questions. How does the reader identify a unit in the first place, or to use the model described in the algorithm, how does he or she know a super entity has been found so that it can be passed onward? In what modality is the unit stored in the output buffer? Clearly, the decisive steps of determining that the contents of a buffer are "complete" requires some sort of search-comparison-retrieval function to take place within the mental lexicon.

---

33 It is unclear whether this kind of algorithm would apply at higher levels of discourse.

Finally, as I mentioned earlier, certain steps in the hierarchy are subject to subsumption. As anyone reading introspectively will discover, as automaticity is achieved, it is sometimes possible to bypass the lowest levels of the orthographic hierarchy. In English, for instance, only beginning readers would try to sound out individual letters or process words as syllables. Some studies suggest that skilled readers simply "see words" as discrete units . How does this subsumption principle apply to the above model? Quite simply, in a language like English, it would mean that processing at the {letter → syllable} and {syllable → word} levels have been collapsed, or "subsumed" as automaticity proponents might put it. This observation implies that there will be an obvious difference in "available mental capacity" to words. Whereas a beginner reader is exerting processing at the lower level, filling low-level buffers, skilled readers may be able to begin with a buffer accumulating entities at the word level. The issue of whether such a process happens in Chinese is explored at the end of this chapter.

5.3 Taking a Closer Look: From Character to Word

Although the above model was neat and tidy, and enabled one to assume that identical processes transpired recursively at multiple levels, one gains more insight from taking a closer look at what the process looks like when applied to Chinese word parsing in particular. How does one determine that a super entity has been identified? When is meaning abstracted from the text? When is the sound internally subvocalized? Answering such questions requires a more realistic algorithm.

To this end, Figure 5.2 provides a hypothetical algorithm in the case where the entity being scanned is at the character level, and where the super entity output consists of words.[34]

A quick glance at the process shows not only that it is more complex, but that it contains activities not previously identified. The first issue of interest is the wordhood decision task. Surprisingly, this question is not asked by asking whether the input buffer contents contain a word, but rather whether they contain a word *plus* something that

---

34 Although I have already pointed out that the notion of wordhood is complex in Chinese (see Packard 2000 for a full discussion of the various definitions), it seems clear that for processing purposes, what matters is how the words serve to form meaning in text, and thus I am assuming that "word" here refers to the syntactic word.

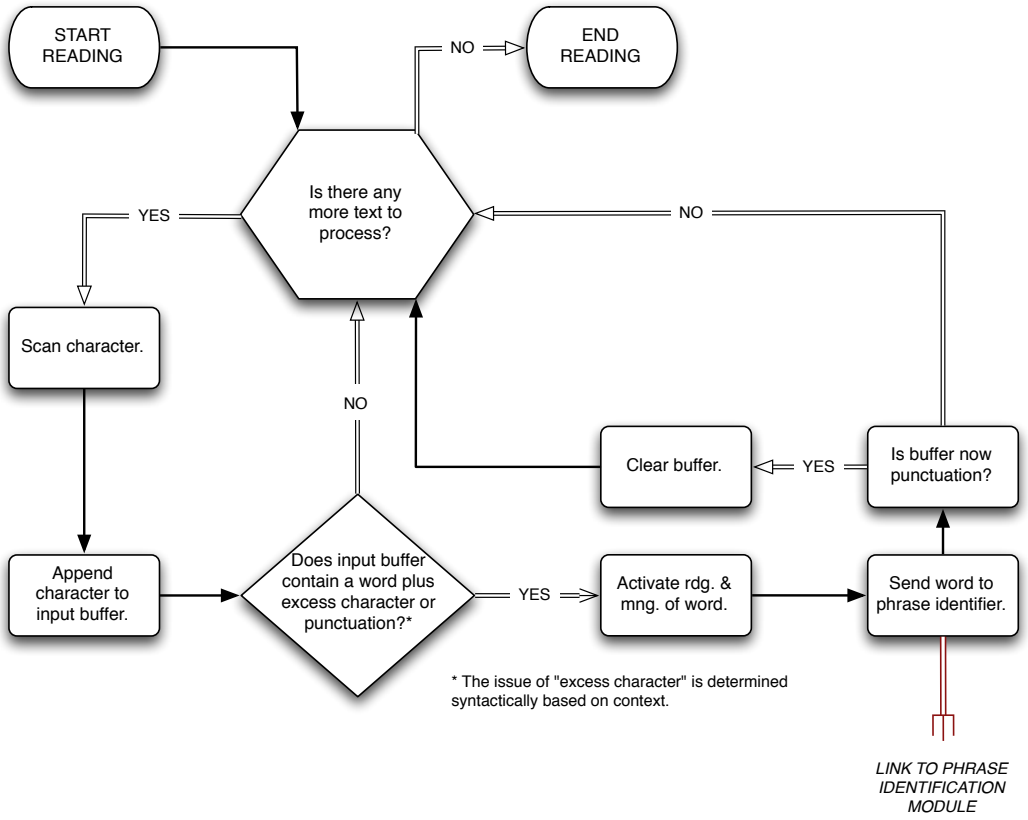


Figure 5.2 Parsing Chinese words from strings of characters

does not belong to the word. This "something else" could either be punctuation or a character belonging to the next word in the sentence. Why not simply identify the super entity straightaway? Because unlike in English and other languages where words are parsed overtly in the orthography, in languages like Chinese, the reader has no way of knowing whether the current character marks the end of the word until he or she reaches something that is clearly not. If one simply asked, do the contents constitute a word, the answer might be yes, but it would not be *the* word represented in the context.

To illustrate my point, imagine encountering a three-character sequence—心理學 *xīnlǐxué* 'psychology'—in the middle of a sentence. Further imagine that the previous word has been unambiguously identified, such that 心 is the first character in the input buffer. As it turns out, 心 is a word meaning 'heart', but it has nothing to do with the meaning in the sentence. Likewise, if one adds 理 to the input buffer contents, the resulting combination, 心裡, is also a word, meaning 'feeling/sentiment'. But again, this is not the word intended. It is only when one scans in the third character that the target

word—meaning 'psychology'—is actually reached. Now the only way to know whether this is the target word is to scan in the next character and see whether the four-character combination of 心理學 plus X, with X being the next character, do or do not form a word. At that point, if X is not part of the word, one deletes X and sends 心理學 to the output buffer as an identified word. If X is a character, it is left in the input buffer for subsequent processing, serving as the first character of the next word.

Although native speakers perform this task automatically, it is definitely different from the wordhood decision task required of native English speakers. The implications are pretty obvious: to become a good reader, one must learn to recognize words quickly, meaning the wordhood decisions must be made instantly. But it also means that to recognize a word, the reader must take a "sneak peak" at the ensuing character to determine whether that is part of the word or not, and only when there is a clearly non-viable character member that the scope of the word is defined. It is only at that point that the word's reading can be safely accessed without risking the misreading of a *duoyinzi* character.[35]

There is something else that may at first seem surprising. Phonology is not activated for each character, but only after the word has been identified. In other words, one first identifies the word, *and then* pronounces it. Why is this so? Because otherwise many words would be mispronounced, and native speakers do not do that. This results from the well-known multivalence of Chinese character readings.[36][37] It also brings us to the next section, which discusses the long and heated debate about the nature of access to readings and meanings in languages like Chinese. One may note that I did not divide access of reading and meaning into separate boxes, but left them as one unit. This "black box" is actually a kind of subtask which has been the subject of much debate.

---

35 In Chinese, a single character may suffice for this "sneak peak" of what comes next, but in Japanese, the need to look ahead before one can make semantic or phonological decisions can be even more pronounced. Even in Chinese, syntactical considerations may come into play.

36 The problem is significantly worse in Japanese.

37 As many beginning students plod along reading one character at a time, it might be interesting as an experiment to encourage them to suppress sound activation until they were sure they had identified a sound.

5.4 Access Paths to Lexical Data

Psychologists have long wondered how lexical data is abstracted from the written word. Major questions can be divided into two issues: How does one get at the sound of the printed word, and how does one get at the meaning of the printed word. But in the case of Chinese, things can get trickier than those basic categories. For example, is the lexicon stored as characters or as words? Is morpheme (i.e. character) data accessed prior to identifying the word, or is the word identified, and then character data analyzed? Another relevant question is which is the salient orthographic unit of native-speaking readers: the character or the word?

5.4.1 Accessing Character and Word Meaning

With respect to accessing word meaning, the major issue at stake is whether, on seeing a word, the reader accesses semantic information in the memory store directly, or whether he or she first processes the phonological representation of the text, and from the sound, then accesses meaning. In this regard, there have traditionally been three main schools of thought.

The first is known as the *universal direct access hypothesis*. According to this theory, words and their meanings are accessed directly in all writing systems, with phonology playing a role only in certain situations.

The second is known as the *orthographic depth hypothesis*. This view asserts that the nature of access depends on the orthographic depth. *Orthographic depth* is a term used to describe the extent to which phonology is explicitly portrayed in the writing symbols of the language. Orthographically "shallow" systems are those where writing symbols portray phonemic information (Finnish would be a good example); by contrast, orthographically "deep" systems are those where the symbols provide little phonemic information (Chinese being a good example). According to the orthographic depth hypothesis, phonology is retrieved more quickly in orthographically shallow systems, and more slowly in orthographically deep systems.

A third hypothesis is the *universal phonological principle*. This view maintains that text activates phonology at various levels regardless of the writing system, with the

nature of the writing system only affecting details of the manner by which the phonology is activated. Thus, one retrieves the sound first, and based on the sound, retrieves the meaning—a mechanism sometimes referred to as *prelexical phonological recoding* (Luo et al. 1998, 833).

Numerous studies have explored this issue using a wide variety of methodologies, with some focusing on the similarities and differences between access in Western language scripts and what many cognitive researchers refer to as *logographic* scripts like Chinese. Although numerous studies show slightly different results, overall, Hoosain reports that words are frequently accessed directly, with prelexical phonological recoding occurring under certain circumstances (Hoosain 1991, 52).

5.4.2 Accessing Character and Word Reading

Frost sums up the question of phonological recoding as follows: Is phonological recoding mandatory? Is it needed for lexical access? Is it needed to access meaning? And how is phonology derived from the orthography? (Frost 1998, 71).

In the case of Chinese, it is actually best to start with the last question. Can the sounds of a character be accessed from the elements? The answer is almost certainly no, except in limited circumstances. As Hoosain pointed out, such a small percentage of phonetics actually match the character reading (25% or less) that accessing sound via the phonetic is usually not viable.[38] This does not mean it is not attempted, especially in cases of rare characters, though various factors come into play, including the type of character (Hoosain 1991, 43). Research tends to point to the fact that phonological activation does take place, though according to some, only upon recognition of the character (Tan and Perfetti 1998, 193), and it should be obvious why: there is no clear subgraphemic mapping from elements to sound. Although one can always set up experiments to show that phonological recoding takes place in Chinese, there is some evidence to suggest that its involvement is lower than in a language like English (Hoosain 1991, 51).

---

38 Many cognitive researchers do not seem to recognize this fact because they appear to believe that the phonetics serve as a possible phonological access route.

5.4.3 The Risk of Relying on Isolated Data

Numerous studies seek answers on the activation paths from graph to meaning/reading. There is, however, a danger in testing characters in isolation. Since many characters have more than one reading, one reading may be activated at the expense of the other when the character is viewed in isolation.[39] Besides, as Hoosain notes, the phonetic component of characters is frequently overplayed. One optimistic account claims 38% of cases are "helpful" (Perfetti and Liu 2006, 228). As mentioned in Chapter 2, Housain notes that only 26.3% of characters are pronounced the same as their phonetic; and if one takes frequency into account, only 18.5% of any encountered character will have a congruent phonetic. By contrast, semantic radicals at best hint roughly at a semantic category with which the character is associated, and much knowledge of graphical etymology would be needed to guess the meaning of a character based on its constituent elements. And even then, the number of characters whose meaning could be guessed based on their constituent components must certainly be low. Therefore, one ought to be circumspect in interpreting any study that compares priming effects of "semantic" radicals vs. those of "phonetic" elements. Even if sublexical activation takes place through the recognition of intracharacter elements, a given phonetic would pre-activate far fewer characters than any given radical. In fact, many radicals would pre-activate so many characters as to render pre-activation cognitively uneconomical.

In reviewing the literature, Perfetti, one of the most active researchers in this field, himself concludes one can expect to see similarities and differences across writing systems when it comes to considering the implications for learning to read. In any event, it seems clear that while access paths from text to comprehension form an interesting avenue of scientific inquiry, nothing learned so far seems solid enough to serve as the basis for prescription in reading instruction.

An undeniable fact is that character meaning frequently requires extra-character information (the context of the associated character in a compound), whereas reading interpretation is not nearly as critical. This suggests that not too much should be read into the fact that sound appears to be accessed somewhat more quickly in some contexts.

---

39 The situation is even worse in Japanese, where each character maps to many sounds in very complex ways based on the extra-character context.

Given that even Chinese native speakers sometimes struggle to identify the meaning of individual characters (Hoosain cites the example of 巴), it is possible that native speakers concern themselves with pronunciation of characters and meanings of words, but not necessarily the meaning of characters. This is probably only natural for any language. Native speakers would naturally be less interested in intraword morphology than foreign learners, who may need to use such intraword morphology strategically.

But if viewed in context, the decision will be made on the basis of identification of wordhood. Therefore, the more interesting question is how wordhood access transpires—first by sound, first by meaning, or both simultaneously. This was clear from Figure 5.2 in the previous section, where reading and meaning could only safely be activated after identifying the word.

5.5 Identification of Word and Activation of Associated Meanings

Recall that following the wordhood decision task, there was a "black box" for activating the reading and meaning of the word. But does the reader access reading or meaning first? The answer is likely to depend on several factors, including whether one is a native speaker and whether is also a highly competent listener. To take a non-experimental approach, consider a spoken utterance that starts thus:

tā suǒ ....

In actual fact, one is not receiving "letters" but streams of acoustic signals, but these are interpreted (somehow) into distinct units recognized as syllables, which in Chinese nearly always also correspond to morphemes. Ignoring this initial phase of complexity, let us examine the process that must take place when the reader identifies individual patterns of sounds that correlate to well-known syllables that are viable in Chinese. Since the pronunciation (reading) and not visual characters are presented as input to the listener, the listener must engage in semantic activation cautiously. I will take this one word at a time to illustrate my meaning.

1. tā

A first best guess would be "third person pronoun" *he*/*she*, though of course this could be just about anything starting with tā. Unless previous context was provided, there is no way to distinguish between *he* and *she* or even *it*. Native speakers are no doubt comfortable with this ambiguity, though it might seem odd to the foreigner learner.

2. suǒ

Assuming that the *he*/*she* interpretation is being followed, this syllable could mean anything. A lexical decision is simply not possible here. It is only upon hearing subsequent material that the listener can begin to guess how *suǒ* contributes to the sentence. Consider these three possibilities, with quasi-literal translations showing the English meaning for *suǒ* in italics.

tā suǒ yào shuō de shì
***what* he wants to say is ...**

tā suǒwèi de hěn shǎo qián shì
for him, *so*-called 'very little money' is ...

tā suǒ le mén yíhòu
after he *lock* -ed door

3. a) yào, b) wèi, c) le

Now a decision can be reached. These three sentences head off in very different directions, but the semantics become a bit clearer as more and more likely meanings are established in the interpretive buffer.

As a rule, then, when listening, one would expect the listener to engage in best guesses, and to revise opinions as necessary as new information becomes activated.

Now consider the case of reading. It is still not possible to make absolute judgments on meaning, but it is possible to "pre-activate" possibilities differently than when listening.

a) 她所要說的話是。。。 **(*what* he wants to say is ...)**
b) 她所謂的很少錢是。。。**(as for him, *so*-called 'very little money' is ...)**
c) 她鎖了門以後。。。 **(after he *lock* -ed door)**

Because the input consists of characters and not the sounds, semantic decisions can typically be made as soon as each word is identified.[40] In a) and b), the characters corresponding to sound *suǒ* are identical (they are both 所) and thus require further input before a lexical decision of its role can be made. By contrast, the character in c) represents a full word that clearly heads in a different direction, and almost certainly constitutes a verb 'lock' (鎖). In other words, although final lexical decisions cannot be made until later, preliminary activation may be invoked upon encountering the character.

So what is the difference between reading and listening? In reading, one sometimes receives semantic clues earlier than in the case of listening, and semantic pre-activation can sometimes take place a little earlier. In both cases, some element of the "sneak peak" principle must be employed before a decision can be made. In the case of listening, one delays decision-making, waiting for the next syllable so that one can determine which word the previous syllable was a part of. In the case of reading, one must look forward to determine not just word boundaries, but also—in the case of *duoyinzi*—which pronunciation should be associated with the previous character.[41]

5.6 Phrase Level Construction

Figure 5.2 showed a hypothetical algorithm for word parsing, with the output being identified words that are passed along for higher processing. One obvious question is, What happens at the next level? Certainly, it is possible to posit that yet another algorithm, akin to Figure 5.1, occurs, whereby words are chunked into

---

40 Occasional exceptions would arise when words have widely varying meanings depending on context, as is the case with "garden path" constructions.

41 Note that a similar "sneak peak" phenomenon takes place in English when deciding how to pronounce sequences of letters. The phonological realization of vowel combinations is especially context-dependent and frequently cannot be activated until the entire word is identified. What distinguishes English, however, is that—except in the occasional case where a single word is written using with a space between two apparent words—there is absolutely no ambiguity about wordhood, so one can proceed directly from word identification to the need to activate the associated reading and meaning.

meaningful units (possibly things roughly corresponding to NP, VP, etc.). Figure 5.3 shows a rudimentary sample that might fit the bill. It mirrors Figure 5.1, and for that reason is a drastic oversimplification of the process. Just as in the case of word parsing, some signal must be received to indicate that a phrase is complete. And when the completed phrase is passed on for interpretation, it is unclear what representation might be involved. This figure also does not indicate how there is an interaction between this algorithm and the lower-level, though as was shown in Figure 5.2, clearly there is one, as completion of a word causes information to pass upward.

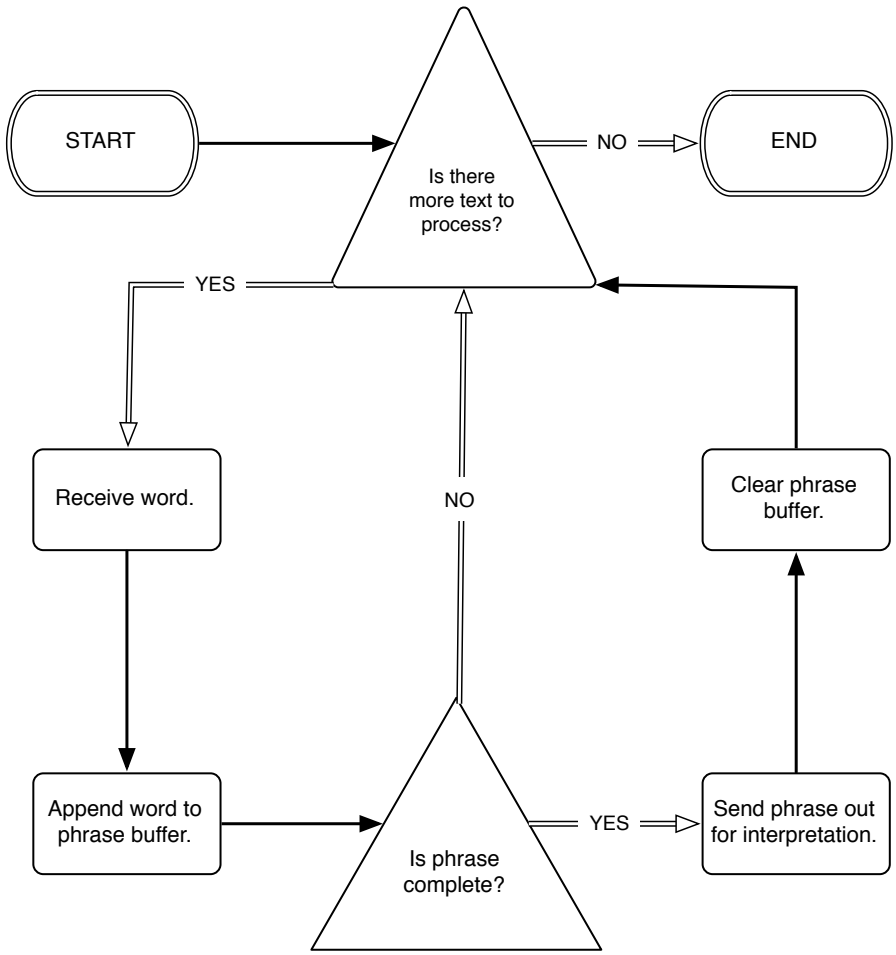


Figure 5.3 Chunking words into phrases

Although Figure 5.3 is oversimplified, it gives the basic idea of receiving input, chunking it, and passing it onward. Decisions such as *Is phrase complete?* are performed instinctively and—unless they happen to be watching their own thinking patterns—

unconsciously by native speakers. As with word parsing, phrase boundaries may be recognized by the presence of an additional word that no longer fits.

5.7 Higher Order Processing: Interpretation and Storage

Presumably at the next level, phrases, or in the case of complex utterances, groups of phrases are interpreted. Whatever its representation, as higher level chunks are completed, a message would be set back to "flush" the phrasal buffer, which is no longer needed. The linguistic interpretation would be some sense that somehow the interaction of phrases is appreciated such that "this NP did X to that NP" or "NP did VP".

The exact manner or representation used for such interpretations is a matter of debate, though some believe that a *semantic* representation constituting an amalgam of perceptual and functional data is used. The idea is not that words and phrases are represented as such, but that "combinations of perceptual information, functional information" and other information would be used for the representation (Smith and Kosslyn 2009, 504).

The interpreted information is then presumably encoded in long-term memory, though the quality of the encoding can vary greatly. (This subject is explored in the next chapter.)

Beyond these higher level processing stages, it is also possible to have metacognitive *reactions* to what is being interpreted or understood, or other thoughts about what one is reading. These metacognitive reactions and thoughts would constitute the domain of activity typically referred to as top-down processing.

Figure 5.4 gives a very crude sketch of the process of reading identified so far. Conceptually, a metacognitive module runs quietly in the background, monitoring the lexical meaning construction sequence. This background monitor is spurred to action under a wide variety of circumstances. With poor readers, a common cause for metacognitive intervention would be a breakdown somewhere along the lexical processing sequence (this could happen at any stage), triggering a sense of "not understanding." An extremely common reason for such a reaction would be if the lexical meaning (or reading) initially retrieved turned out to be incorrect given the context, at

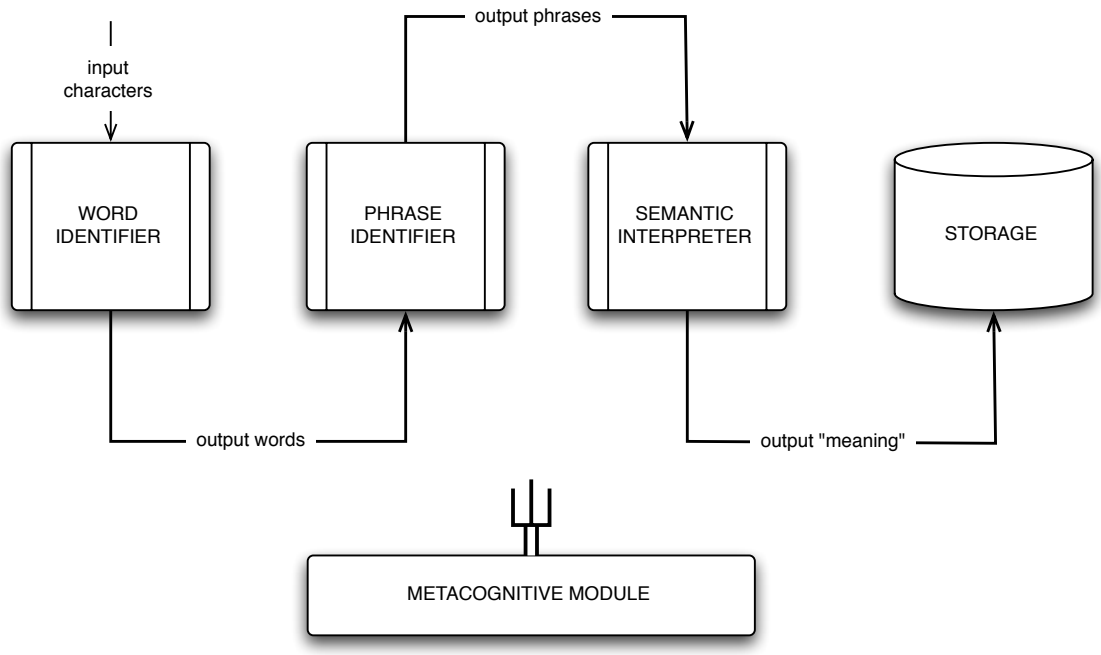


Figure 5.4 Meaning construction and background monitoring

which point backtracking and a new search would have to be performed. Excessively slow processing might trigger a realization that one is not "grasping the big picture." Or it could be that the text itself provokes a reaction to the material, causing the reader to reflect on what he or she has just read. Otherwise, there is no need to get involved, and reading proceeds along the lines of construction suggested by the figure.

5.8 Working Memory and the Limits of Processing Capacity

The buffering of information has so far been shown at three separate levels: characters are placed into a buffer to form words, words are placed into a buffer to form phrases, and phrases are pooled together to create interpretable relations among the actors and agents in a sentence. The general rule is that at any given level, entities at the input level are chunked until they form a super entity, and this super entity is output so that it can serve as the input-level entity for the next higher-order process. The process of chunking or constructing meaning continues until interpreted semantic information is somehow stored in memory.

Who or what performs all these tasks? Typical cognitive models ascribe such processing to what is usually referred to as *working memory*. Since models of *working*

memory seems to be evolving (see Gathercole and Baddeley 1993; also Smith and Kosslyn 2007), and because other possible models have been posited, its exact nature will not be spelled out in detail here. However, certain basic notions that persist in most models are that it contains a *central executive* that manages information, a *phonological loop* that retains auditory input, and a *visual sketchpad*, that stores visual information. Executive processing determines how the information in the auditory and visual stores are to be manipulated.

Day and Bamford claim that words are converted to sounds, which are placed into the phonological loop so that lexical (semantic) data can be accessed, at which point meaning is understood (Day and Bamford 1998, 12-14). Gathercole and Baddeley state that the picture is more complex than that, however, in that whereas very simple sentences might be interpreted "on-line" (that is, as they come in), more complex material would require rehearsal and backtracking (Gathercole and Baddeley 1993, Chapter 8).

It is less important to explore the details of this working memory than to note one thing that cognitive scientists all agree on: whatever form it may take, working memory's processing capacity is limited. In other words, if one looks at the functions identified in Figure 5.4, the processing capacity available to perform word identification, phrase identification, and semantic interpretation is shared. Presumably, for a fluent and rapid flow of text processing, there must be a certain balance of interactions among these three modules that maximizes the rate of flow from one module to the next.

All the previous figures from Figure 5.1 to 5.3 depicted ideal cases wherein certain information processing functions simply worked. In 5.2, lexical readings and meanings were accessed by a black box whose nature was not clarified; in 5.3, groups of words are recognized as constituting a phrase. But what happens when something goes wrong?

To take a simple example, imagine the word identification algorithm when a character is not recognized. Because an unknown character is added to the buffer, the decision as to whether it belongs with the previous stream of characters is no longer certain. Conscious *attention* (the operation of the metacognitive module) awakens to intervene. Because the character is unknown, there is no ability to activate a reading,

though the metacognitive module might try to guess if it recognizes a phonetic component in the character (and that strategy in itself assumes a pretty sophisticated understanding of character elements). Not only will phrase recognition be hampered, but semantic interpretation will be impaired, since the exact meaning of one of the phrases is no longer precisely known, and even the encoding in storage is likely to be affected. If the encoding in storage is affected, then any attempts to reflect on what has been read will be impaired since the metacognitive module trying to reflect will not be quite sure what it is reflecting on.

This is why in Chapter 1 I asserted that the construction of meaning from small linguistic units to larger units is "absolute"; it cannot be done away with, and no amount of "top-down" manipulation or expertise can fully undo the damage of gaping glitches in the flow of information from one module to the next. In brief, the critical lexical data must be retrieved for the reading to take place in anything of a meaningful fashion.

The second important realization is that flow must be smooth. Imagine, for instance, a case where for some reason (perhaps the presence of an unknown idiomatic construction), failure occurred at the level of semantic interpretation. Words were processed, phrases were identified, but to the non-native reader, the resulting phrases do not make sense in the context. At this point, one can expect an intervention, "stop signals" would be sent to lower level modules, and some intervention will be performed —likely candidates being an instruction to reread the passage, attempts to engage in contextual guessing, or a decision to ignore the misunderstood passage and move on. Again, flow is interrupted.

In the context of Figure 5.4, the insistence of automaticity advocates becomes all that much more compelling. If, for instance, within 5.2 lexical meaning/reading retrieval is slow, then since the phonological loop storing information can only hold information for so long, then only so many words will be passed along in a given period of time. That means that in a given number of seconds, only one or two phrases will be input. And that means that semantic interpretation will be taking place with smaller chunks of data. The reader will have a harder time getting a feel for the "big picture" of the discourse as a result. Here again, the issue is not some need to practice top-down

strategies, but the fact that the information is percolating up too slowly to be of full use. This is precisely how the widely disparaged "effortful decoding" comes about.

In short, for fluent reading to proceed, lexical retrieval must succeed, and the flow of retrieval and processing must be fast and smooth. Fortunately, it is by recognizing areas where things become bogged down that one can attempt to decide where intervention is in order.

5.9 Serial vs. Parallel Processing

One perennial subject of debate has been with respect to the notion of whether mental processing operates serially or in parallel. With respect to the issue of the temporal sequencing of activity during reading, there are obviously three logical choices: the first is that the various activities take place serially, with one task being accomplished at a given time, and the next task beginning only when the first task is completed. The second logical possibility is that the mind is capable of parallel processing, performing a variety of tasks simultaneously. Finally, as both of these models were largely influenced by information processing in the field of computer science, it is also possible to envision a third, hybrid model, one akin to the strategy known as multithreading in computer programming. In this model, any number of tasks may be attended to within a certain timeframe; but at any given time *t* only one of those tasks is actually being worked on. Rapid switching between tasks leaves the impression that the tasks are being attended to simultaneously, whereas in fact, only one action is taken at a time.

The human brain is clearly capable of performing a host of activities simultaneously. We move, breathe, see, listen, and perform numerous other activities simultaneously. More difficult to answer is whether consciousness or conscious attention is a like a spotlight or more like a light show. In cognitive psychologists' terms, the question is whether the system is limited to one controlled activity (that is, one requiring conscious attention) and a certain unidentified number of automatic activities, or whether it can engage in multiple controlled activities.

For the purposes of this dissertation, I do not believe it is critical to choose from among these three possibilities. As an approximation, it is sufficient to pretend as if it

acts serially, or in a multithreaded fashion. That said, the fact that metacognitive interventions can lie dormant and suddenly kick in when something goes wrong seems to suggest that there is some capacity to monitor passively and yet be capable of judging if another activity requires help. This strongly pushes the argument in favor of parallel or multithreading models of processing.

5.10 Subsumption in Lower Level Processes

Looking back at the model presented in Figure 5.4, one might feel that the entire process is extremely laborious. Characters are serially chunked until a word is identified, words are put together to form coherent phrases, and so on. The equivalent in English would be to assume that readers use the letters to "sound out" the word, at which point they would access the meaning. Scientists have shown that that is not true, however. Because the spelling system of English is so haphazard, with many phonological relations possible for any given combination of letters, fluent readers tend to recognize words almost logographically, as basic units. If that is the case, would not something similar happen for Chinese?

The answer is almost certainly yes. Rather than accumulating characters one at a time, native speakers almost certainly "see words" on the page despite the lack of overt parsing in the orthography. If words are seen directly (essentially eliminating the left-hand component in Figure 5.4), then far less processing needs to take place at the lowest level. Instead of *Scan character*, the lowest level becomes *Scan word*. The ability to "see words" on the page would go a long ways to enhancing the mental processing capacities available for higher levels, and fits with the chunking or subsumption model of automaticity.

* * * * *

Chapter 4 explored paradigms from reading theory, and while that branch of inquiry has turned up intriguing notions such as that of automaticity, the reasoning behind the nature of that activity was somewhat suspect. The rough "constructionist" model of processing described in this chapter accords with the observed behavior: namely that accurate and rapid identification of lexical units is needed to read fluently.

One must be able to identify words, know how words fit together to form meaningful phrases, and of course be able to see the relationships occurring between the phrases represented in a passage of text. The better one performs these lower level tasks, the easier it is for the mind to handle larger amounts of information at the higher levels. Failure, or system processing breakdown, could theoretically occur in any of the modules described. Breakdown is likely to invoke intervention by a metacognitive function that attempts to deal with the problem, and during that time, processing comes to a stop. Thus, if any of these tasks cause the system to bog down, comprehension will suffer as a result. The next chapter in part seeks to explore what one might do to help avoid such potential failures. To this end, one must look at encoding, and also develop a model for measuring the quality of lexical storage that would aid in fluent reading.

The scenario presented here differs subtly from previous interactive models of bottom-up/top-down processing. Rather than envisioning a fluid interaction between equal partners, the model depicted here asserts that, under normal conditions, all readers work upwards, the only difference being that skilled readers sometimes work upwards from a higher starting point, due to the chunking or subsumption of certain lower-level identification algorithms. In any event, the lowest level of elemental symbols must be correctly interpreted and fused into larger units for reading to take place. The only question is between whether that process has been automatized or not. The fact that lower-level processes are done automatically or unconsciously by the native speaker, and only through effortful decoding by the novice learner or non-native, only attests to differences in skill levels and the need to apportion cognitive resources to different levels. But there is no denying that the processes must take place.

Chapter 6
Memory Concepts Relevant to Lexical Acquisition

The survey of reading theory and the analysis of reading as process in the last two chapters illustrated why successful reading is so heavily dependent on lexical processing. The overall conclusion was while that metacognitive interventions and reactions inevitably take place during reading, fluent reading only transpires when the flow of information from lower levels to higher is smooth. Two prerequisites were identified: longevity and automaticity. As these key prerequisites are byproducts of memory, this chapter looks at important memory concepts that shed light on the acquisition process.

There are a host of concepts associated with memory—including plasticity, inhibition, transfer, context, and others. However, in its most basic form, memory is typically broken down into three fundamental components: *encoding*, *persistence*, and *retrieval*. Although the delineation between these facets of memory is not as clear-cut as one might expect, this traditional subdivision of the "basic three" provides a good structure for exploring how memory processes interface with the lexical data to be acquired.

6.1 Encoding

Encoding is the alteration of brain structures that permits an event or information to persist. Many cognitive scientists believe that encoding is not "special" in the sense of requiring separate cognitive processes to transpire; rather, it is the byproduct of perception and comprehension of sensory input (Hasselmo 2007, 127). During the encoding process, a loop takes place between attention and perception, during which the subject selects some portion of the target entity to attend to, and via the senses, engages in perception. In response to sensory perception of the object, a number of potential responses may be triggered automatically: a sense of familiarity, a

sense of total unfamiliarity, or a recognition response.[42] The cyclical acts of attention/ perception in and of themselves are said to result in encoding.

Despite the temptation to view it as "an event," encoding is not a single-shot affair. Even upon the initial learning session, a student is apt to perform multiple "mini reviews" until he is satisfied that he has learned his target. Different theories exist as to what may be happening neurologically, and even the notion of an engram or memory trace is sometimes in dispute. But while the notion of a single encoding is perhaps realistic for episodic memory, it certainly does not apply to the learning of semantic knowledge. Moreover, the acquisition of lexical knowledge to be encoded entails an ongoing process, which some scholars describe as a continuum between reception and production (Melka 1997). Students of Chinese do not learn all the meanings and all the readings associated with a given character at once. As is the case with second language learners of any target language, the sophistication of vocabulary understanding only grows with time, as appreciation of the context and manner in which words are applied increases. Thus, what might be viewed as a single linguistic item will inevitably undergo multiple rounds of encoding.

6.1.2 Strategies to Enhance Encoding

Some of the documented strategies used to enhance encoding include *attention*, *generation*, and *spacing* (Smith and Kosslyn 2009, 202-209). Paying full attention at the time of encoding is, unsurprisingly, very important. Generation refers to the act of reproducing or utilizing a target, and contrasts with passive exposure to the learning target. Spacing refers to the amount of time spent between repeated study or testing incidents; in the long term, distributed study usually works better than massed study. These three encoding-related parameters are discussed further in the next chapter.

Other encoding facilitators include *organization*, *elaboration*, *meaningfulness* (Schunk 2004, 160-161). Organizing the information and making it meaningful can greatly aid retention, as is elaborating on the data, such as by drawing associations to

---

42 As mentioned in Chapter 3, whereas vocabulary experts view word familiarity as a continuum along which successive levels of knowledge about a word lie, cognitive scientists argue that familiarity activates separate brain regions from recollection (Yonelinas et al., 3002 - 3008).

known information.[43] These encoding-related parameters are discussed at length in this chapter.

The fact that what one perceives depends on what one already knows is highly relevant to the issue of learning characters and character compounds. Depending on what information is made available to the student, a certain character component may be viewed as an arbitrary selection of strokes or as meaningful symbol representing a certain sound or meaning or both. The same is true of the morphemic relations between constituent characters in a compound. If one knows the constituent character meanings and readings, the compound's meaning may make sense if the word has some level of morphological transparency. If the student does not know those individual meanings and readings, then he is learning the word more or less logographically. Thus, the distinction between mandatory and potential facets of character acquisition discussed in Chapter 3 can weigh heavily on how encoding takes place.

While the exact nature of the representations of memory traces in the brain may be hard to pin down, whatever form they take, they clearly depend on how one goes about trying to memorize the target. A student who tries to memorize a character by writing it a hundred times will be creating a sequential psycho-motor pattern of behavior that he or she tries to associate with the cue. At the same time, assuming he is not writing with eyes closed, he will also be observing the writing, and so attending to and perceiving the shape multiple times. One should expect this to have a very different encoding effect than, say, an analytical approach that subdivides the character into meaningful components. The analytical approach involves the subsumption or "chunking" of strokes into elements to which some form of meaning or at least familiarity is attached, reducing a collection of say 15 strokes to a compounding of one or two previously known elements.

6.2 Persistence

Eichenbaum defines persistence in memory as *the temporal extension of a modification in neural representation or behavior resulting from experience* (Eichenbaum 2007,

43 The fact that organization helps goes a long ways to account for the effect that grouping characters by core component greatly aids meaning recall (Child 2004).

193). In some ways, persistence is the defining characteristic memory, as without persistence, memory simply could not exist. By definition, persistence implies that one can distinguish between accessibility and availability. A memory trace may exist in the brain, but at certain points in time, it may not be accessible. Persistence is also a feature of memory that is almost impossible to study at the neurological level. It cannot really be perceived as is, but rather only inferred on the basis of successful retrieval.

In recounting William James' belief that a good memory was achieved by forming diverse and multiple associations with targets for learning, Eichenbaum states that persistence depends "in part on a cellular persistence mechanism and in part on the size and organization of the network of persistent associations in which a memory resides" (Eichenbaum 2007, 195). The associative structures are therefore considered a key to facilitating persistence. What does this imply in the context of Chinese lexical acquisition?

6.2.1 Lexical Networks

Cognitive scientists hypothesize that lexical data is stored in the form of networks of linked nodes of information. The nature of these nodes is a matter of speculation, but experiments have shown that related information is indeed linked (Schunk 2009, 169). When the proper cue is presented, the target node is activated, but so are a host of related nodes.

The associations formed between discrete pieces of information will depend on what information is made available to the student. Specifically in Chinese lexical acquisition, the associations a Chinese learner forms are a direct byproduct of the material he is learning from. Lexical data exposure can be classified as either contextualized or decontextualized. Contextual exposure entails the presentation of words in context.[44] Decontextualized exposure provides no such context, as would be the case in a random word list with translations. Some foreign language instruction

---

44 Note that in cognitive studies, context can have a different sense, and that is the composite of environment, location, style of presentation, or even surrounding material provided at the time of the encoding. Context in this sense also plays a role in the encoding process, but may result in unintended context-dependencies that are unhelpful.

researchers strongly advocate ensuring that words are always learned in context on the assumption that it is only through context can one understand the word's usage.

It cannot be denied that one must see examples of a new word in context to begin to appreciate its usage, and much of the natural or "organically generated" knowledge that children acquire for their native language will be a mix of haphazard associations and coherent or thematic associations based on real-life experience. To take a trite example, children may learn the names of several animals at roughly the same time, and may often "experience" them nearly simultaneously, whether at a zoo or within a children's book.

Likewise, with foreign language materials, it is not uncommon for textbooks to use chapter themes to help introduce related vocabulary. Such an approach should in principle help to generate "organic" networks of semantically related information. Such networks are liable to attach to other words that are not necessarily related, but happen to arise in the same material. Lexically, the result of such an approach is that when students learn the associated vocabulary, their associations will be based on this material. Consider a simple example from common words that might be found in an introductory textbook. Figure 6.1 shows the associations that might result from thematically organized materials. The figure is based on a textbook chapter whose theme is travel.

There are a number of things to note about the figure. Each lexical item contains three components: a meaning (in English), a reading (expressed here in pinyin), and the orthographic realization in Chinese characters. Bidirectional solid lines show the bidirectional connections among these three components. In addition, the figure proposes the possibility of loose associations (shown by dashed lines) between concepts/meanings in English; here, the examples are an association between *travel* and *tour guide*, as well as between *travel* and *Mexico* (a plausible assumption for American students). Such semantic associations might be attributed to cognitive *schema* resulting from life experiences.

At the same time, the figure posits the possibility of the student drawing orthographic connects between some of the familiar characters in the words, notably the characters for west (西) and elder brother (哥) in Mexico (墨西哥). The connection is

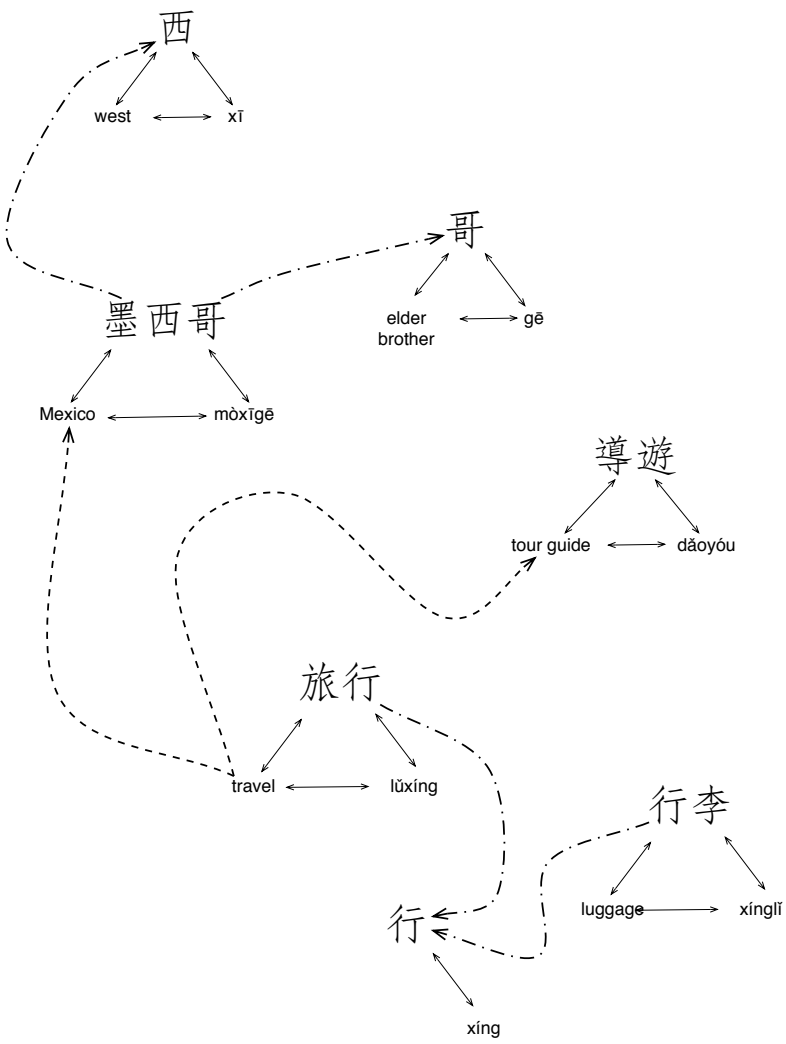


Figure 6.1 Lexical associations based on textbook materials (Chinese) and semantic schema (English)

based purely on the form of the graph itself in this case, as there appears to be no semantic motivation for the use of the characters. At the same time, there is a character (行) shared between two of the vocabulary items, and its reading (xíng) can be deduced, though it is quite possible that a beginning Chinese student may not be aware of the meaning of the character on its own. If he or she happens to know that the character means "OK" in some contexts, then its usage may appear to be incongruous, or perhaps purely phonetically motivated.[45] Other characters are learned simply as part of the word

[45] The semantic significance of this character to a beginning Japanese student would be totally obvious since it is used for the everyday word meaning 'to go'.

without a sense of their semantic contribution, and with their phonetic contribution being deduced based on that of the corresponding syllable in the compound. In sum, one would expect two forms of connections to be made: pre-existing connects due to the student's *schema*, and newly formed associations / connections formed as a result of the presentation of the text.

Natural thematic associations are not the only sort of information that can link lexical data. For a given character, knowledge about its constituent elements, multiple meanings and readings, and compounds the character is found in can also be linked. Figure 6.2 shows an example of a network constructed based on orthographic principles. (Character elements are shown in blue.) Some schema-based semantic associations

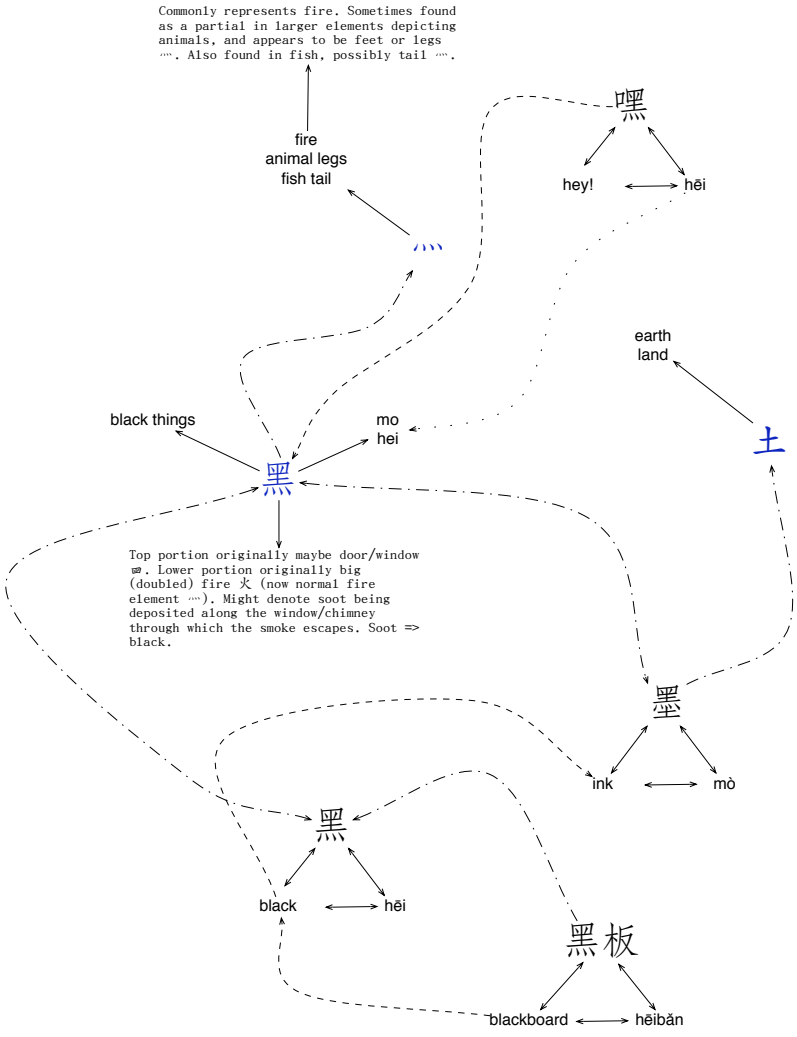


Figure 6.2 A network organized based on orthographic principles

(dashed lines) remain, as can be seen between *black* and *ink*, or between *black* and *blackboard*. But most of the the connections are based on graphic principles (dash-dot lines), connecting elements to characters that employ them, or characters with compounds that employ them.

Both types of networks are plausible organizing strategies to enhance encoding, and each has its advantages and disadvantages. The advantage of the sort of thematically based network that one might expect to result from standard textbook presentations is that the words learned will be part of a theme that forms the basis for learning contextualized material. A whole swathe of everyday Chinese vocabulary and expressions will be covered at one go. Dialogs can be natural, and students will learn much of the vocabulary needed to handle specific real-life scenarios. The downside to this form of thematic organization is that most of the orthographic connections between characters and words is opaque, and that between elements and characters is not provided at all.

Looking at the network organized based on orthographic principles, the obvious disadvantage is that if one extended such a network, semantically, the characters and words might be connected in less expected ways. Some semantic connections still exist, but these are a byproduct of some of the semantic principles in the compositional rationale of the constituent characters. If information about phonetic realizations is provided for elements, then phonetic connections may be drawn as well (see the dotted line between "mo/hei" associated with 黑 and the "hei" associated with 嘿). But characters constituting pure phonetic compounds will be totally unrelated on the semantic level. Although this network is a very fine organizing principle from the standpoint of learning the orthography, it is terrible as a basis for developing natural reading materials.

Organization is an important strategy for enhancing encoding. Moreover, scientists believe that material is actually stored in the form of semantic (and other) associations within the brain. But if one wishes to translate this to real life issues such as learning the Chinese orthography, one is faced with a dilemma between one that is better suited to an "organic" approach to learning lexical items in context, and a purely

orthographically motivated organizing principle that would assist in in the recall of character and compound graphical forms.

6.2.2 Consolidation

Beyond the notion of lexical networks, another key component of persistence for semantic information is *consolidation*. Consolidation is the notion that processing subsequent to the initial encoding takes place to change the way that a memory is stored. In that sense, consolidation is essentially a process by which memory traces become—to borrow Wickelgren's phrase—less fragile.[46] It can occur without attempts to strengthen the memory, and is not related to the strengthening of traces through such things as testing or repeated exposure.

Table 6.1 Types of consolidation cited in the literature

Type	Memory Phases Involved	Hypothesized Process
(brief store) consolidation	WM => LTM	transfer from prefrontal cortex to medial temporal lobe independent of protein synthesis
molecular consolidation	STM => LTM	protein synthesis facilitating conversion from STM to LTM
systems consolidation	LTM => LTM	consolidated LTM in medial temporal lobe (specifically, the hippocampus) moves to other areas, possibly neocortex

Confusingly, consolidation is used in three different senses in the literature. These different usages are attributed to different portions of the brain, and affect different stages in the learning process, as shown in Table 6.1. *WM => LTM consolidation* is inferred from the fact that it is possible to disrupt information stored in working memory so that it does not get transferred to long-term memory. *Molecular consolidation* refers to the transfer from short-term to long-term memory, and is thought to involve protein synthesis that helps solidify memories. Finally, *systemic consolidation* is refers to the moving of memories from dependence on the medial temporal lobe (the

46 The fact that surviving memory traces grow less fragile over time has been recognized at least since the late 1800s.

hippocampus, in particular) to a form of persistence that is less dependent on that region (LeDoux 2007, 171 - 172).

Although it was scarcely recognized by experimental psychologists, until only recently the notion of consolidation has widely been accepted as fact by neuropsychologists, (Wixted 2004). Now, however, researchers are troubled by a number of issues, including the fact that as more and more experiments are conducted, the time span of consolidation has begun to grow longer and longer. This has in turn lead to alternative theories such as multiple trace theory, which argues that standard consolidation ("systems" consolidation in the above table) operates for episodic memory but not semantic memory (Nadel 2007, 178-179). New notions such as "re-consolidation" have also entered the fray, to the effect that some researchers believe the notion of consolidation is a problematic construct. Even the notion of an engram or memory trace has been called into question.

6.2.3 Negative Factors in Persistence: Decay and Interference

For centuries, scholars have tried to pinpoint the primary cause of forgetting. Historically, three theories emerged: the first is that forgetting resulted from the decay or degradation of memory traces; the second is that it resulted from interference from other memories; and the third is that it resulted from the fragmentation of memory traces over time (Baddeley 1990, 243 - 247). The first two of these theories seem highly relevant to lexical acquisition, and so will be discussed further.

6.2.3.1 Memory Trace Degradation

In decay theory, items remembered fade over time, exhibiting a form of decay like that found in the physical sciences. Intuitively, this seems like a plausible notion. Anyone can think of bits of information he or she has learned in the past that can longer be recalled, seemingly because there has not even been occasion to use that information in the intervening period. The information appears to simply be "gone," and appear to have decayed through disuse.

Unfortunately, there are sufficient counterexamples to undermine this straightforward theory. Many experiences are experienced only once, and yet are

remembered for years to come. This is especially true of episodic memories, but can occur with semantic memory as well. For various reasons—mainly due to the fact that not all memories necessarily decay over time—decay theory was rejected early on (Wixted 2004, 237), and scholars largely focused on the notion of interference.

6.2.3.2 Interference Theory

Interference theory argues that what causes forgetting is the occurrence of some other mental activity that interferes with the memory's consolidation or its retrieval. During the many decades that interference was in vogue, two types of interference were postulated: proactive and retroactive. Proactive interference occurs when learning a first group of items interferes with the learning of a second group of items. Retroactive occurs when learning a second group interferes with the consolidation of a previously learned group of items.

Evidence for the existence of interference is found in a number of experimental approaches. In the typical experiment, subjects are given a paired list (A-B), and are instructed to retain the association between the paired items. Then, at some interpolated time, some of the subjects are given a second paired list to learn. A control group will only be given the first list (A-B). Finally, either the first or second list is tested. If learning the first list impairs learning of the second, then proactive interference is thought to have taken place. If learning the second list impairs memory of the second list, then retroactive interference is thought to have taken place.

Strong support for the notion of proactive interference was initially provided in 1957 by Underwood, who discovered that variability in studies on forgetting related to the number of prior lists the subjects had learned. The more sets of items subjects had learned previously, the harder they found it to learn new sets of items (Wixted 2004). It was later discovered, however, that proactive interference was exhibited only when the

study was massed rather than spaced over a period of time.[47] Thus, proactive interference as a factor for everyday memory and forgetting fell into disfavor.

As for retroactive interference, there were two primary avenues of study. Initially, scholars felt that there was a temporal gradient of retroactive interference, meaning that a second list of items would interfere less the older the previous list was. In other words, the longer the first set of items was given a chance to consolidate, the weaker the impact of retroactive interference. However, as Wixted (2004) points out, there were problems with the testing methodology. As early as the 1920s, certain scholars had pointed out that there could be two mechanisms responsible for so-called retrograde interference, the first being that any mental exertion during the interpolated period would cause a decline in memory performance due to its interference with consolidation of the previous material; and the second being that learning of a list immediately prior to recall would interfere with the recall traces (A-C items would cause last-minute confusion with A-C associations). Thus, by exploiting a cue-overload design (A-C and A-B both sharing A as the cue), there would be a mixing of mechanisms. According to Wixted, after separating these two effects, the bottom line appears to be that subsequent mental exertion can interfere with the consolidation process, and cue overload interference will take the form of an inverted U, such that second lists studied right after the first list is learned or right before the list is tested will be more pronounced than second lists learned at some intermediate point, where consolidation was given a chance to proceed and immediate recall interference is not so pronounced.

Although interference appears to be extremely helpful in the context of lexical acquisition in real-world contexts, interest in the topic is said to have waned in the early seventies (Loftus 2007, 323), as the focus of inquiry shifted to inhibition and misinformation (Rubin 2007, 327). However, as Baddeley notes, the sudden lack of

---

47 An easy way to experience this phenomenon of proactive interference with massed "study" is to play one of the many on-line memory games available wherein one the player tries to pick matching pairs of images in a grid. I find that memory performance is acceptable in the beginning, but deteriorates the more I play. This seems to be due not just to fatigue, but due to the fact that memory of image positions from a previous game interfere with memories for positions in the current game. Spacing the "study" by, say, playing only once per day, eliminates this problem.

interest stemmed mainly from the fact that proponents of interference were largely "dustbowl empiricists" whose theoretical framework held little promise. That said, Baddeley himself feels that the effects of interference are powerful and incontrovertible, and that it would be a shame if this fact were ignored (Baddeley 1990, 248).

The fact that proactive interference was rejected seems like an overreaction. It is a fact of life that students often cram for exams, and therefore often perform precisely the kind of massed study that is held responsible for proactive interference. Interference certainly matters in the real world when large sets of vocabulary need to be learned one after another, as is the case in language curricula, and so certainly cases of retroactive interference can also be found in the language student's life.

Interference and the notion of overload are closely related, and while interference should not be construed to be the only cause of forgetting, it is certainly a possible source of retrieval failure. Just as the empirical effects of interference are well-documented and incontrovertible, so too is the fact that memories require time to become less fragile. It seems clear that interference and lack of consolidation are closely interrelated.

6.3 Retrieval

The third principal component of memory is *retrieval*, which Roediger defines as *to access a residue of past experience and (in some cases) convert it to conscious experience* (Spear 2007, 216). Behaviorally, testing retrieval is the only way prove that persistence has occurred, as there is no way to probe the mind and find proof of persisted memories that cannot be accessed. As noted earlier, by positing a distinction between persistence and retrieval, researchers assert that there is a distinction between available and accessible memory traces. In theory, a trace may exist, but for whatever reasons, at a certain point in time, a cue may fail to elicit the desired retrieval mechanism.

Cognitive scientists note that it is difficult to fully distinguish retrieval from encoding, because the act of retrieving an engram is said to reactivate it, causing it to be re-encoded (McDermott 2007, 226). In some cases, reactivating a memory trace can make it unstable, making it susceptible to alteration, usually in the direction of strengthening (Spear 2007, 217).

6.3.1 Vector Nature of Chinese Lexical Retrieval

It is tempting to think of the memory for a Chinese character or Chinese word as constituting a "thing," a memory trace in storage that can be retrieved. In fact, the intuitively sensible notion of learning a character actually entails a complex construct involving multiple stimulus-response vectors. Given that a character is defined as the intersection of three linguistic components—namely, a sound (P), meaning (S), and graphical (G) representation—and that the relations are bidirectional, one obtains six vectors: G→P, P→G, S→P, P→S, S→G, and G→S. These vectors are independent: it is possible to retrieve information easily in one direction (e.g. from graph to meaning) but struggle to retrieve it in the other direction (e.g. from meaning to graph). Thus, learning characters and words must be conceptualized as a multifaceted process, one containing multiple input modalities. Although this is true of encoding, it is most obviously observed in retrieval.

Vectors are not equal in specificity. For example, in Mandarin, any given graphical representation will typically correspond to somewhere between one and two readings and three or four meanings. But any given syllable in the language can correlate with dozens of characters. Therefore, presentation of a stimulus (cue) in one direction may lead to more or fewer candidate responses (targets) depending on the vector path invoked.

All the debate about lexical access paths notwithstanding, in reality, learners of Chinese may at times experience multiple-modality dependency in word recognition, meaning that they require more than one modality for lexical access. An example would be a non-native speaker who has not yet attained superior competence. If given merely the sound of the word, nothing may be activated in his or her mind due to lack of a sufficiently strong vector from P to S. But if that person sees *and* hears the word, access might be successful. If this were the person's experience with much of the material at hand, then that material would be suitable for reading, but be too difficult for listening practice.

6.3.2 Lexical Retrieval in Reading

How will lexical retrieval transpire during reading of Chinese by the learner? Judging from the above discussion, it clearly depends partly on the manner in which the encoding was performed, whether the initially encoded material persisted, and whether the contextual triggers provided are sufficient to elicit a retrieval response.

If a cue succeeds in activating some response in long-term memory, then scientists believe it can set off a chain of activations of related links—a process known as spreading activation. Presumably, it is this fact that accounts for why it is that more elaborately encoded information is better retrieved: there are more ways to arrive at a given node. Spreading activation also accounts for the phenomenon known as priming, whereby response speeds to cues for related information are faster when a priming target has been activated.

Recall that character data is vector-like, containing both strength and direction between the cue and the target. Recall also that the network that is established may depend on how the presenting material is organized, whether based on semantic themes (Figure 6.1) or orthographic principles (Figure 6.2). The so-called spreading activation phenomenon by which related information is activated will therefore have different effects depending on the organizing principle of the lexical network that was encoded and persisted.

6.4 Forgetting

Though it may seem redundant to look at both memory and its approximate opposite, forgetting, it turns out that looking at failure to recall is extremely helpful for addressing lexical acquisition issues. This section looks at postulated rates of forgetting and what light they shed on the lexical acquisition process.

6.4.1 Forgetting Curve Theory

Forgetting curves were first postulated over a century by Ebbinghaus, who noted that forgetting followed an exponential curve (Ebbinghaus,1885), wherein forgetting was rapid in the beginning, and gradually slowed over time for the surviving items. Various

formulas are suggested to describe the phenomenon. In a formula adapted from Loftus, it is described as follows (Loftus 1985):

Eqt. 1 $R = I * e^{-t}$

Here, *R* is the fraction (between 0 - 1) remembered at a given time, *I* is the initial fraction of items remembered, and *t* is time. Needless to say, the exact form of the parameter will be subject to various features. One area of contention is whether the rate of decay is the same regardless of the amount of initial learning or the strength of initial learning (Loftus 1985, 308). The idea is that if if the initial learning strength is greater, then attrition or forgetting will proceed more slowly. Graphically, this is represented as a flattening of the curve, achieved by dividing the exponent -*t* by some value. This alternative form is postulated as follows:

Eqt. 2 $R = I * e^{-t/s}$

where *R* is retention, *t* is time, and *s* is the strength of the memory. In this case, the function of s is to flatten out the curve (slow decay) if s (strength) is high.[48] Figure 6.4 compares the two equations.

Numerous mathematical expressions have been proposed to achieve a closer fit,[49] though the basic principle is the same: the rate of forgetting is initially high but tapers off over time. A corollary to this observation is that memories that survive are more likely to continue to survive as time goes on.

The description of forgetting as a function of time as shown above and in other similar formulations make assumptions that are often not true in the context of learning words. First, these curves express forgetting for the *rote* memorization of meaningless items that are assumed to be homogeneous and roughly equally difficult. As shall be

---

48 When I speak of "decay" here I am not siding with decay theory as opposed to other possible mechanisms of forgetting. Here, decay simply refers to the apparent loss of retrievability, regardless of the underlying mechanism.

49 For additional possibilities, see Wixted 2004, 243.

discussed at length in Section 7.4, this is not at all the case for lexical acquisition. The interrelations and associations of lexical data create imbalances in learning difficulty.

Second, different types of material are much easier than others. For example, my own experiments strongly suggest that, at least for novices, semantic associations with characters will last far longer than phonetic associations with characters. Teach beginners the meanings of a few characters, and they are apt to remember those meanings for several days. Teach them the readings of those characters, and they are apt to forget the majority in under an hour.

Third, these functions assume that all learners perform at the same level, whereas in fact it is well known that in real-life contexts, subjects exhibit different levels of capability, depending on a wide variety of factors. One particularly obvious factor is familiarity with the subject matter. A chemist would typically learn and retain new chemical formulas more easily than someone who saw in them nothing but an abstract series of symbols. Similarly, it has been shown that the alleged "eidetic" quality of chess masters remembering a chessboard is quickly lost if the arrangement of pieces has no strategic significance.[50] This last observation has implications for the learner of Chinese.

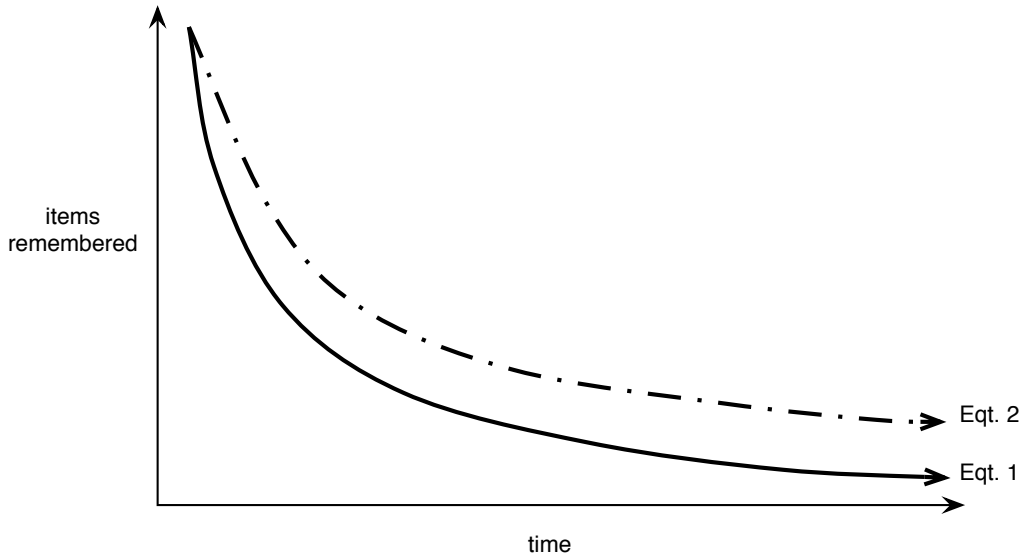


Figure 6.3 Approximate shape of forgetting curves

50 Components of memorability are discussed at length in the next chapter.

If one likens the configuration of graphemic constituents to pieces on a chess board, or symbols in a chemical formula, then familiarity with the individual components and their relative spatial orientations is liable to reduce the tendency to forget.

Now, it is important to keep in mind that forgetting curves portray the fraction of items retained at a given point in time. In a sense, then, if one were to depict a chart showing various items as either retrievable or not, one can postulate a threshold above which items are retrieved, and below which they are not. Imagine a case where six of ten items (numbered 1 - 10) are recalled, as in Figure 6.4.

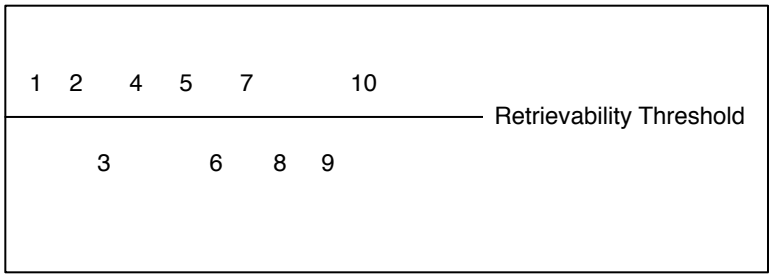


Figure 6.4 Forgetting curve as statistical measure of retrievability

In the chart, items 1, 2, 4, 5, 7, and 10 and recalled, and items 3, 6, 8, and 9 are not.[51] Forgetting curves measure the fraction of items above or below the threshold of retrievability (or recognition) over time. They do not, however, say anything about the inherent memory persistence or endurance of individual items in the list, despite suggestions to the contrary when proponents attempt to place 1/s in the exponent, or when they label "strength" as a "probability of recall." To put it another way, there is simply no way of knowing which of the items 1, 2, 4, 5, 7, and 10 will next fall below the retrievability threshold as more time elapses. Statistically, one can only expect some form of decay—whether according to an exponential equation (Eqts. 1 and 2) or some other function.

* * * * *

51 Note that the retrievability threshold should be taken with a grain of salt in the sense that it is possible to have retrieval failure on one occasion with subsequent success, despite the lack of ostensible reinforcement of the item. Strictly speaking, the threshold represents a line below which retrieval is highly improbable.

Drawing from cognitive literature on memory, one can see that there are different areas of strategic significance to the Chinese language instructor. For example, it is possible to utilize organizing principles that enhance encoding. As seen in Figure 6.1, a thematic "contextualized' approach works well in in the context of developing naturally organized instructional materials. Figure 6.2, by contrast, shows an approach that should work better for orthographic acquisition. The manner in which presenting materials are organized—whether by thematic or orthographic motivations—may well affect the lexical networks persisted in memory. Finally, the three-part vector nature of Chinese lexical data suggests that, for any given item, differing retrieval strengths will exist depending on the modalities of the cue and target. This feature of the orthography also interacts with the above-described lexical networks in complex and unpredictable ways.

As implied by the principles of consolidation and interference, instructors must take care not to produce lexical overload. Surprisingly, the notion of consolidating lexical items—a fairly self-evident goal for language instructors—does not seem to correspond exactly to the notion of systematic consolidation as described above. Whereas systematic consolidation is allegedly independent of any form of review, practical language instruction requires review and study to help consolidate vocabulary. As with interference, macro-level empirical observations such as the decay of vocabulary retention or the interference experienced in intensive courses are not fully understood at the underlying neurophysiological level. But if the calculations of target vocabulary made in Chapter 1 are any indication, to attain anything near advanced proficiency within a four-year curriculum, the student will want to learn something like 3.5 characters and 12.5 words per day. While that may not sound like much, given the rate of degradation expected based on forgetting curves, and compounding that with the issue of proactive and retroactive interference, each and every set of characters and words will need numerous reviews. While the available research on interference might suggest that one should "do nothing" for a while to allow learned materials time to consolidate, naturally that is not a practical approach. Measures are clearly needed to deepen encoding, facilitate consolidation, and minimize interference when enabling the curriculum to continue on pace. The fact that the mechanisms underlying the processes have uncertain theoretical underpinnings is secondary.

One important question for SLS researchers is just how long it takes to achieve real-world lexical consolidation .This is an area requiring much more research under more realistic settings than those typically employed in cognitive psychology. The next chapter explores issues to consider when testing interventions in this area.

Chapter 7
Intervention and Assessment

This chapter seeks to establish a psychometric model that accounts for multiple facets of lexical acquisition. To this end, this chapter identifies factors to consider during experimentation, groups method interventions into amplifying and attenuating categories, identifies measurable properties of recall, discusses variables that constitute properties of the target item itself, and places the roles of the previously explored concepts of automaticity and longevity in perspective. The eventual goal is to point the way to an approach to assessing the relative efficacy of different interventions, and to find ways to mitigate contributors to forgetting identified in the previous chapter.

7.1 Contextualizing Experimentation

To combat the tendency to ignore the fact that control variables greatly influence the results of the experiment and therefore color the various propositions that might be made in conclusion, Jenkins proposed a "tetrahedral model" of memory experimentation. Roediger's slight revised version of the Jenkins model groups the variables into four major categories, as shown below (Roediger 2008, 229).

Table 7.1 Jenkins-Roediger tetrahedral model of control variables affecting experimentation outcomes

Encoding	context, settings, instructions, activities, strategies
Events	pictures, words, sounds, sentences, videos....
Retrieval	free recall, cued recall, recognition, fragment completion ...
Subjects	age, abilities, knowledge, expertise, traits, disorders ....

As Roediger explains, differences in any of these four greater parameter categories can greatly impact the results of any given experiment. Put another way, every experiment is sensitive to the context (set of control variables) in which it was

conducted. The value of the model is clear: it serves to warn against rushes to judgment in asserting that a particular action or approach constitutes a law of memory or learning. And using this model, Roediger goes on to refute a handful of "laws" that had been proposed in the past.

Although this model is an excellent starting point, I believe it can be refined. First, the term encoding subtly gives the false impression that the initial learning of the target material is necessarily a one-time event. It could be, of course, but in most contexts, the subject is most likely given time to rehearse within the so-called "first exposure" period. To emphasize this point, the first parameter will be renamed to "initial encoding". By this is meant the initial approach to encoding the information until the learning session stops, regardless of how many repetitions were permitted during that session.

Second, the original model makes no mention of various post-initial encoding events that might occur, but in practice, it is quite common to provide opportunities for subsequent review or testing and retesting of the material. A variety of tactics are found to assist learning—including repeated study, intermediate testing, review, generation, and other forms of subsequent exposure. These essentially constitute enhancements to the initial encoding experience, and are therefore a second attempt to provide learning. Given that these are clearly separate from the initial encoding session, I propose a new category called "reinforcement".

Third, "retrieval" is a logical category, but the title itself is vague. It should be emphasized that this category refers to the manner of retrieval, and not other factors like time of retrieval. In most experiments, this manner of retrieval will also reflect the manner of evaluation. To prevent any misunderstanding, this category is renamed to "retrieval mode."

Fourth, "events" pertains to the materials learned and their manner of presentation, including instructions given during the experiment. As such, this term mixes and matches things that are of a different nature, and should be separated into two categories: experimental procedure and exposure/target medium. Experimental procedure first states explicitly what kinds of activities were done, when they were done, what instructions were given, etc. In some cases these may take the form of control

variables, and regardless of whether they do, they should be recognized as important parameters affecting the results. Target material and exposure mode refers to what is learned and how it is presented.

Fifth, there is no mention of a timeframe to be considered. In the above table, even if one presumes that Jenkins is referring to long-term memory, then what length of period is really involved? While it is fairly obvious that most experimentation in language learning would concern itself with long-term retention, to be meaningful, longevity timeframes need to be defined differently than they are in cognitive models, for which more or less anything above a few minutes would be considered "long term." To underscore the fact that various interpretations of the timeframe may well guide the results and thus constitute unstated control variables, I add a category of parameters for timeframe. The issue of defining appropriate timeframes is addressed under Section 7.4.

The remaining category, "subjects", is perfectly fine. Note that by varying target subjects, one can strengthen a finding if it proves that the finding holds across age groups. Taking the above arguments into consideration, the revised experimentation model can be represented as shown in Table 7.2.

Table 7.2 Revised model of memory testing parameters

Learning	Initial Encoding	context, settings, instructions, activities, strategies
	Reinforcement	testing, repetition, study time, generation, re-exposure
Method	Target Material	pictures, words, sounds, sentences, videos....
	Exposure Mode	instructions, manner of presentation
	Subjects	age, abilities, knowledge, expertise, traits, disorders ....
Testing	Retrieval Mode	free recall, cued recall, recognition, fragment completion ...
	Timeframe	short-term, long-term, further specificity of timeframe

In the figures, the darkened horizontal lines in the table subcategorize these parameter sets into those primarily concerned with the *learning*, *method*, and *testing*. Initial encoding and reinforcement are both concerned with help with the learning process. Subjects, target material, and experimental procedure focus on the actual steps

taken when conducting the experiment. Evaluation and timeframe both refer to testing. Timeframe could theoretically be considered part of the experimental procedure, but it is such a key element in the model that I have provided it separately. Also, because evaluation can be made at more than one timeframe (1 week, 2 weeks, 2 months) in the sort of experiments targeting Chinese lexical acquisition, it is better to place it separately.

This revised model can serve as a guide in the selection of control variables, and can also be used to caution against hasty judgments with respect to the implications of experimental results.

7.2 Amplifiers and Attenuators

Because the field of memory is attacked from so many angles, it can be difficult to get a clear view of the forces at play in learning. As seen earlier, cognitive psychologists have examined various approaches to enhancing memory. Neuroscientists have explored various mechanisms for forgetting. Here, I propose that, for the purposes of constructing a psychometric framework, the basic factors affecting learning be categorized into *amplification* and *attenuation*.[52]

7.2.1 Amplifiers

Of the many amplifiers recognized in the literature, many have at one point or another been proposed as embodying a law of memory and/or learning. Roediger argues convincingly that none of these potential amplifiers produce consistently positive results in all contexts, and therefore must always be promoted with qualification with respect to context, target material, or subject (Roediger 2008). Even so, as a rule, they tend to have some positive benefit. Table 7.3 (following page) lists commonly reported amplifiers (Wixted 2004, 239-243).

---

52 It would be tempting to apply the terms facilitation and inhibition, which are commonly used in cognitive psychology. The problem with these terms, and inhibition in particular, is that they are already used in particular contexts that have nothing to do with the macro-level senses of positive and negative reinforcement implied here. Inhibition is often used in the specific sense of functions that prevent neurons from firing and therefore inhibit the recall of events in certain circumstances, and in particular, to attempts to regulate unwanted memory traces (Anderson 2007, 295). By contrast, attenuation as discussed here is the unwanted diminishing of desirable memory traces.

Table 7.3 Memory amplifiers and their effects

Amplifier	Alleged Effect	Actual Effect	Qualifications and Comments
Repetition	Early accounts spoke of memory as being "engraved on some mental substratum" (cf. Wixted)	Some experiments have shown that it doesn't necessarily help.	One study showed that it was the act of organizing target items that mattered, not repetition per se. Rote repetition definitely not helpful for recall, though may be of some help for recognition.
Study Time	Initial proposal of law based on list-learning experiments.	Effect highly dependent not just on duration of study but also its timing.	Studies on distribution of study show that equal amounts of time have different effect depending on the amount of spacing between the study periods.
Spaced Review	Increasingly spaced repetition better than massed repetition.	Usually true, but sometimes it proves not to be the case.	If the testing is shortly after study, massed study is actually better than spaced review.
Generation	Active learning presumably more effective than passive learning.	The quantity of material matters. Large lists can reverse the effects.	Fine within-subjects, but not between-subjects. Also not true for implicit memory.
Imagery	As postulated by the Greeks and Romans, the use of imagery helped with retention.	Works for within-subject experimental design Does not work between subjects.	Cannot really comment without seeing exactly how the experiments were conducted. Not useful for implicit memory tests.
Testing	Robust results showing that testing far more effective than repetition, even if no feedback given for items not recalled.	Typically, testing helps recall, and studies have shown it even generalizes from the laboratory to the classroom. Retrieval difficulty on a first test enhances learning, as does some spacing before the first test is applied (Roediger and Karpicke: 2007).	Testing has been shown to be less helpful than repeated study for very short-term evaluation (e.g. 5 minutes later), but typically does work better for long-term than repeated-study. True-false or multiple choice testing can induce the "negative testing effect" wherein students end up absorbing some of the incorrect answers and adopting them as truth. The delay in the first test should not be too long. Providing feedback on errors further enhances the testing effect.
Orienting Tasks	Analytical processing (graphemic, phonemic, semantic) of the target material enhances recall.	Some forms of processing will improve recognition for some forms of materials.	Many of the orienting tasks were rather artificial from the standpoint of SLS application. Moreover, some of the alleged counter-evidence relied on highly artificial testing procedures that test priming effects and not basic recognition, as if to show that it did not aid with implicit memory, but only explicit memory. For CFL/SLS purposes, this qualification is not a concern. The bottom line, then, is that orienting tasks and analysis should orient the learner to the type of material and information targeted for recall. Irrelevant tasks are unhelpful.

Roediger's assessment of the various alleged laws of memory appeared to have debunking as a primary aim. And so while he succeeded in showing exceptions to a "law of memory" can always be found to a given principle if one modifies one of the control variables, much of the counter-evidence provided constituted facts of experimentation and testing that are of no practical relevance to the context of SLS instruction. For example, the fact that imagery is of no help in enhancing implicit memory is rather irrelevant. As it turns out, implicit memory was frequently cited as "proof" that the amplifier in question did not always achieve a desired outcome. But is this of any relevance to the learner of Chinese? It seems unlikely. Some of the counter-evidence should be seen as a helpful guide on how to not implement an amplifier wrongly. Thus, for example, the fact that testing is unhelpful if performed too soon after the initial learning only encourages instructors to delay testing to a more appropriate interval. The value of all these prior studies, one can only hope, would be to glean which facts are applicable to classroom instruction as well as whether there might be synergistic effects when more than one amplifier is provided.

Table 7.3 frames possible amplifiers in terms of activities., but the manner of presentation of the material can affect the quality or depth of the encoding, and therefore affect amplification.

7.2.2 Attenuators

Just as one can expect certain activities to constitute amplifiers to memory, one should expect the converse to be true as well: that certain attenuators can hamper learning of material. It might be tempting to consider the absence of known amplifiers a kind of attenuating factor, and in some senses that is true. But if the aim is to eventually develop a psychometric model that permits calculations of the effects of interventions, then this is not a good approach. There are also attenuators one intentionally constructs. Cognitive scientists do this all the time when they employ distraction to see how it interferes with learning. This it is of interest to cognitive psychologists conducting experiments, but not to SLS researchers. Finally, poor presentation of materials could also be identified as an attenuating factor, but it too is actually simply the absence of organization, the absence of amplification, not a true attenuator.

Among frequently reported causes of forgetting, Craik lists increased rate of presentation, fatigue, sleep deprivation, aging, intoxication, and certain chemicals (2007, 133). Others would include various forms of amnesia and other mental impairment, poor diet, and distraction or lack of attention at the encoding phase. From the standpoint of the SLS researcher, the attenuating factors worthy of attention are those that can be controlled by the curriculum and teaching method, those that result from unintended consequences, and those that are simply par for the course.

Out of the factors just mentioned, a major attenuating factor under the instructor's control is rate of presentation, and in particular, overload. Given the tendency to want to pack as much lexical data into a curriculum as possible, the potential for retroactive and proactive interference looms large. The relationship between the timing of presentation and interference with consolidation processes is something that needs to be further explored *in realistic settings*, not the laboratory. Whatever the results of such research may be, it is clear that the presentation of a first set of data may interfere with and thus attenuate memory for a second set, and vice versa.

The other significant factor is degradation of the memory over time in the absence of further amplification (Wixted 2007, 335). Although neuroscientists have had trouble pinpointing the exact nature of memory trace degradation in neurological terms, there is no question that, whatever the underlying mechanism may be, memories often become harder to retrieve over time if not amplified in some way. As was mentioned earlier, degradation without interference can be roughly calculated for sets of items using forgetting curve formulas like those explored in Chapter 6. An alternative approach is discussed later in this chapter.

7.3 Need for Establishing Metrics for Testing

Years of research on memory have identified a wide variety of important "effects" —such as the spacing effect, repetition effect, mass study effect—on the learning of items under experimental conditions. Earlier, these effects were grouped into two categories, amplification and attenuation, so as to enable to point to away to assess their relative merits. A common approach is to then conclude that X has such and such an effect on Y, where Y is typically either short-term or long-term memory. Unfortunately, such

experimentation performed under ad-hoc experimental designs do not readily translate to the classroom, where far more data must be learned over a much longer period of time.

The following sections therefore seek to establish a framework for performing psychometric testing of the various factors affecting lexical acquisition identified so far. In principle, the first step should be to establish a scale of reference by which—all other things being equal—the various amplifiers and attenuators can be compared. Here, "all other things" refers to the various contextual parameters identified earlier in this chapter (Table 7.2). However, there is one factor that is typically held constant in cognitive experiments on learning, and that is the material itself. Therefore, an important question to ask is whether all lexical items are equal, and how would one differentiate grades of difficulty. This is the subject of the next section.

7.4 Variability in Memorability of the Target Material

Unlike with cognitive experiments where paired items often consist of nonsense syllables or words that are well-known to the subject, Chinese characters and words are liable to vary in the ease by which they are remembered. Thus, it is helpful to define an experimental factor—*memorability*— as the ease with which something is remembered. Memorability can logically be divided into subcomponents:

- the inherent degree of difficulty of an item when perceived in isolation
- background information that facilitates encoding and persistence
- incidents that positively affect retrievability (subsequent amplification)
- negative factors adversely affecting retrievability for the item (subsequent attenuation)

In short, one can expect memorability to have intrinsic, extrinsic, and subsequent exposure components. Let us define these terms precisely.

Intrinsic memorability refers to the inherent memorability of an item when exposed to subjects having identical background knowledge. If a group of subjects is exposed to two items, and are given no special information or advantage for knowing these two items, then the item that, statistically, is remembered best will be the one with

greater intrinsic memorability. As it turns out, it seems some Chinese characters are indeed more memorable than others.[53]

Extrinsic memorability refers to the inherent memorability of item as a function of outside knowledge given about the item. For example, if additional background information about a item is given to enhance the ability to memorize that item, then that enhanced memorability is extrinsic to the item itself. To take a rather simple example, if subjects are told that the phonetic 義 tends to carry the reading *yi* when found in characters as a phonetic, and if the subjects retain that information, then the memorability of character readings for characters containing that phonetic should increase. The principle of enhancing the extrinsic component of memorability is tested in the following chapter for semasiographic characters (Experiment 1) and phonographic compounds (Experiment 2). Technically, extrinsic memorability can be a function of background knowledge provided at the time of initial encoding, or background knowledge already known from prior information acquired by the subject.

Subsequent amplification refers to the timing, frequency, and manner of exposure or other form of amplification of an item subsequent to its initial encoding. Repetition, study duration, frequency of re-exposure, and manner of re-exposure can all provide different forms of amplification. Recalling the vector-like nature of character knowledge, different exposure types will affect the various memory vectors differently. Reading will reinforce recognition qualities, writing by hand will enhance meaning-to-graph vectors, reading out loud would reinforce graph-to-reading vectors, and so forth.

Subsequent attenuation could include such things as the interference brought about by excess study in a short period of time, or irregular or prolonged absences from re-exposure. Mathematically, one would expect different types of attenuation to have different effects on the rate of decay.

What does this mean in the context of reading Chinese as a foreign language? First, certain characters and vocabulary are likely to have greater inherent memorability

---

53 Inherent memorability is essentially the flip-side to Wikelgren's notion of fragility. Stating this factor in positive terms, however, makes it possible to establish a model that tests for the effect of positive interventions on a target. Given that the goal of instructors is to induce positive interventions on lexical acquisition, using a positive frame of reference is far more convenient.

than others. (This is strongly suggested, but not proven, by the experiments described later.) I also believe that memorability for characters and vocabulary can be enhanced by adopting certain practices. The enhancement of data used when teaching character meaning is illustrated in Experiment 1; and the enhancement of data when teaching readings is tested in Experiment 2.

Finally, and rather obviously, subsequent exposure is critical to keeping a memory target alive. Subsequent exposure is in turn a function of the exposure the item receives over time, which suggests that frequency ranking of characters and vocabulary is a critical factor in memorability, common items being frequently revisited through reading of texts, and therefore less likely to suffer fragility than rarer characters and vocabulary. Assuming an average sort of reading material, memorability will be stronger for the more frequent items. The unexpected result from the standpoint of curriculum design is that—provided they are within the targeted vocabulary domain required to attain independent reading capability—somewhat less common characters and words will need more attention than the common ones.

7.5 Assessable Properties of a Target Memory Item

Another step to establishing a psychometrical model for lexical acquisition is to identify properties that form part of the framework. Before offering my own solution I would like to examine candidates in the existing literature.

Wickelgren defines memory in terms of two properties: *fragility* and *strength*. In Wickelgren's model, *fragility* measure's an item's susceptibility to time decay, and *strength* indicates an item's probability of correct recall or recognition at a given point in time. *Fragility* clearly identifies a critical issue for lexical acquisition, as cursory self-reflection should enable anyone to recall cases where certain memories died easily and others persisted without hardly any intervention at all. Figuring out why there is so much variation among items is useful, and so fragility might work, except that it expresses the memory trace property in negative terms.

By contrast, using *strength* as a property is wholly problematic. Whereas Wickelgren defines *strength* as indicating the probability of correct recall or recognition, other scholars sometimes use strength in a manner synonymous with retention span.

Worse yet, the layman's notion of strength might make things even more confusing. To a layman, what would a "strong" memory be like? Would it be enduring, vivid, or rich in associations? The inherent ambiguity of the term makes it liable to misinterpretation.

Even couching strength in probabilistic terms has its issues. If one recalls that forgetting curves by nature predict the percentage of a set of items likely to be recalled, the probability of correct recall is not what researchers typically measure; instead, they measure the fraction of items likely to be recalled. Strength defined in Wickelgren's terms does not outline a quality that has meaning from the perspective of a learner's or reader's objectives. SLS researchers should be interested in strength for individual items, not sets of items.

Another candidate property for memorability is a term that was explored repeatedly in earlier chapters, that is, *automaticity*, the notion being that retrieval of lexical information is instantaneous, spontaneous, and autonomous, or some other such quality. While automaticity can be applied to a variety of bottom-up facets of reading beyond vocabulary, and in that sense may be a helpful concept from the standpoint of foreign language reading acquisition, when it applies to the retrieval of lexical data, it is not so much a quality of memory as a constellation of ideals. The lack of a single agreed upon definition makes it an inappropriate metric for lexical acquisition.

Because of the inherent ambiguity in terms such as strength and automaticity, I would like to propose three measures of memory that are, in principle, assessable for individual items: *retrievability*, *accessibility*, and *longevity*.

The first two terms are straightforward. Retrievability is a binary: it succeeds or it fails here and now in the present time. One was either able to access the target data or one was not.

Now if retrieval succeeds, a speed is associated with that act of retrieval, and as defined here, that speed reflects the item's accessibility.[54] This speed is roughly measurable so long as one recognizes that the speed measured would be a composite of four activities: time to perceive the target, time to search for a match, time to access the data, and time to indicate what one has accessed (see Gathercole and Baddeley 1993).

---

[54] Obviously, I am defining *accessibility* differently than the definition used to contrast persistent (available) and retrievable (accessible) memory traces.

Finally, longevity here refers to the length of time that the target remains retrievable without further amplification.[55] Longevity is essentially the flip-side to Wikelgren's notion of fragility. Stating this factor in positive terms, however, makes it possible to establish a model that tests for the effect of positive interventions on a target. Given that the goal of instructors is to induce positive interventions on lexical acquisition, using a positive frame of reference is far more convenient.

Of the three measurable properties of retrieval, the second two are dependent on the first. This becomes apparent when one reduces retrievability, accessibility, and longevity to binaries (see Table 4.4). Retrievability is measured simply in terms of success and failure, and as such, it constitutes the bottom line for lexical access. If retrieval fails (the first column in Table 4.4), then all other measures fail and are therefore inapplicable (N/A in the table). If retrieval fails, there is no speed because there is no retrieval, and there is no longevity, because the memory is dead and has no future, at least, not until revived with reinforcement of some kind.[56] On the other hand, if there is success in recovering the target lexical data, one can have combinations of speed and

Table 7.4 Measurable properties of a target memory item

	Result	Result	Result	Result	Result
Retrievability	fails	succeeds	succeeds	succeeds	succeeds
Accessibility	N/A	fast	fast	slow	slow
Longevity	N/A	long	short	long	short

longevity. The speed at which it can be recovered now is of interest, as is the future longevity of the memory. From the RFL perspective, then, retrieval success (however slow) is the first and foremost criterion to achieve. After that, one looks for secondary properties such as speed and longevity.

---

55 Note that a performance slip might make such information temporarily unavailable. This, however does not really complicate the model, and merely suggests that the item in question is hovering just above and below the critical retrieval threshold described earlier.

56 Spontaneous recovery of memories is reported in the literature, but from an RFL standpoint, one needs the retrieval to occur during the act of reading, not some hours or days later, when it i's far too late to be of use.

Accessibility relates to one of the characteristics idealized by automaticity advocates: instantaneous retrieval. Naturally, access speed lies along a continuum. It is also likely to relate to amplifiers such as recency (of testing, rehearsal, and the like) as well as frequency of reinforcement. Rapid accessibility is crucial for audition, but somewhat less so for reading, where, excepting special cases like movie subtitles, the reader can afford to pause to recall an item. Even so, most words need to be recalled automatically for fluent reading to transpire.

One should note that RFL researchers' focus on automaticity, and its association of instantaneous response, can tacitly condone unwanted study practices. It is perfectly possible to develop a temporarily fast access speed (as when cramming for certain materials), only to have longevity suffer because of the massed nature of the study. This phenomenon is perhaps all too common in the real world, where students need to pass exams.

Finally, Wickelgren's fragility—the positive correlate of longevity—is essentially a byproduct of the separate components of memorability alluded to earlier. What would cause fragility, and what would counteract it, thereby enhancing longevity? Clearly, the intrinsic and extrinsic factors of memorability, and the amplifying and attenuating factors.

In addition to reducing ambiguity, employing *retrievability*, *accessibility*, and *longevity* in place of previously used terms such as *strength, fragility*, or *automaticity* may help add precision to intervention assessment. These new metrics are target-item-specific, and therefore differ from the statistical assessment of sets of items as found in forgetting curves. As such, these properties provide a means for assessing the impact of encoding strategies and subsequent interventions (amplifications) on memory for a particular item, rather than the statistical aggregate of impacts on a set of items. This ability to identify and assess individual items is useful, because it allows the researcher to modify experimental parameters and tease out their effects more precisely. Whereas the components of memorability essentially define factors that affect encoding and persistence, these three properties are assessed in terms of their predictions or actual results at the retrieval stage.

7.6 Applicability of the Psychometric Model

The issue of measuring memorability in precise terms is beyond the scope of this dissertation, but computerized control of material exposure and tracking of the number and types of exposures over a wide number of subjects would in principle make it possible to identify the intrinsic component of memorability for individual target items. Once that is established, mathematical models could be designed so that amplification would be timed to help keep the less common words remembered. At the same time, as overloading could result from interference when numerous lexical targets are studied day after day, tests would be needed to factor in interference as a function of number of words learned per unit of time, and specifically, to determine its attenuating effect over time.

In the previous section, the commonly used notions of memory strength and automaticity were replaced with three measurable properties: retrievability, accessibility, and longevity. Two of the three properties can only be tested "after the fact"—meaning that one can only obtain their values through testing subsequent to previous encoding. Strictly speaking, longevity cannot be tested at all. One can only establish that longevity continues to exist at a given time if retrievability is positive. That said, given sufficient statistical data, it should be possible to come up with expected values for each property for a given target item given the target item's intrinsic memorability, initial encoding, subsequent amplification inputs, and subsequent attenuating factors.

With respect to retrievability, factoring in the nature of exposure at initial encoding, the nature and timing of subsequent amplification, the extent of interference and degradation during the intervals, and the intrinsic memorability of the item itself, then if one sets up a retrievability threshold as a probability of recall, then it should be possible to estimate how likely retrieval is to occur at any given time. A retrievability threshold could be established to define "largely forgotten"—say, $P_{\text{retrievability}} < 1\%$. One could also establish retrievability benchmarks to define "certain to remember"—say, $P_{\text{retrievability}} > 99\%$.

As for longevity, by factoring in the nature of the initial exposure, tallying subsequent amplification incidents, calculating degradation and interference between amplification incidents, and noting the current scaled position above a pre-established

"largely forgotten" retrievability threshold, it should be possible to estimate expected longevity from the current point in time under the assumption that no further amplification is provided.

Finally, while models have not been built to test the rate of decay of accessibility, there is no reason such a model cannot be built through experimentation in the same way it is built for longevity. Upper and lower thresholds corresponding to "usefully fast response" and "hopelessly slow" could be established on a scale measured in seconds. Testing could establish the effect that various amplification and attenuation types have on accessibility. Then, if one could determine the time course of accessibility due to degradation, as well as the effects that interference has on accessibility, a similar curve to that found in forgetting curves might be established. Then, by tallying amplification incidents, taking recency of exposure into consideration and the vector of the exposure, and based on previous estimates of a person's access response speeds, one should be able to estimate accessibility at a given point of time.

Note, one should expect each target to have different retrievability, longevity, and accessibility values for each vector (G => P, G => S, S/P => G). More intriguing will be discovery of the way that each form amplification contributes differently to each vector. This too is beyond the scope of this dissertation.

7.7 Expected Effects of Amplifiers on Memory Properties

With an informal but pragmatic model of memory for testing lexical acquisition in place, it is now possible to hypothesize about the effects of various types of amplification on both a statistical level (how the memorability curve will be shaped) and in terms of the properties of individual target items. The table below shows some of the possibilities, with the check marks representing my current "best guess" as to the expected effect of the amplification technique in question. The list of possible amplifiers here is rather general, and distinguishes some activities in terms of whether they are massed (repeatedly performed under a short period of time) or spaced (performed over increasingly longer intervals).

Table 7.5 Expected effects of common amplifiers on memory properties

Amplifier	Enhances Retrieval	Enhances Speed	Enhances Longevity
massed study	√	√	
spaced review	√		√
massed testing	√	√	
spaced testing	√		√
visual analysis	√		√
generative use review	√	√	√
passive review	√		√

The selection of these three properties as proposed earlier might strike some as illogically mixing levels of analysis. After all, if one succeeds in retrieving something, there will be some speed to it (say on the order of a few milliseconds up to two or three seconds). And if retrieval fails, then it's longevity could say to have fallen to 0, and its speed (measured in time duration) as having blown up to infinity. But I believe that this strange mix better represents actual RFL goals, and that the mix is not as illogical as their mathematical plots might seem. Thus the terms warrant clarification.

Retrieval measures the ability to retrieve an item at any point in time. It is a pure binary: one fails or succeeds. As the table above shows, all amplification activities are expected to improve retrieval at some point in time, even if it is only immediately after the amplification. Thus, whereas passive review may do some good to enhancing retrieval and may slightly enhance longevity, it is unlikely do much for speed. On the other hand, massed study will greatly increase speed for a period of time, but may not guarantee longevity. And visual analysis (as one might do in analyzing the components of a character) might increase longevity while not doing anything for access speed. Thus, pretty much any activity will temporarily help retrieval (= short-term longevity), but many will not significantly extend longevity (some target period of time that is significant from an RFL standpoint), and some contribute little to accessibility

One might ask: why is retrieval worth bothering with at all? The answer is that it is the fundamental goal. If after three months of a summer break, even if the student is

"rusty" he or she can more or less pick up with 95% - 98% of the same vocabulary level, then he or she should be capable of reading at more or less the same level as before, even if the reading is initially a bit slower. If 25% - 40% of the words are lost, then reading is no longer practical, as too much time will be spent looking up words. The disincentive to reading becomes too high. Delays (lack of "automaticity") are less a concern than basic retrieval.

In real-time audition, failure to quickly retrieve will result in much meaning passing the listener by. By contrast, with reading, some lag is acceptable so long as the lags only occur for less-frequent words. Therefore, to encourage Chinese RFL, retrievability is the primary goal, with rapid access being a secondary goal. So long as the retrievability persists (longevity is attained), some decay in speed for less common items is satisfactory. Thus, for example, if a learner must pause slightly to recall the sound or meaning of a character or word, that is certainly far preferable than having no recall and having to resort to a dictionary or electronic search tool.

Note also that by defining the general term amplification for any attempt to enhance memorability, it becomes possible to expand one's field of inquiry to far more precise cases then the rather generic terms used in the literature of experimental psychology. For example, in Chapter 3, I proposed a set of subcomponents of lexical acquisition. Now, with the current model, I could specifically test, for instance, what effect rehearsal or testing of connections has on speed and longevity; or, whether discrimination activities enhance longevity. Such experimentation will be far more restrictive than the type of lab work expected in experimental psychology, but should bring far more useful answers for those concerned with Chinese lexical acquisition.

Placing the new model in perspective enables users to assess the effect semantic or phonetic analysis has on speed or longevity. This sort of information can alternately fall under "visual analysis" or "componential analysis," but typically would be identified as either semantic or phonetic processing, in other words, an "orientation task."

7.8 Effect of Amplifiers on Longevity and Accessibility Curves

Forgetting curve theory essentially postulates a rate of decay that can be characterized by a logarithmic, exponential, or power function. Pure forgetting curves

assume no reinforcement—reinforcement being one of the parameters stipulated for my proposed revised model of contextualizing parameters. In real life, and especially in the context of studying a language, there is reinforcement. Just as obviously, there are also factors that hamper memory.

Now forgetting curves are typically postulated under the assumption that no form of reinforcement (amplification) has taken place. They also describe the statistical decay of a set of items learned at an initial encoding incident. A key question is, What happens to the forgetting curve when amplification is provided?

To begin with, in the framework proposed here, the emphasis should not be on statistical averages provided by forgetting curves, but on target-item-specific longevity and accessibility curves. Without further research, it is not possible to know how these curves would look, but as a first approximation, one could postulate that they behave similarly to a forgetting curve with reinforcement. In other words, each amplification incident would result in a spike in the dependent variable at the point in time when the amplification incident occurred, and the extent of the spike would reflect the nature of the amplification incident and its relative effect on the dependent variable in question (that is, longevity or accessibility). (Table 4.5 provided a list of expected effects in this regard.) Also, as is clearly implied by the experiments described in the next chapter, the rate of decay, and not just the amplitude of the spike, would be altered depending on the type of amplification provided.

Figure 7.1 shows a crude diagram of how the time course of effective memorability might look like for a single target item. As can be seen from the figure, as the target is amplified over time, subsequent decay curves should grow flatter. Moreover, the extent of amplification differs depending on the type of amplification. L1, L2, and L3 show the expected longevity following the first, second, and third amplification incidents, respectively. As indicated, longevity is essentially the amount of time required to decay to the retrievability threshold if the item were neglected. A fully consolidated memory would have a theoretical value of infinity (though inconveniences like death might intervene!).

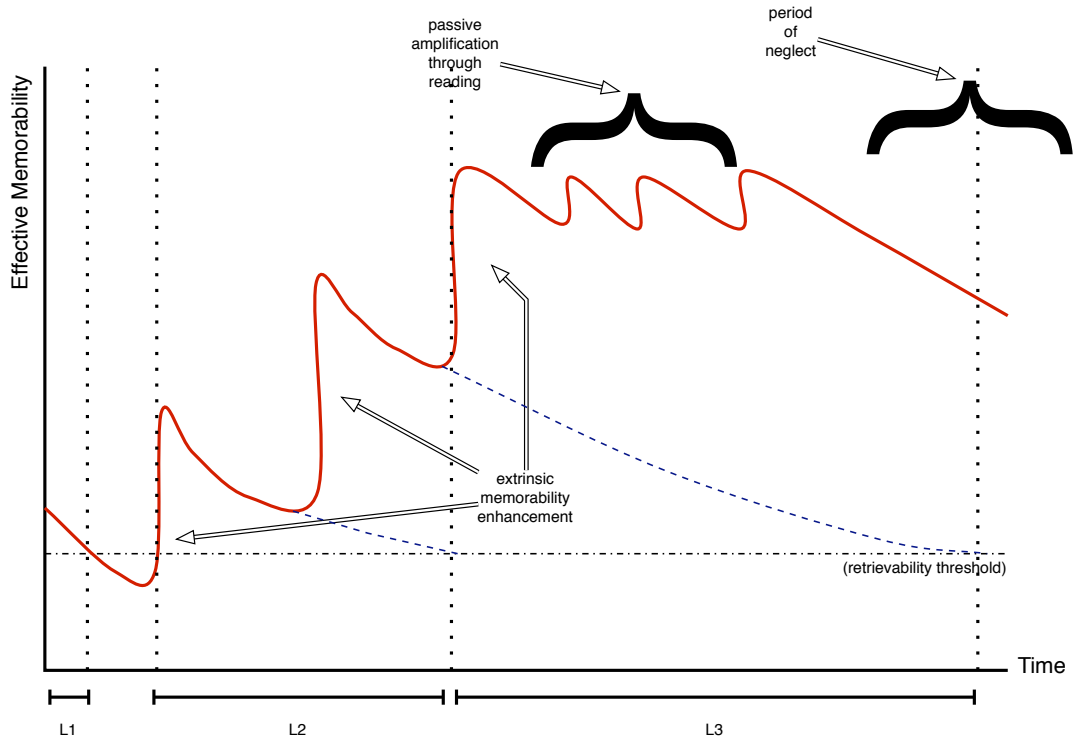


Figure 7.1 Longevity and effective memorability

In the previous section, I hypothesized that different types of amplifiers would affect memory properties differently, with some lending more weight to access speed, and others doing more to enhance longevity (they all help with immediate retrieval by virtue of the recency effect). If one were to device a visual conceptualization for accessibility, one should expect that as a rule amplifications increase accessibility, with the understanding that certain amplifications will have a far greater effect than others. Little is known about the time course of accessibility as a function of amplification. But in principle, by plotting 1 / t (time to respond) for the y-axis, one should be able to construct similar charts plotting that time course for a given item subjected to various grades of amplification. It is also possible that when working to improve automaticity, it may help to employ a rehearsal "package" that continues until a target response time is achieved.

7.9 Redefining Longevity Timeframes: LTM => ULTM and CLTM

A final step toward improving the framework available for experimentation and intervention assessments is the issue of defining adequate timeframes. The need to revise longevity timeframes becomes obvious when, for instance, one examines typical research results in the field. In an article on the relative effects of spaced and expanded retrieval practice, Karpicke and Roediger concluded that spaced retrieval practice improved long-term recall, and expanded retrieval practice improved short-term recall (Karpicke and Roediger 2007, 704). The timeframes in question? Short-term was a few minutes later; long-term, two days later.

Since a duration of two or three days, let alone ten minutes, cannot possibly be considered meaningful from the standpoint of language acquisition, it is important to redefine longevity in more meaningful terms. How would one go about this?

As a basic starting point, one can *in principle* distinguish between unconsolidated long-term memory (ULTM) and consolidated long-term memory (CLTM). According to the traditional view of systemic consolidation, unconsolidated memories are dependent on the hippocampus whereas, thanks to a process of reintegration, consolidated memories do not. There is therefore presumably a neurological basis for lasting memories, even if the exact mechanisms are in dispute.[57]

Unfortunately, today's science has not quite yet caught up to the everyday needs of the SLS researcher, as the actual time/effort needed to reach this alleged consolidated state has not been elucidated. As should be obvious from the various topics in this chapter, there is not going to be an easy answer—like "fifty reviews" or "six months continuously above the retrievability threshold." The time to consolidation will clearly depend on the nature and frequency of amplification, the extent of attenuating factors, intrinsic memorability, extrinsic memorability (and the extent to which it is altered by the provision of enhanced data), and perhaps other factors yet to be identified.

That said, for teachers aiming to get students to achieve a critical vocabulary threshold, at the very least, they should at least understand that a goal of lexical

---

57 There is also a theory that that this view applies only to semantic memory and not episodic memory (Nadel 2007, 178-179).

acquisition is the consolidation of what is learned, and not simply exposure to a wide variety of materials and vocabulary items. Time will tell whether this pragmatic usage of the term consolidation ends up corresponding to underlying neurological mechanisms.

Chapter 8
Sample Interventions:
Altering Extrinsic Memorability

The identification of potential targets for intervention discussed in Chapter 3 and the preliminary framework in Chapter 7 point to a large number of avenues for experimentation. This chapter shows but two of those possibilities. Further avenues of research are discussed in Chapter 9.

8.1 Effect of Understanding Compositional Rationale on Character Meaning Recall

A preliminary experiment was conducted to test the effect that an understanding of compositional rationale might have on character meaning recall.

8.1.1 Method

Subjects ranging in age from 9 to 71 and having no experience with East Asian languages were presented two sets of Chinese characters, each set having five characters. One set contained raw data, and the other contained enhanced data. Raw data consisted of the Chinese character and its corresponding English meaning, with no additional information. Enhanced data consisted of raw data but additionally included a description of the character's compositional rationale.

All characters whose compositional rationale was explained were semantographs (their constituent elements were employed to convey meaning). Moreover, all characters had a largely "vertical" construction, meaning that, unlike many phonetic compounds (characters containing a semantic radical and phonetic component with a left-right relationship), these characters did not have obvious divisions. Horizontal divisions between one or two elements were sometimes present, but these would not be apparent to the novice.

Subjects studied the data using a computer program that presented 10 characters (half raw, half enhanced) in a randomly ordered list. Subjects were instructed to click on each item in the list to learn the meaning of the character. Highlighting within the list

indicated which character had been visited. Subjects were told they did not have to memorize the enhanced information about compositional rationale, but they were told to read it to understand why the character meant what it did. Their job was to learn the character's meaning.

Subjects were free to review items more than once. In addition, the amount of time spent examining character information was measured.

When subjects felt they had learned the material, they were given a test within two minutes of completing study. Correct and incorrect answers were noted, as was the time from onset of presentation to the typing of the first character of the response. Subjects were given the impression that there would be a follow-up questionnaire, but there was no mention of a second test.

A second test was given a week later, with testing being conducted as before, only without study time. To eliminate a potential first-last effect, during the initial study session, the order of characters in the list was randomized between subjects. In both the first and second tests, the order of item testing was also randomized.

Note that, to acclimatize subjects to the software used for learning and testing, students engaged in an identical study and testing procedure for dummy data that consisted of alphanumeric characters with obvious or easily remembered meanings assigned (e.g. 2 = two, a = alpha). This also enabled the system to estimate response speed for cases where the response was automatic (e.g. subjects would have no trouble recalling that 2 is "two"), and so response time would roughly gauge their ability to manipulate the keyboard as well as their cognitive response speed.

8.1.2 Hypothesis

Given that enhanced data induces a more complex "understanding" of the components of a character and its graphemic significance, the expectation would be that enhanced data should improve recall. In terms of principles discussed under 6.2, the enhanced data provides both elaboration of the encoding and makes it meaningful. Thus, a one-tailed hypothesis was assumed:

$$\mathbf{H0:}\ \mu_{\mathbf{raw}} = \mu_{\mathbf{enhanced}}$$
$$\mathbf{H1:}\ \mu_{\mathbf{raw}} < \mu_{\mathbf{enhanced}}$$

8.1.3 Analysis

Oddly, there are two separate ways to view this data. In the first, the subjects are viewed as if having been given two treatments: raw and enhanced. By this reasoning, the experiment is a repeated measures design, only the "repeated testing" using identical subjects occurs simultaneously—meaning that each subject engages in the raw-method of study and enhanced-method of study simultaneously. In this case, one must presume that there is no interaction between the raw and enhanced data, meaning specifically that none of the information given through enhanced explanations would help students learn raw-data characters. As it turns out, care was taken to ensure that there was no overlap between the elements found in each set of characters.

8.1.4 Results

The raw data is summarized in the Table 8.1. The results are striking for a number of reasons. First, while the difference in correct recall is not that great immediately after study, one week later, correct recall was more than six times greater for the enhanced data.

Table 8.1 Raw data for Experiment 1

n = 9 (session 1); n = 7 (session 2)	**RAW**	**ENHANCED**	**RATIO**
correct responses 1 (just after study)	5.60	6.80	1.21
correct responses 2 (1 week later)	0.40	2.60	6.50
mean study duration (sec)	14.0	30.2	2.15
SD session 1	1.67	1.48	
SD session 2	0.44	1.95	
variance session 1	2.80	2.20	
variance session 2	.30	3.80	

As the data distribution did not meet the requirements for normalcy required of a repeated measures t-test, a Wilcoxon signed rank test was employed for the analysis. The results are as shown in Table 8.2.

Table 8.2 Statistical Measures: Wilcoxon signed rank / Cohen / r^2

	SESSION 1	SESSION 2
sum of signed ranks W	12	26
standard deviation of W, σW	10	12
significance level z	1.21	2.16
significance level (directional)	not significant	significant ($p < .01$)
Cohen's d	0.47	1.26
size of Cohen's effect	moderate	large
r^2	0.20	0.64
percentage variance from treatment	moderate	large

As can be seen, the Wilcoxon test indicates that the the difference was not significant for session 1 but was significant for session 2. The reason for the failure to indicate significance in session 1 is fairly clear from the calculations: three subjects scored the same with raw and enhanced in session 1, eliminating them from consideration. Of the remaining six subjects, one did better with raw data. With such a small value of n, even one failure to improve with treatment will eliminate significance. By contrast, the significance level was $p < .01$ for session 2, indicating a very strong effect from data enhancement.

This large effect is confirmed for session 2 by the values for Cohen's measure of effect (d), and percentage variance from treatment (r^2). This accords with a previous experiment (Child 2006) wherein logical grouping of related characters and providing compositional rationale data made little difference in recall accuracy immediately after study compared to raw data, but decreased forgetting by a factor of three 24 hours later.

The implications for forgetting curve theory are significant in that, whereas the initial level of learning is almost the same (the ratio of correct responses with enhancement to correct responses with raw data was 1.21 : 1.00), there is a marked difference in recall one week later (the ratio of *enhanced : raw* correct responses rose to 6.50 : 1.00). If forgetting is described using an exponential function, this signifies a change in the rate of forgetting (the factor associated with exponent -t), and not simply a

difference in starting amplitude (defined by the coefficient of e). This seems to contradict the assertion that verbal lists are forgotten at a constant rate despite the degree of initial learning (Slamecka & McElree; see also Baddeley 1990, 254). Since the initial correct response rates were nearly even immediately after study but substantially different a week later, it is clear that the encoding process has altered the forgetting rate.

8.1.5 Correlation Between Study Time and Correct Recall

One obvious consideration is the fact that enhanced data tended to take more time to process. Whereas for raw data the subject needed to simply examine the character's shape and somehow associate that shape with the meaning presented, processing enhanced data required reading a passage of text that explained the possible significance of each character's constituent elements. In short, enhanced data required additional cognitive processing.

Now it is well known that deeper cognitive processing enhances memory strength (Craik and Lockhart, 1972), but given that this additional processing requires an additional amount of study time, some might argue that the effect is simply due to the additional study, and not to the presence of additional information per se. Consider, for instance, the difference in mean study times for raw and enhanced data as shown in Table 8.3.

Table 8.3 Study time: raw vs. enhanced

	Raw Character Mean	Enhanced Character Mean
Study Time	15.6 sec	33.5 sec

To test the proposition that correct responses were simply a by-product of enhanced study time, a regression correlation was applied to the data for characters by comparing the mean study time for each character with the correct responses for that character. The results are as shown Figure 8.1, where the y axis shows study time in seconds, and the x axis shown the number of correct responses for that character (maximum value of 9).

Since additional study time could only expect to increase recall, a one-tailed hypothesis was adopted. As shown in the following table, the correlations are positive—

as expected, increased study tends to increase recall—but the correlations are relatively low.

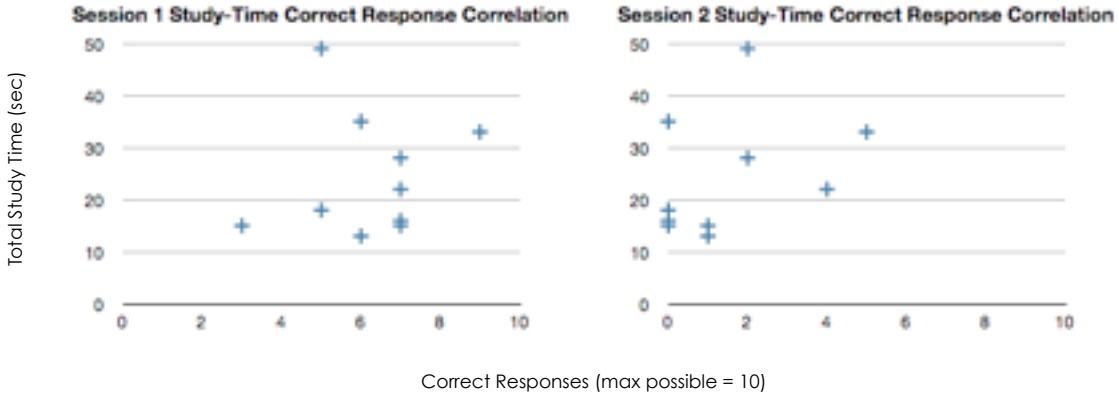


Figure 8.1 Correlation between study time and number of correct responses

To have a significant correlation, the Pearson coefficient with df = 7 (n - 2) should be 0.549. Since both session 1 and session 2 values are well below this level (see Table 8.4), the correlation between study time and correct response is not significant.

Table 8.4 Correlation between study time and mean correct recall

	Correlation
Session 1	0.11
Session 2	0.37

8.1.6 Conclusions

We can infer several things from this initial experiment. First, treatment (providing enhanced data) does not really help that much in the short term. This is probably because, with only a short time lag of a few minutes, rote memory is sufficient to produce adequate results. This is not surprising. The typical approach to passing exams through cramming, only to forget the material a short while later, is supported by various cognitive experiments on massed study, with spaced repetition being identified as generally superior as far back as Ebbinghaus's research a century ago. By contrast, the

enhanced cognitive processing induced by learning about the significance of a character's constituent elements goes a long ways to improving recall in the longer term. This indicates a difference in the rate of forgetting, not the initial amplitude of learning.

Second, this enhanced longevity effect is not simply a product of study time, as correlation between study and recall rates was rather poor. Moreover, from a practical standpoint, simply giving more time to the student to stare at a raw character-meaning pairs is likely to be wasteful: with nothing to process, after a short period of time the student will be ready to move on to the next item. Forcing students to study each type of item the exact same amount of time would therefore be unnatural.

Although a wide range of subject ages was intentionally selected to establish a range of memory capabilities, as it turns out, the effect of enhanced data was almost perfectly uniform. Only one subject did better with raw data, and that was only on the first session; for the second session, he performed better for enhanced data. Thus, while subjects of age 20 - 40 generally performed better than older or younger subjects, the effect of treatment helped everyone.

Third, although I initially intended to test subjects three days after study, my first subject was not available until the seventh day. Since she ended up remembering nearly half of the characters, I concluded that one week was a good outer-limit for testing decay, and used that timeframe for the rest of the subjects.

Fourth, the software was designed to assess not just recall accuracy but recall speed. This was done by averaging recall speed for dummy data (2 = two) requiring no effort of memory for comparison with characters, for which recall would not be automatic. As it turns out, however, many of the youngest and oldest subjects were not good typists, and so their response times were too uneven. I concluded that the recall speed data was unreliable, and discarded it from the analysis.

In future experiments, voice recognition could be used to eliminate typing competence as a factor. However, the problem with using voice recognition to enhance the measure of response speed is that voice recognition software must be trained for each subject, and the training period takes five to ten minutes per subject. This approach could only work for paid volunteers.

Finally, there is an intriguing level of variance in recall rates between characters. Because the study was not designed to test the inherent memorability of characters, such variability was treated as random and uncontrolled. But the discrepancies in recall for certain characters were intriguing nonetheless (9 max. vs. 3 min. in session 1; 5 max. vs. 0 min. in session 2). Clearly, there is some difference in inherent memorability, but it is unclear at this stage what factors might be involved. It would be tempting to suspect a difference, say, between characters portraying concrete meanings vs. abstract meanings, or nouns vs. verbs and adjectives, but the correlation is not apparent from this experiment. Several subjects reported that the description of "legs" in the character for elephant helped sustain a vivid association with the graph, though this advantage might go away if other animal characters were included. Therefore, such idiosyncratic advantages might well be offset once the domain of characters to be learned expanded.

8.2 The Value of Phonetic Components in Learning Character Readings

The ability to recall a character pronunciation in response to the presentation of a graph is one of the key skills demanded of those learning the Chinese writing system. A strong memory in the {graph —> reading} vector is important for two reasons. First, it enables the learner to read the text aloud or subvocalize it. As many believe that word meanings are activated after the reader "hears them" internally through subvocalization, the ability to recall the pronunciation should assist with text comprehension. Secondly, in cases where the learner does not happen to know the meaning of a word, knowing the pronunciations of the word's constituent characters provides the easiest avenue for finding the world's meanings in a computerized or printed word dictionary.

The degree of consistency in pronunciation (reading) of characters within a phonetic series varies considerably owing to assorted historical phonological developments. In his aim to identify pedagogically useful phonetics, Kraemer (1980, 1991) describes seven types of phonetic series. This classification is based on the syllabic components that maintain consistency throughout the series in question, as shown by Table 8.5 below.

Kraemer's earlier work actually sought to identify phonetic character series whose readings were similar enough to make their constituent phonetic of pedagogical

value. Rather obviously, one could posit series types beyond those listed in Table 8.5, including series that do not enjoy the types of consistency described.

Table 8.5 Typology of phonetic series

	Kraemer Phonetic Series Classification	Description
1	totally perfect	All characters have the same pronunciation, including tone.
2	segment perfect	Everything (initial and final) is the same except for the tone, which may differ.
3	initial perfect	The initial is consistent within the series.
4	final perfect	The final is identical for all characters within the series.
5	tone perfect	The tone is identical for all characters.
6	initial and tone perfect	The initial and the tone are identical, but the final may differ.
7	final and tone perfect	The final and tone are identical, but the initial may differ.

Although his typology of phonetic series is logically constructed, Kraemer does not actually test the alleged "pedagogical usefulness" that such consistency affords the learner. As a corollary to the subject examined in section 8.1—namely, the effect of learning compositional rationale on meaning recall—the experiment described in this section examines the degree to which training subjects to know the pronunciation of a phonetic might help with the recall of character readings. In other words, this experiment seeks to answer a simple question: Is learning the pronunciation of Mandarin phonetics helpful in remembering the readings of characters containing that phonetic?

Departing somewhat from Kraemer, this chapter does not concern itself with complex subtypes of phonetic series, and instead defines pedagogical usefulness in terms of two variables: the frequency rank of the target character, and the percentage of characters within a phonetic series conforming to the phonetic independent of tone. This works as follows. The number of characters shared by a phonetic will be a function of the total number of characters examined. For example, the number of characters in domains containing the phonetic 昌 will depend on the maximum frequency rank used as the cutoff for the set of characters under scrutiny. Within the domain employing that

phonetic, those having a consistent reading then indicate the congruence of the phonetic, and that total number of phonetically consistent characters would define the relevant fraction out of the total number of characters obtained at the cutoff frequency. Examples are shown in Table 8.6.

Table 8.6 Pedagogical usefulness of phonetics as a function of reading congruence

Characters	Frequency Rank	Matches	Congruence	Pedagogical Usefulness
唱, 昌	< 2500	2	2/2 = 100%	2/2500 = .00080
倡	< 3500	3	3/3 = 100%	3/3500 = .00086
猖, 娼	< 4500	5	5/5 = 100%	5/4500 = .0011

In this case, two characters (both consistent) are found below 2500, three total are found below 3500, and five are found below 4500.[58] As shown in the table above, in this particular case, the phonetic 昌 becomes slightly more useful as the frequency cutoff point is raised.

Of course, as Kraemer points out, the domain of Chinese characters presents far more complicated cases. Consider 同. This phonetic is usually read as 'tong', but in some cases it has a voiced counterpart, 'dong'. While various options are possible, one could assign half points for a second reading that shows congruence, such that if, say, x characters share reading A (the primary reading) and y characters share reading B (a secondary shared reading), the congruence would be valued as: x + y/2. Similar formulations could be devised for the other phonetic series types.

Now, in light of the lexical acquisition subcomponents identified in Chapter 3, one obvious question is whether explicit training to recognize phonetic elements and their associated common readings can enhance memory for character readings. The experiment described in this chapter was designed with just this purpose in mind.

8.2.1 Method

To test the proposition that explicit training in phonetics might help with reading recall, four phonetic components were selected for study. For each phonetic, a set of four

58 In the table under the "Characters" column, only new additions are shown at any given level.

characters containing that phonetic was selected. A fifth set of four characters having no phonetic in common served as the control.

Within the four phonetic series, there were differing degrees of congruence between the reading of the phonetic and the reading of the character. For the four "character families" with a common phonetic, the reading congruence values were 0%, 25%, 50%, and 75%. In the experiment these four grades of congruence were to be treated as differing degrees of "data enhancement," with the control character set representing absolutely "raw data."

In part 1 of the experiment, subjects ranging in age from 33 to 73 and having no experience with East Asian languages were asked to study four phonetic elements (Figure 8.2) corresponding to the above-described character families and learn their associated readings. Because the subjects were unfamiliar with Chinese, the readings were spelled in an ad hoc "English-style" spelling that would best approximate the actual sound in Chinese if read by a native English speaker with no Mandarin training.

No.		meaning	reading	consistency
0	Φ	null	not applicable	0
3	失	lose	sher	0
4	瓜	melon	gwaah	25
2	同	same	tong	50
1	求	fox tail	cheeoh	75

Figure 8.2 Phonetics used to group characters (φ indicates the group with no shared phonetics)

To further strengthen recall of the phonetics' readings, subjects executed a rehearsal program that presented phonetics on the screen and asked them to type in the corresponding reading. When the subject was able to reach a certain degree of accessibility for a given phonetic, that phonetic was removed from the pool presented for rehearsal. Since the input mode was typing, a skill at which subjects varied significantly, automaticity was necessarily relative, with faster typists being demanded

to respond more quickly, and slower typists more slowly (the requisite response range was approx. 1250 - 4500 ms). Nonetheless, for each subject, it is felt that a modicum of accessibility and memory strength was achieved in a single sitting.

Once the subjects had learned the phonetics sufficiently well, in part 2, they were asked to learn the readings for the characters in the associated character families, as well as those of the four control characters which were not associated with any phonetic. To this end, the families were presented as sets, with the phonetic in the middle and the corresponding characters in four corners. This provided a way for the subjects to easily crosscheck and compare characters that, because they all shared the same phonetic, would otherwise look rather similar to one another. (In the terminology of Chapter 3, the aim was to enhance discrimination.)

Finally, for part 3 of the experiment, a test was administered by presenting individual characters in a random order and asking the subjects to type in the reading. Because it was expected that decay would be rather fast for unfamiliar readings in a foreign language, a follow-up test was not administered.

8.2.2 Test Data

The nature of the phonetic series and the control data is summarized in Table 8.7. Note that out of the total of 20 characters, 16 were associated with a phonetic, and 4 (the control) had no associated phonetic. The total number of characters having a reading exactly matching the phonetic was 6 (see righthand column: 3+2+1).

Table 8.7 Phonetic congruence of test data

Family	Phonetic Congruence
phonetic #1	75% (3/4 characters)
phonetic #2	50% (2/4 characters)
phonetic #3	25% (1/4 characters)
phonetic #4	0% (0/4 characters)
no phonetic	N/A

8.2.3 Hypothesis

It is hard to know *a priori* what results might come of the above-described experiment. On the one hand, because the rehearsal program drilled subjects on the association between a phonetic graph and its reading, one would expect a slight advantage in reading recall for characters having the same reading as the phonetic, or a similar reading. One would therefore perhaps expect that families having greater reading congruence would fare better in terms of recall. In other words, characters with more "pedagogically useful" phonetics should be easier to remember.

On the other hand, because of the rather ambitious learning task involved—twenty total characters in a totally unfamiliar script with unfamiliar pronunciations all had to be learned in one sitting—it could be that the grouping of characters into phonetics would prove overwhelming. In short, a kind of mental breakdown is also a possible outcome, and it could well be that the unaffiliated control characters might be the easiest to remember because there was nothing to confuse them with. Also, in the case of a family with 75% percent reading congruence, if the subject forgets that one shared reading, all three character readings are forgotten simultaneously. This would reduce recall rates for the more congruent families.

Despite the caveat in the last paragraph, I assumed that enhancement would improve recall, and therefore adopted a one-tailed hypothesis for the analysis.

8.2.4 Analytical Approaches

8.2.4.1 Testing the Value of Providing Phonetic Information

In the simplest approach to viewing the experiment, information about the phonetic is considered an enhancement of data regardless of whether the character reading matches its phonetic. In this case, the number of enhanced characters would be 16 (characters in the four phonetic family groups), and the number of raw characters would be 4 (those in the control group). This approach ignores the fact that reading congruence varies by character family from 0% to 75%. From this perspective, data could be analyzed as a repeated samples test where both "treatments" (raw vs. enhanced) are administered simultaneously. The significance of the test results would be assessed using the t-test for normal distributions, and a Wilcoxon signed rank test for non-normal

distributions. However, because of the difference in the total number of raw and enhanced characters (4 versus 16, respectively), instead of applying raw scores, one would have to employ percentages—comparing the percent of raw character readings recalled correctly against the percent of enhanced character readings recalled correctly.

8.2.4.2 Testing the Effect of Reading Congruence within Sets on Reading Recall

A second way to look at the experiment is to view enhancement as being dependent on whether phonetic information is applied *and* the extent to which there is phonetic congruence as being relevant. In other words, the degree of phonetic congruence is treated as a quasi-independent variable. However, unlike case A, each family (level of congruence) is treated as a separate treatment. Just as one might run a scientific experiment at different temperatures, one can view this experiment as testing different degrees of data enhancement, from zero enhancement (the control data) to tepid (a family with incongruent readings) to warm (family with more congruent readings) to hot (a family with mostly congruent or totally congruent readings). In this case, if the test results indicated a normal distribution, one would adopt a single-factor repeated measures ANOVA with four levels for the test statistic. For non-normal distributions, the Friedman test would be appropriate.

These two approaches correspond to different hypotheses or questions one can make about the effect that using phonetics to enhance data might have on reading recall. Consider the differences. Approach A asks: Does training learners to recognize the phonetic and know its reading significantly help character reading recall? Approach B asks: Does the usefulness of providing phonetics lie along a continuum in the sense that the greater the level of congruence in character readings, the greater the effect on reading recall?

Table 8.8 Summary of raw data: percentage of correct responses

no phonetic	has phonetic	phonetic does not match	phonetic matches	total
9%	41%	19%	65%	35%

8.2.5 Results

Basic results from the experiment are summarized in Table 8.8. As one can see from the table, teaching phonetics made it a little more than four times more likely that the character reading would be called (41% vs 9%). Phonetic consistency between the phonetic reading that was learned and the target character showed an even stronger correlation, with 65% of characters whose reading matched that of their phonetic being recalled correctly. Overall, the subjects found this experiment to be extremely demanding: only 35% of all readings were recalled. In light of the dictum that to aid recall the material should be *meaningful* (Section 6.4), this is not surprising. As far as the subjects were concerned, there was no meaningful connection between the graph and its sound, other than the possible consistency with the phonetic, which to them was also arbitrary and only learned by dint of repeated review and testing executed until some degree of accessibility was achieved.

The small number of subjects (n = 17) made it unlikely that a fully normal distribution would be obtained. As seen in the histogram in Figure 8.3 below, certain important characteristics were present: there was a high score and a low score both representing approx. 5% of students, there was a clear mode, and the median and mode were just shy of the mean. The lack of a perfect bell shape, however, and the obvious gap in scores between 7 and 11 forced a non-parametric approach to the analysis. Accordingly, instead of a t-test, a Wilcoxon signed rank test was used for the analysis.

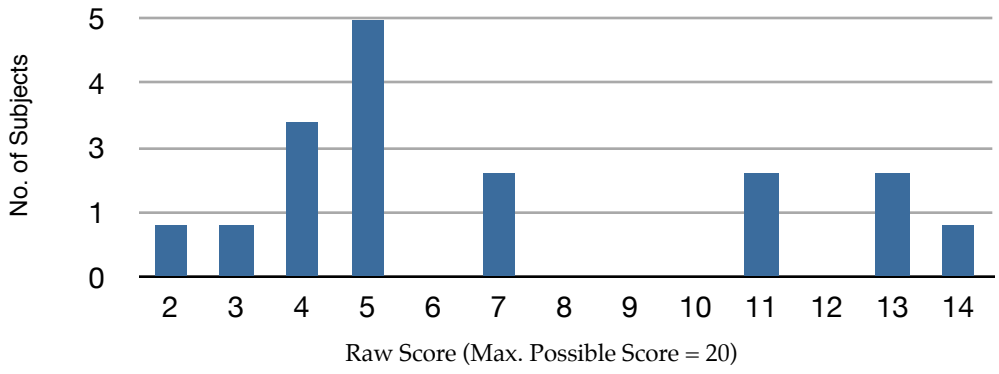


Figure 8.3 Subject scores

The results of the Wilcoxon signed rank test are shown in Table 8.9. The sum of signed ranks, W, is 135. For a one-tailed hypothesis, the corresponding test statistic ($z = 3.18$) indicates an extremely large effect ($p < 0.0007$) at $n = 17$. Cohen's d and r^2 also indicate that data enhancement plays a significant role in reading recall.

Table 8.9 Analysis results for effect of providing phonetic information

Subjects (n)	17
W	135
σ_W	5.57
z	3.18 ($p < 0.0007$)
Cohen's d	1.35 (large effect)
r^2	0.66 (large effect)

8.2.6 Phonetics and Consistency

Beyond seeing whether teaching phonetics is useful, the other obvious line of inquiry is what effect reading consistency within character families has on the learning of character readings. Looking at the basic data, one can observe a largely positive correlation between the amount of phonetic information provided (taking consistency as a factor) and recall rates. Correct answers by character type are shown in Table 8.10 below.

Table 8.10 Summary of raw data: recall as a function of reading consistency

	no phonetics	**0% consistency**	**25% consistency**	**50% consistency**	**75% consistency**	**total**
No. of correct answers	6	15	21	45	31	118
% answered correctly	9%	22%	31%	66%	46%	35%

The first question is whether these differences are significant. In other words, are the sample means significantly different to argue that consistency is a driving factor. As

with the previous analysis, because of the lack of a normal distribution, a nonparametric test was applied, this time a Friedman test statistic

As the table shows, the test results were highly significant (Q = 0.0001), indicating there is a greater than 99% likelihood that differences in sample means for families of different levels of consistency were statistically significant.

Table 8.11 Friedman test calculations

	no phonetics	phonetics00	phonetics25	phonetics50	phonetics75
sum of ranks	72	62.0	51.5	27.0	42.0
(sum or ranks)^2	5184	3844	2652.25	729	1764
No. of columns, k	5				
No. of rows, n	17				
Sum of R^2	14173				
12/nk(k+1)	0.02353				
3n(k+1)	306				
M	27.49				
Q (significance) assuming df 4	0.0001				

8.2.7 Correlation

If one looks at the raw data shown above, one sees a slight anomaly: in the progression from no phonetics to characters with phonetics of different phonetic consistency, the number of correct answers follows a smooth progression except at 50%, which if trends were consistent, would have a value between that for the 25% and 75% character families.

A Pearson correlation and r^2 value were calculated for characters containing a phonetic (see Table 8.12). Although the correlation is not perfect, there is a strong correlation such that as consistency increases, then the number of correct answers does increase. The same markers were also calculated including the control (Table 8.13).

Table 8.12 Correlation between consistency and correct answers (phonetics only)

Pearson correlation	0.71
r^2	0.50
effect size	large

Table 8.13 Correlation between consistency and correct answers (including control)

Pearson correlation	0.84
r^2	0.71
effect size	large

Along a similar vein, one can treat the control group where no phonetic information is provided as being even "less consistent" than 0%, and while there is no certain way to establish a value, I employed -25% for the control data. (Recall from Table 8.8 that even characters whose phonetics did not match were twice as likely to recalled than characters with no phonetic information at all (19% vs 9%), so it makes sense to assume that control data is less helpful than 0% consistency data.) Following this reasoning, if one calculates the Pearson correlation and r^2, one finds similarly significant positive correlations between the provision of phonetic information and the number of correct character readings recalled.

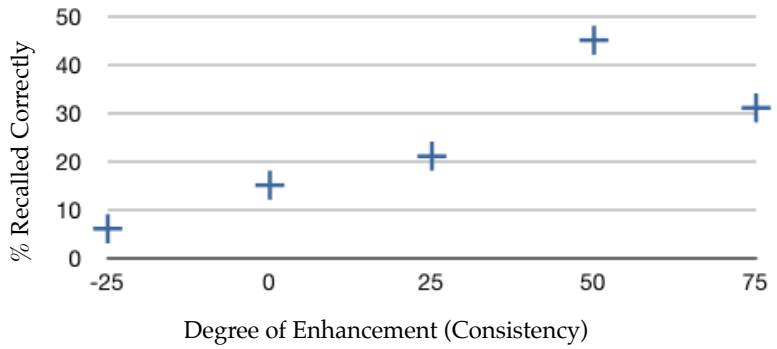


Figure 8.4 Consistency vs. recall.

Note that in the progression ("-25" representing no phonetic information), one sees an almost perfect linear progression except at 50, which is something of an outlier. Although it might be tempting to posit some ideal level of consistency at the midway point, in fact, I believe the answer to this unexpected result is uncovered through qualitative analysis of the data.

As it turns out, several subjects noted that the phonetic representing the group with 50% consistency, 同, actually looks like a pair of tongs. Since that matches the spelling of the phonetic as presented ('tong'), this made characters containing this

phonetic easier to remember. Combined with the fact that graphically 同 was significantly different from the other phonetics, this accidental correlation between appearance and Chinese reading made characters in this series especially memorable. By contrast, some subjects appeared to find 'qiu' 求 (the phonetic for the group with 75% consistency) similar enough to 'shi' 失 (the phonetic for the 0% consistency group) that they confused characters from the two groups. In other words, the addition of radicals distracted them enough to lose sight of the difference in the shape of the phonetics.

Two observations are in order. First, as hypothesized earlier, there is an intrinsic memorability factor at play, one which interferes with the consistency effect to some effect. Secondly, this memorability advantage is extremely idiosyncratic, and could end up changing depending on the total domain of phonetics and characters. For example, if several of the phonetics had been "box-like" (e.g. 回, 田), the advantage enjoyed by 同 may have been erased.

As it turns out, if the value at 50% were in its expected position (roughly 25), then the correlation would be nearly perfect, at 97%. The fact that it proved to be lower really doesn't matter: it is still strong enough to assert that phonetic consistency has a clear effect on recall. Further, when subjects are exposed to a small domain, the most easily discriminated elements will enjoy better recall rates.

Finally, it should be noted that a modicum of training in the concept of voicing alternation might have gone a long ways to enhancing perceptions of phonetic congruence. Such training, however, was felt to be beyond the scope of such a preliminary experiment involving novices.

8.2.8 Qualitative Observations

There are other qualitative observations to be made. First, it was clear that several subjects wanted to know more about the characters they were learning. In fact, it was apparent that they felt disadvantaged not knowing what the non-phonetic component stood for. Take for example the phonetic series represented by the phonetic 'gwa' 瓜, that is, {呱, 狐, 孤, 弧}. With the exception of the first character, which has the radical 口, all the others contain something that look like a curved or squiggly line with or without one or two slashes going through it. Distinguishing such elements is not easy

for a novice, even though such elements would be extremely familiar to a more experienced student. I think it is pretty clear that recall scores would have increased significantly if the subjects had been trained in advance to discriminate such elements. Being able to associate a character by recognizing that it contains 'bow' or 'dog' or 'child' would obviously make things far easier than trying to remember what otherwise appear to be slight variations of the "squiggly line" motif. This accords with the natural tendency to want to encode *meaningful* information.

However, teaching these elements in advance would have no doubt confounded the interpretation of factors affecting recall, begging the question whether they were recalling because of the semantic element or because of the phonetic. It is safe to say, however, that a greater ability to discriminate lookalikes and to associate them with meanings would have been helpful, further underscoring the role that the subcomponents of Chapter 3 play in retention.

8.3 Modeling the Effects of Providing Enhanced Data

In light of the previously discussed notion of lexical networks and the issue of activation of information during decoding, it is possible to hypothesize how it is that providing enhanced data at the level of intracharacter associations helps to improve recall. The following two subsections explore the probable ramifications of the interventions just discussed in 8.1 and 8.2.

8.3.1 The Effect of Compositional Rationale Information on Meaning Retrieval

The experiment of Section 8.1 essentially tested the possibility of using compositional rationale to provide an alternative path to character meaning retrieval. As shown in Figure 8.5 (next page), a rote memorization approach hypothetically leads directly to the meaning. There may, of course, be intermediate steps, but if they exist at all, these steps tend to be unconscious and unknown to the subject.

In contrast, the alternate path is a convoluted affair, wherein one must identify elements, recall associated meanings, and then recall the rhyme and reason for their juxtaposition. If the subject progresses that far, he or she then hopefully arrives at the graph's meaning. The additional complexity of the alternative path, ironically, is

precisely what enhances the retention. It also explains why the retrieval times were significantly slower for memorability items. Equally ironically, this additional complexity is also probably why some students enjoy the rote approach so much. It helps them directly improve accessibility, and recognizing that it is speed that they are going to need when actually reading, for some, this is the practice mode of choice.[59]

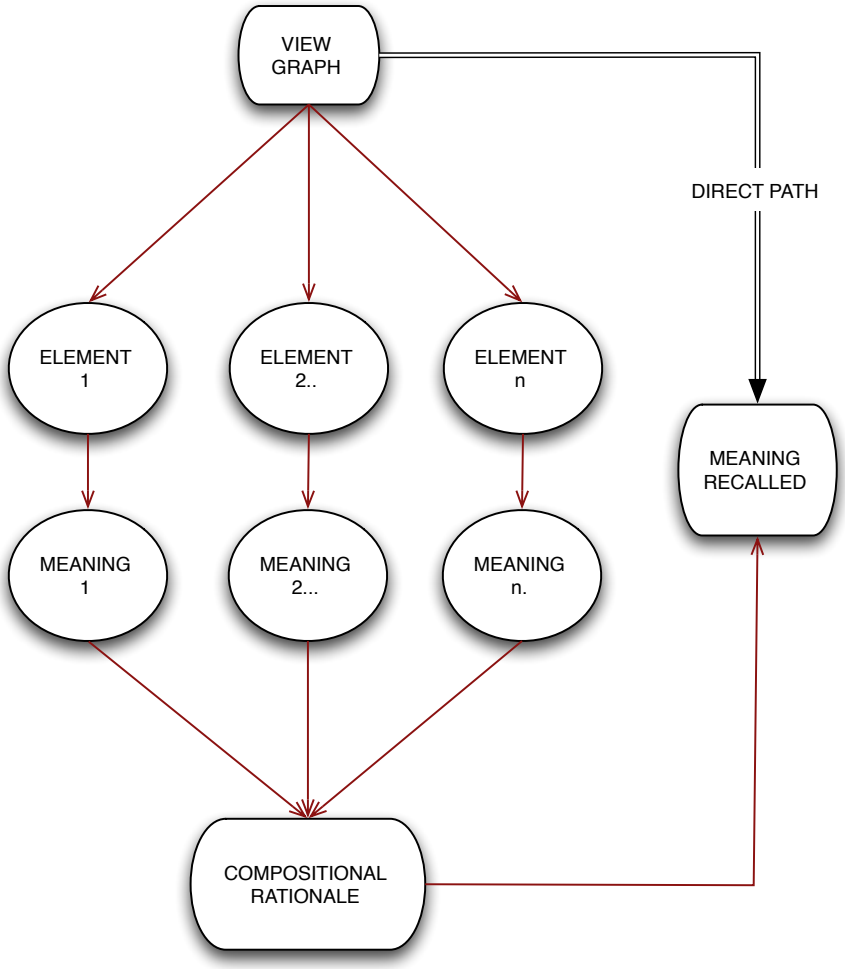


Figure 8.5 Paths from graph to meaning recall: direct and alternative approaches.

The obvious question, at this stage, is whether it is worth the trouble to learn all this additional information. Section 5.3 indicated that researchers suggest that the

59 Some younger students prefer rote memorization to "meaningful study" of compositional rationale (Child 2004).

standard access paths for Chinese graphs may be from graph to meaning (direct) or from graph to reading to meaning (the phonological route). However, considering the common issue of forgetting for characters, let us hypothesize what would happen if a new path existed to arrive at meaning: one that employs the standard (or "direct") path (whatever that is), and one that proceeds from graph to compositional rationale to meaning. On the one hand, the standard, direct, path is easily speeded up through massed study. However, it is also subject to rapid decay. If, on the other hand, a second, more enduring path also existed, then if direct retrieval should fail, it might be possible to retrieve meaning via an alternate path (see Figure 8.6). If the alternative path

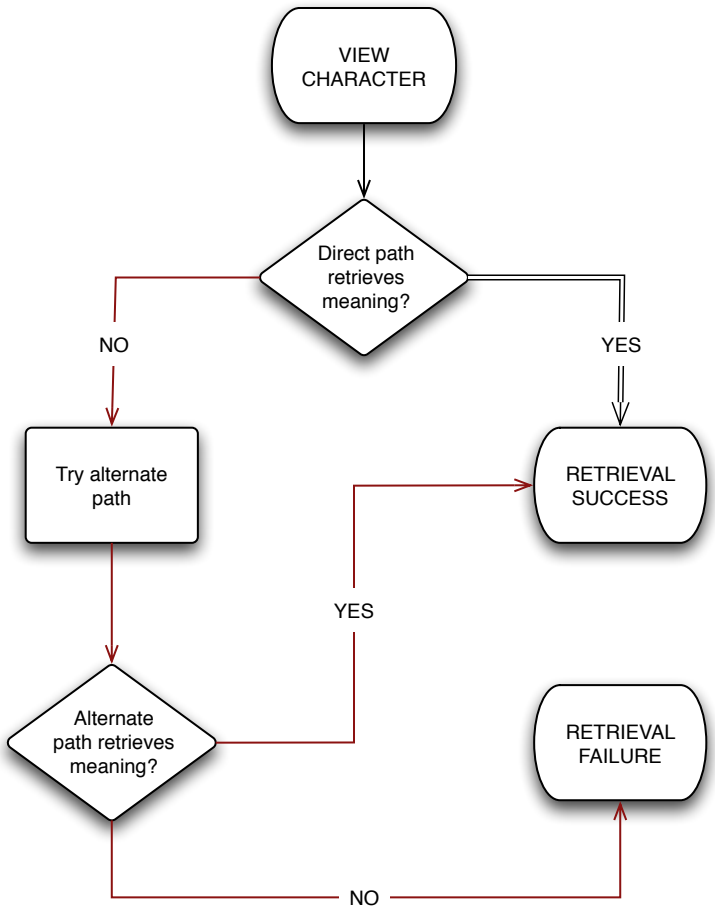


Figure 8.6 Algorithm for character meaning recall when direct and alternative paths exist simultaneously.

succeeded, one could expect the direct path to be restored above the retrievability threshold, at least for some time. Moreover, even though there is perhaps a half-second

time lag when using the alternative path, in pragmatic terms, if that saves going to a dictionary, then that slight additional time is certainly a better alternative to having to look the character up.

There is no way to draw firm conclusions at this stage. Although the results from Section 8.1 were extremely promising, there is simply no way to know what would happen if the procedure were expanded to large numbers of characters. On the one hand, proper learning of character compositional rationales might do much to help students discriminate the many lookalike characters, and to help keep track of their meanings. On the other hand, streams of "stories" about character composition might, if poorly implemented, simply result in overloading the brain, resulting in one of the forms of interference identified in Section 6.5.2. Longitudinal studies using data on a larger scale would be required to tell for sure. What is certain, however, is that compositional rationale has a quite powerful effect when tested in limited quantities. It is therefore worth determining whether this effect can be replicated on a larger scale.

8.3.2 The Effect of Phonetic Element Training on Reading Retrieval

Looking at the flip side of the retrieval process with the graph as stimulus, Section 8.2 tested the effect of phonetic element training on reading retrieval. Many of the processes are similar, though there are slight differences. Only phonetic readings are obtained, as other elements do not have readings. More importantly, after recall of the phonetic reading, there has to be a check to see whether that reading actually matches the character. This can only be done by checking the other elements in combination with the phonetic, and determining whether that combination reads the same as the phonetic or as something else.

Figure 8.7 really only presents the most basic case imaginable. For example, it is quite possible that the phonetic corresponds to two or more readings (there is often a 50-50 split or other fractional split among the characters sharing a phonetic element), and so the path from [phonetic —> phonetic reading] will in practice often be more complex. The subject may recall one or both readings (or none). But if both readings are recalled, then a check of the other elements is needed to perform the critical step of determining the character reading.

As with the case of meaning recall (Figure 8.5), Figure 8.7 is not meant to suggest that the direct recall path (upper right in the figure) is a simple, one-shot affair. It may be, or it may have a number of steps. However, it is quite likely that these steps are not

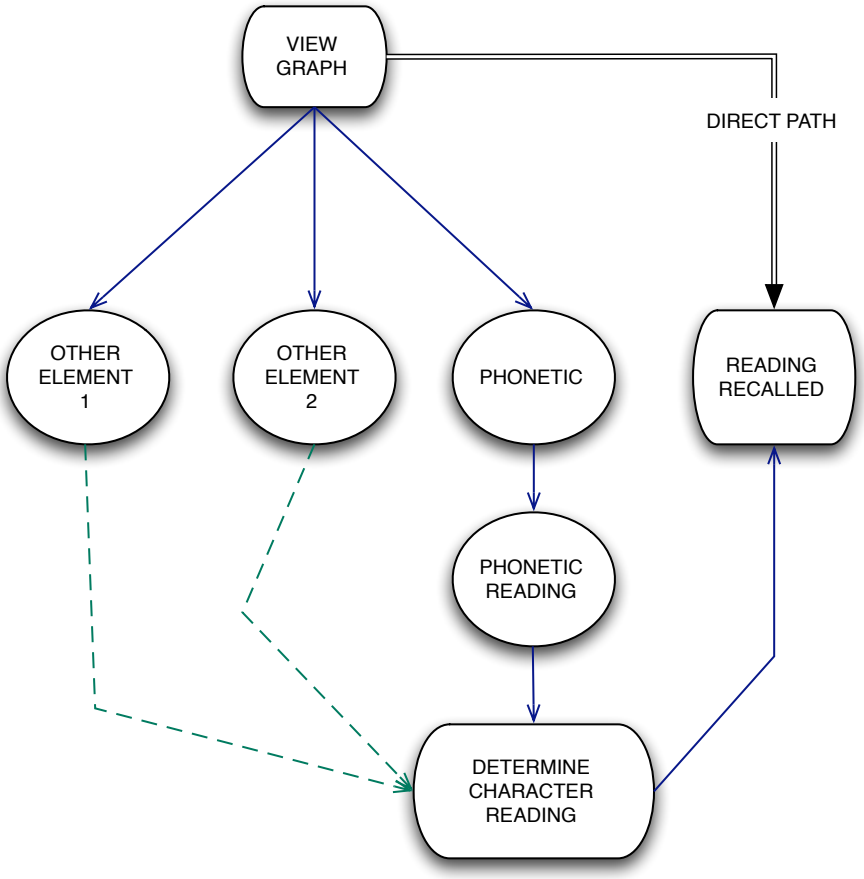


Figure 8.7 Paths from graph to reading recall: direct and alternative approaches.

performed consciously, and because they are not trained based on the provision of information, they are also apt to be highly idiosyncratic.

The algorithm of Figure 8.8 shows whether to employ a direct—or perhaps I should say, idiosyncratic and unconscious—path vs. one that consciously employs the alternative path. As expected, this algorithm is nearly identical to that for meaning (compare Figure 8.6). In essence, one way to look at the experiments of this chapter was that they tested two alternatives: employing an idiosyncratic and possibly unconscious approach to recall vs. one that was made fully conscious and given explicit information,

some of which was quite logical (in the case of compositional rationale), and some of which was utterly random and had to be memorized.

It seems quite likely that explicit compositional rationale work and phonetic information play slightly different roles. Because compositional rationale work is something that can actually be grasped and thought about, it does provide a significant

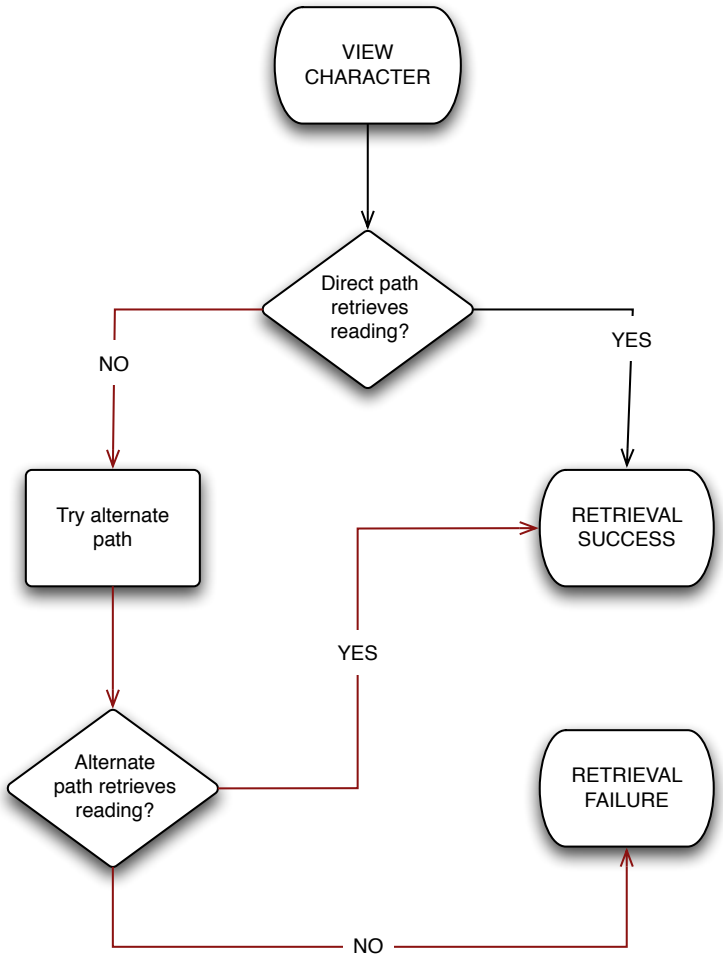


Figure 8.8 Algorithm for recall for character reading when direct and alternative paths exist simultaneously.

stimulative effect, one which may enhance the strength of encoding. By contrast, the association between phonetic element and reading is of course far less memorable, and therefore difficult to recall. This is why repetition had to be used in the experiments to achieve some measure of automaticity. However, training with phonetic elements did help with character part identification, and thus appeared to assist with recall. It is likely

that if the subjects had also known the meanings of the other elements, which were semantic in nature, they would have had a far easier time recalling character readings, as the ability to discriminate between similar characters would be enhanced.

Chapter 9
Implications of the Framework and Avenues for Future Research

9.1 Overall Implications

This dissertation has been written under the working hypothesis that lexical acquisition—and, in particular, issues pertaining specifically to the Chinese orthography—lie at the heart of students' difficulty in attaining superior competence in the language. Using that hypothesis as a starting point, I have sought to identify the features that distinguish Chinese from more easily acquired orthographies, and have then sought to describe the types of information that might be of use in interventions designed to facilitate acquisition.

Lexical acquisition itself is a complex process, and the vector model employed in this dissertation belies this complexity. For example, when describing the notion of a graph to meaning vector, I am obviously pretending that there is one target result (e.g. a meaning) that should be elicited when the stimulus (e.g. a graph) is applied. However, not only is it obvious to anyone that any graph will have a number of meanings, but the search process itself is bound to be complex given the fact that the search may need to be performed recursively until the correct meaning is found given the context. In focusing almost exclusively on the more basic issue of how to better attain longevity for a single item, I do not intend to belittle the importance of such higher-order contextual concerns. I simply mean to emphasize that such higher-order concerns cannot even take place until the foundation of essential vector data is established. Simply put, you cannot choose from among two or three meanings of a word unless you are aware of the individual meanings from which you choose.[60]

Chapter 2 looked at perspectives by which to analyze the orthography, and sought to identify what sort of information may be of use in establishing positive

---

60 That may be a slight overstatement. When facing a context where usage of a target character (or word) is unclear to them, language learners sometimes eliminate the one or two meanings they know as being inappropriate given the context, and either guess based on context or realize that they simply do not understand the word's role in context.

interventions for orthographic acquisition. My conclusions in this regard were as follows:

- Information from historical phonology will be of only indirect interest to the average student. Such information matters to the extent that it helps to identify which elements constitute phonetics, and thus can shed light on compositional rationale. But details on various reconstructions themselves will be of little pedagogical value.

- Understanding of phonological alterations and the modern phonetic inventory should prove extremely helpful in understanding readings. The raw test results on the phonetic element training experiment (Section 8.2) also indicated that understanding phonetic alterations would have helped avoid confusion among similar-sounding characters (commonly confused alternations included *jiu* and *qiu*, *tong* and *dong*, etc.).

- If the results of the experiments discussed in Chapter 8 continue to hold when tested on a greater scale—that is, using more input data as the subject matter and testing retention over longer timeframes—understanding readings and compositional rationale can be a significant aid to character meaning retention. Note, however, that although the more complex encoding scheme provided in the first experiment did increase retention by a factor of six after one week, even complex encoding should eventually decay below retrievability if not reinforced. Therefore, as a practical matter, assessments of how and when to review such information is needed. It would at least be periodically needed until the direct semantic recall route is consolidated.

- As my 2004 study pointed out, grouping of similarly composed characters significantly enhanced retention for readings as well as meanings. Therefore, organizing the data in ways that can enhance character discrimination will help with recall of character meanings (though see caveats below).

- Alas, the issue of whether knowledge of character readings and meanings and attention to morphological connections likewise enhance vocabulary retention remains

untested at this stage. In terms of the acquisition subcomponents elicited and vector relations deployed, this proposed intervention would be exactly parallel to that of the compositional rationale and phonetic element experiments, only this time the targeted interface of the orthography would be one step higher. For that reason, I suspect that such an intervention would likely bring similar benefits. In other words, knowing the meanings and readings of characters that make up a compound and analyzing how the characters interact to elicit that meaning should make character compounds more memorable. I hope to test this at some future date.

In Chapter 3, factors having a potentially direct bearing on cognitive activities such as perception, attention, and memory were specifically identified in the context of the Chinese orthography. The identified "subcomponents" of orthographic acquisition—exposure, familiarity, discrimination, connection, association, and retention—operate at both interfaces of the orthography, namely that between element and character, and between character and word. These subcomponents are not all obligatory at all interfaces. For example, native speakers are frequently unfamiliar with the significance of recurring components within the standard character domain, nor do they spend much time attending to the connections between words. But the act of identifying actual and potential subcomponents enables researchers to systematically test various hypotheses within a structured framework. Thus, one can, for example, hypothesize that discrimination at the recurring element level will prove to be extremely important in a system that contains thousand of constituent recurring elements, and where the target linguistic unit in the writing system is on the order of several thousand characters. Other subcomponents, such as association, connection, and familiarity, may also prove to be important. Although these skills have now been identified, it was unfortunately beyond the scope of the dissertation to test the effect of interventions at every level, though Chapter 8 did sample two possible relations: the effect of data enhancement (associations and connections) at the element-to-character interface. Future studies may help clarify the effect interventions targeting other subcomponents at other levels of the orthography.

Chapter 4 showed how current reading theory continues to encourage the use of top-down skills as a strategy for reading comprehension, even if such activities now appear under the guise of "*metacognitive processing.*" Not only does research caution against over-reliance on these strategies, but the insistence of always *learning in context* can lead to excessive context-dependence. On a more helpful note, automaticity theory (another major component of reading theory) identifies successful lexical processing as the foundation on which fluent reading takes place.

Chapter 5 examined lexical acquisition from the perspective of information processing theory, and provided a model whereby decoding or "bottom-up" processing is an act of constructing meaning by recursively identifying lexical units and passing them on up for further processing until meaning is interpreted and stored in memory. The possibility that lower levels of processing are subsumed is hypothesized as a way to free up the limited processing capacity of working memory for higher-order concerns. Within this model, rather than acting as an aid to understanding, metacognitive interventions are applied when things go wrong or some other triggering event occurs. This metacognitive module acts like a daemon that runs silently in the background until something prods it to action: either something in the text is of interest and provokes a reaction, or constructive lexical processing (the recursive act of building ever greater lexical units) breaks down somewhere along the line, prompting the need for intervention. The implications are obvious: learn the words, learn how phrases are put together (that is, understand the syntax), and more than half the battle will be won.

Meanwhile, in light of this model, much of the discussion of acquisition paths becomes largely a red herring. The alleged differences in acquisition speed among native speakers when various forms of priming are enacted are far below any threshold that should concern the student. Besides, the nature of the data itself heavily biases priming for phonological elements over (semantic) radicals, therefore skewing the results obtained in such studies. Further, the need to delay phonological activation until a second character is scanned in cases of words containing *duoyinzi* strongly suggests that activation is delayed until a word, not a character, is identified. In any event, once recall automaticity is achieved, whether humans access meaning directly or via a phonological route is not likely to be something under conscious control, making the issue moot.

Chapter 6 examined the nature of encoding, identified the vector nature of character knowledge, and relayed the significance of those factors to the lexical networks that are thought to exist in the mind. The possibility of such networks being allowed to form organically is contrasted with the intentional construction of highly structured networks. A distinction is also drawn between links that could be formed if supplementary information were provided and those that form when only minimal lexical data is provided. In short, by intentionally targeting subcomponents identified in Chapter 3, one could ensure that more complex and robust lexical associations are formed.

After demonstrating that experiments with an eye on identifying possible interventions need to be contextualized so as to better appreciate the many variables at play, Chapter 7 sought to tailor existing information on memory to a framework that will be more helpful in the SLS context, and specifically, for measuring lexical acquisition. To that end, the vague notion of memory strength was rendered by three variables: retrievability (a binary), accessibility (access speed), and longevity. At the same time, the many contributing factors to longevity and accessibility are categorized into two classes—amplification and attenuation. This will eventually be of benefit when scalars are defined for a system of assessing interventions, enabling quantitative comparisons of the effects of each type of intervention. And while more detailed work needs to be done in the future, this model points to the need to define timeframes from a more pragmatic standpoint, such that the term *long-term memory* as usually invoked in cognitive psychology studies be classified, at the very least, as consolidated or unconsolidated. If the ultimate pedagogical goal is to ensure that students reach a satisfactory vocabulary level, then identifying unconsolidated items which are prone to being lost, and providing just-in-time amplification, will be an important step along the way.

Chapter 7 also suggested that forgetting curves, which are typically thought of as representing memory strength, be redeployed as predictive models of longevity. Although a full implementation is not provided here, the implication is that, at least in theory, one should be able to predict memorability (the ease with which a given target is liable to be remembered) on the basis of three components—namely, its intrinsic,

extrinsic, and effective components. Doing this, and redefining forgetting-type curves in terms of not retrieval percentages for a set of learned items but expected longevity for a single target should set the stage for identifying the potential relative contributions that each memorability component might make. Eventually, if one can identify the relative weights of intrinsic, extrinsic, and effective components of memorability while establishing scalars for various types of amplification and attenuation, it should be possible to calculate longevity predictions as a function of these factors.[61] Having some measure of predictability will in turn eventually prove useful in the design of curricula.

The same process could be applied to accessibility. In other words, as accessibility is also affected differently by different types of amplification and attenuation, and may also have intrinsic, extrinsic, and effective components, it may be possible to measure and predict accessibility as a function of the same or similar parameters. One of the more interesting features of this new model is that depending on the manner of encoding, there can be a trade-off between accessibility and longevity, as some amplifications favor one over the other.[62] Taking the example of a summer intensive course, although the inherent massed study approach found in such courses can result in impressive gains, these gains in retrieval may be temporary, and if sufficient follow-up is not provided, knowledge degradation may readily ensue.

Finally, Chapter 8 discussed two experiments that essentially apply the notion of intracharacter associations and connections to test recall of meaning and reading. In both cases, positive results were obtained. Elaboration on inter-element relationships and making those elements meaningful helped to improve encoding for character meanings. Training to associate readings with phonetic readings helped to improve encoding for character readings. These results "make sense" from the standpoint that providing structure and elaborating on the data during encoding should improve encoding.

---

61 In simplistic terms, longevity predictions could be expressed as the expected time until retrieval is unlikely (below some form of established retrievability threshold). A more refined approach might express longevity in terms of the time until probability of retrieval falls below a certain threshold (e.g. 99%).

62 This finding is backed by the data from the first experiment described in Chapter 8.

9.2 Avenues for Future Research

Considering the many different topics and perspectives that bear on orthographic acquisition, it should be unsurprising that most of the features of the model were not explicitly tested. Here I would like to briefly describe some possible future tests that would be needed to advance researchers' understanding of the retention of lexical data. The list is almost discouragingly long, though on a positive note, I believe that the clear identification of such items is partly a result of having established a model that adopts multiple perspectives on the topic.

1. Parallel exploration at the word level

Chapter 8 explored single vectors at the interface between the lower two levels of the orthography: namely, the recurring element and the character. Exactly parallel tests could be run at the next higher interface, namely, between the character and the word. I suspect that it can be shown that consolidation of character meanings and readings, combined with focused attention on the morphological interactions between constituent characters in a word, may well enhance memorability.

2. Exploration of interference factors

The interventions tested in Chapter 8 were applied in isolation, that is, using total novices outside the context of a realistic CFL learning environment. As I mentioned earlier, it is clear that repeatedly applying such interventions could lead to some form of either proactive or retroactive interference among the bits of information to which the learner is exposed. This raises some questions. First, is this interference significant and problematic? If the answer is yes, then is the interference worse than the naturally occurring interference arising among characters when no intervention is applied? And if interference is problematic when applying interventions on a larger scale, then what steps can be made to reduce the interference that is being experienced? Here, obviously, one wants to consider the rates of exposure to and review of new material. These need to be applied to determine the consolidation time required for permanent or semi-permanent storage of the learned material. The need to assess interference exists

regardless of the level examined, whether at the interface between elements and characters, or the interface between characters and words.

3. Psychometric studies of amplification

The experiments of Chapter 8 were limited to specific vectors, namely graph to meaning and graph to reading. They also only tested limited forms of input. However, in real-life contexts, students do not simply "study" vocabulary lists, they also read materials, use words they have learned in dialog, and listen to materials. Each of these alternative forms of input can reinforce the memory for discrete lexical items. One interesting question becomes, How much amplification is achieved by each type of input? Further, input (amplification) of one kind is likely to affect different vectors differently. For example, writing a character by hand is likely to affect memory for a graph's form and features differently than reading that same character on a page in text. Is the recall of form amplified more by one form of input than another? And once attention is set at a different level (e.g. in reading, one focuses on words, not characters), is there any amplification benefit received at lower levels? In other words, if one's only input after some time becomes reading of words in text, is it possible for knowledge at lower levels in the orthographic hierarchy (characters and elements) to decay to the point that alternative retrieval paths are lost? If so, then for items not likely to be encountered regularly, strategic interventions devised to overcome decline of rare characters and/or elements may be indicated.

4. Quantitative assessment of intrinsic, extrinsic, and effective components of memorability

The holy grail in a fully functional psychometric model for lexical acquisition would be to establish actual values for the intrinsic, extrinsic, and effective components of memorability. This can only be done by means of a major study of all characters in a given domain, together with precise measurement of the various times and types of exposure provided. Given the inevitability of variation in exposures received by different subjects in real life, it seems unlikely that a perfectly accurate assessment could be devised for all characters at all levels. But software could in principle be designed to approximate this separation of memorability factors for certain subsets of characters

under certain strict experimental conditions (e.g. a first-year Chinese class where students have minimal extracurricular exposure to Chinese).

5. Differentiation of the effects of various forms of amplification

As was mentioned in Section 7.7, different types of amplification will have different effects in the degree to which they enhance both longevity and access speed. Massed study can quickly enhance retrieval speed, but may have an ephemeral effect on longevity. Cognitive interventions like that applied in Section 8.1 may be the opposite, having a strong correlation with longevity but possibly a negative (or at last, initially negative) correlation with access speed. Scales need to be established to put this form of study on a proper footing. Once that footing is established, curricula could be designed taking these results into consideration.

6. Establishment of meaningful timeframes for longevity

The previous avenue of research will be important in helping to establish meaningful timeframes by which to come up with a formula for re-amplification that prevents forgetting. If, for example, it becomes clear that the rates of first exposure and re-exposure employed in a certain curriculum lead to consolidation issues, or, as predicted by the model, will result in significant numbers of items falling below the retrievability threshold, then something needs to be done to revise the curriculum. The proposition that long-term memory be divided into *consolidated* and *unconsolidated* segments is an arbitrary first guess. Cognitive scientists are not in agreement as to whether—from a cognitive neuroscience standpoint—memory is a unary process or truly segmented into short-term and long-term components. Nonetheless, a pragmatic model is needed for SLS researchers, and the timeframe division proposed here is a first step in that direction.

7. Interplay of intensive and extensive reading

As a follow-up to the previous item, while there was no room to test the proposition in this dissertation, a likely solution may be the deployment of extensive reading activities *targeting a specific vocabulary domain* as a follow-up to intensive

readings conducted in that domain. The cycling between intensive reading (which is best for learning new words) and extensive reading (which helps with consolidation) is likely to be a very good curriculum design strategy.

* * * * *

In closing, I should point out that in this dissertation, I totally ignored two extremely important factors in lexical acquisition: motivation and affect. Because they are of a completely different nature than the material explored here, motivation and affect were left outside the scope of this study. Even so, they must not be ignored in future research seeking to develop positive interventions. An intervention may work well in theory and under limited trials, but if it leads to boredom or reluctance to study, then it will not be the solution one wants.

One of the overarching implications of this study is that extensive materials would need, in part, to be *lexically* motivated rather than *content* motivated, a notion that many proponents of the proficiency-oriented approach might cringe at. But if a critical threshold for independent reading ability is ever to be achieved, then one must identify and consolidate a core vocabulary, after which successive subsidiary vocabularies can be established on a field-dependent basis. To that end, extensive reading materials would have to cover that core list. At the same time, it is critical to establish optimal lexical densities for both intensive and extensive materials, while determining just how to cycle between the two genres in a way to maximize the likelihood of lexical consolidation. While developing materials that attain lexical goals, curriculum developers will nonetheless want to find ways to make those learning experiences meaningful to the student.

Also not resolved in this dissertation, though certainly implied by the above list of avenues of future research, is the complex interplay that must exist between the manner and timing of character data learning and that of word data learning. By learning the amplification and attenuation affects of learning by various approaches while accounting for the various factors of memorability, it is at least theoretically possible to greatly enhance lexical acquisition efficiency. Again, this is not simply a matter of learning lists of elements, and lists of characters, and lists of words, but rather

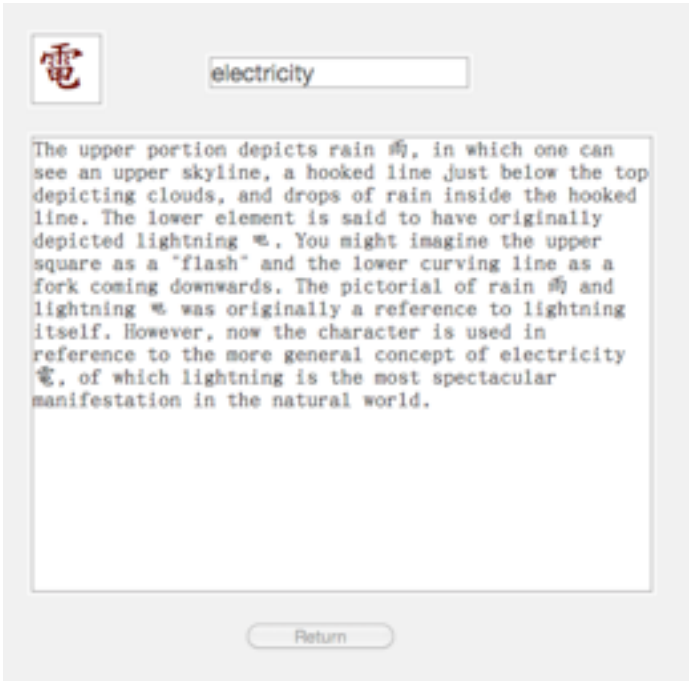
engaging in a dove-tailed sequence wherein elements are learned for a few characters, which are applied to a few words, which are then used in context. One cannot possibly front-load the entire orthography without bringing it to life for the student along the way.

Although such issues were beyond the scope of this study, it is my hope that the framework presented here will enable researchers to move forward in these directions.

Appendix

Subjects studied enhanced data or raw data for a character. Enhanced data provided a plausible compositional rationale for the character. Raw data simply provided the character meaning.

Enhanced Data



Raw Data

Enhanced Data

Characters Used for Raw Data, Showing Breakdown Between Raw (5) and Enhanced (5) Characters

Raw Data and Analyses of Correlations Between Correct Responses and Study Duration and Study Frequency

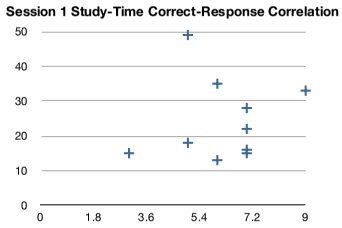
Correlations between amount of time studied for each character and the number of correct answers for that character are shown on bottom left half of page. The lack of close correlations suggests that extra study time was not the only, or even main, reason for improved correct response.

graph	meaning	is raw?	times	times std /	duration	dur std /	attempts 1	attempts 2	att total	correct 1	correct 2	correct	percent 1	percent 2	percent
美	beauty	FALSE	16	1.78	200850	22316	9	7	16	7	4	11	78	57	69
電	electricity	FALSE	16	1.78	253968	28218	9	7	16	7	2	9	78	29	56
象	elephant	FALSE	14	1.56	299125	33236	9	7	16	9	5	14	100	71	88
筆	writing	FALSE	21	2.33	436966	48551	9	7	16	5	2	7	56	29	44
菜	vegetable	FALSE	19	2.11	318094	35343	9	7	16	6	0	6	67	0	38
春	spring	TRUE	18	2.00	120957	13439	9	7	16	6	1	7	67	14	44
算	calculate	TRUE	19	2.11	147459	16384	9	7	16	7	0	7	78	0	44
事	thing	TRUE	19	2.11	131494	14610	9	7	16	3	1	4	33	14	25
乘	ride	TRUE	20	2.22	135792	15088	9	7	16	7	0	7	78	0	44
素	element	TRUE	28	3.11	165589	18398	9	7	16	5	0	5	56	0	31

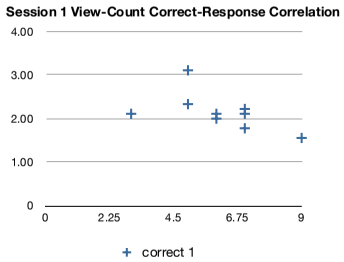
graph	meaning	dur std / sub	dur std (secs)	correct 1	correct 2	X-Mx 1	X - Mx 2	Y-My	SP 1	SP 2	SSx 1	SSx 2	SSy	r 1	r 2
美	beauty	22316	22	7	4	0.8	2.5	0.8	0.64	2.0	2.6	3.2	133.6	0.034	
電	electricity	28218	28	7	2	0.8	0.5	0.8	0.64	0.40					
象	elephant	33236	33	9	5	2.8	3.5	2.8	7.84	9.80					
筆	writing brush	48551	49	5	2	-1.2	0.5	-1.2	1.44	-0.60					
菜	vegetable	35343	35	6	0	-0.2	-1.5	-0.2	0.04	0.30					
春	spring	13439	13	6	1	-0.2	-0.5	-0.2	0.04	0.10					
算	calculate	16384	16	7	0	0.8	-1.5	0.8	0.64	-1.20					
事	thing	14610	15	3	1	-3.2	-0.5	-3.2	10.24	1.60					
乘	ride	15088	15	7	0	0.8	-1.5	0.8	0.64	-1.20		avg enh study	33.53		
素	element	18398	18	5	0	-1.2	-1.5	-1.2	1.44	1.80		avg raw study	15.58		
MEANS		24558.3	25	6.2	1.5										

By contrast, as seen on the bottom right half, there was actually a negative correlation between number of times viewed and correct answers for each graph. This is because people tended to view the raw data more frequently (they had to look at the raw data more often) to try to remember it. Despite the extra views for raw data, enhanced data performed better.

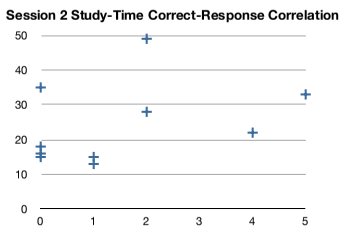
SESSION 1	correct 1	time studied (s)
美	7	22
電	7	28
象	9	33
筆	5	49
菜	6	35
春	6	13
算	7	16
事	3	15
乘	7	15
素	5	18
correlation	0.11	should be .549



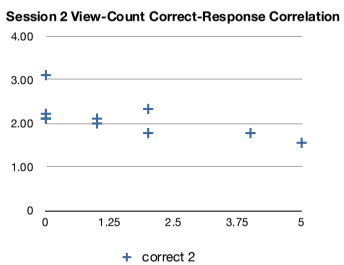
SESSION 1	correct 1	No. times studied
美	7	1.78
電	7	1.78
象	9	1.56
筆	5	2.33
菜	6	2.11
春	6	2.00
算	7	2.11
事	3	2.11
乘	7	2.22
素	5	3.11
correlation	-0.56	should be .549



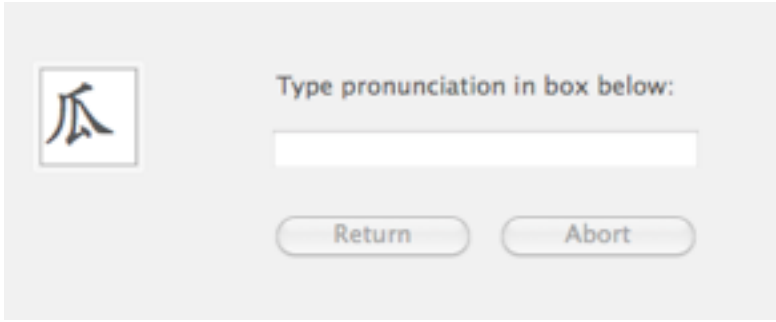
SESSION 2	correct 2	time studied (s)
美	4	22
電	2	28
象	5	33
筆	2	49
菜	0	35
春	1	13
算	0	16
事	1	15
乘	0	15
素	0	18
correlation	0.37	should be .549



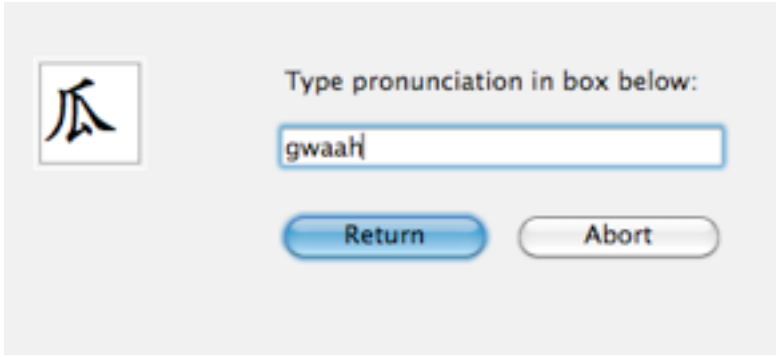
SESSION 2	correct 2	No. times studied
美	4	1.78
電	2	1.78
象	5	1.56
筆	2	2.33
菜	0	2.11
春	1	2.00
算	0	2.11
事	1	2.11
乘	0	2.22
素	0	3.11
correlation	-0.66	should be .549



Part 1 of the experiment involved clicking on a list of phonetics (see left) to learn its reading. The Φ symbol is a place marker for "null" (no phonetic).



Part 2 of the experiment involved testing the readings for phonetics. Subjects were shown a phonetic (top right) and asked to type the pronunciation. They would continue to be tested at random from among the phonetics until they typed each one within a certain period of time. Feedback was given as to correctness by indicating correct answers or (if they had made a mistake) showing the right answer, so that eventually every subject knew the phonetics and could provide the pronunciation quickly. The purpose was to make this information "known" in a way that might enhance extrinsic memorability for character readings.



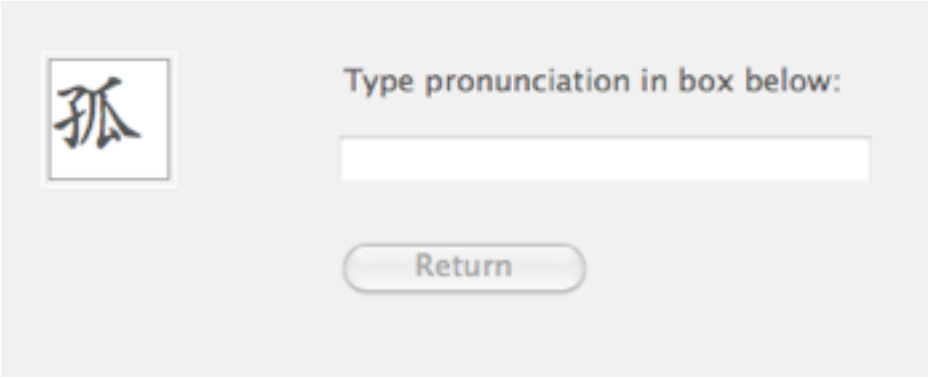
Part 3 involved having the subjects study groups of phonetics. The hide pronunciation buttons enabled them to test themselves during the study period. (See "gwaah" group below.) They did not know the meanings of the characters or the significance of the accompanying radicals, as this was a "pure" test of the value of phonetics.

Finally, in Part 4, subjects were tested in random order for the readings of the individual characters. The number of correct responses was evaluated with respect to the following potential variables:

- whether a phonetic had been provided (all groups vs. the control "Φ" group)

- the extent of phonetic consistency within the group on average (75%)

- whether there was a match between the character in question and that groups phonetic



Bibliography

Alber, Charles J. 1989. "Information Technology and Chinese Language Instruction—A Search for Standards." *Journal of the Chinese Language Teachers Association* 24, no. 2: 85-108.

Allen, Joseph Roe. 1984. "Chinese Script and Lexicography for the Uninitiated—Pedagogical Notes." *Journal of the Chinese Language Teachers Association* 19, no. 3: 35-86.

Astor, Captain. 1970. "A Phonetic-Inductive Approach to Chinese Character Recognition." *Journal of the Chinese Language Teachers Association* 5, no. 2: 30-66.

Baddeley, Alan D. 1990. *Human Memory*. Boston: Allyn and Bacon.

———. 2001. "Is Working Memory Still Working?" *American Psychologist* 56: 849-864.

Bamford, Julian and Richard R. Day. 1998. *Extensive Reading in the Second Language Classroom*. New York: Cambridge University Press.

———. 2004. *Extensive Reading Activities for Teaching Language*. New York: Cambridge University Press.

Barnard, Noel. 1978. "The Nature of the Ch'in 'Reform of the Script' as Reflected in Archeological Documents Excavated Under Conditions of Control," in *Ancient China: Studies in Early Civilization*, edited by David T. Roy and Tsuen-hsuin Tsien, 181-213. Hong Kong: Chinese University Press.

Baxter, William H. III. 1992. *A Handbook of Old Chinese Phonology*. Trends in Linguistics Studies and Monographs 64. Berlin: Mouton de Gruyter.

Baxter, William H. III and L. Sagart. 1998. "Word Formation in Old Chinese," in *New Approaches to Chinese Word Formation: Morphology, Phonology, and the Lexicon in Modern and Ancient Chinese*, edited by Jerome L. Packard, 35-76. Berlin: Mouton de Gruyter.

Boltz, William G. 1994. *The Origin and Development of the Chinese Writing System*. New Haven, CT: American Oriental Society.

Boodberg, Peter A. 1937. "Some Proleptical Remarks on the Evolution of Archaic Chinese." *Harvard Journal of Asiatic Studies* 2: 329-372.

———. 1940. "'Ideography' or Iconolatry?" *T'oung Pao* 35: 266-288.

Bottéro, Françoise. 1996. *Sémantisme et classification dans l'écriture chinoise*. Paris: College des Hautes Études Chinoises.

Brooks, Lee. 1977. "Visual Pattern in Fluent Word Identification." *Toward a Psychology of Reading: The Proceedings of the CUNY Conferences*, edited by Arthur S. Reber and Don L. Scarborough, 141-183. Hillsdale, NJ: L. Erlbaum Associates.

Brown, Douglas H. 1993. *Principles of Language Learning and Teaching*. Englewood Cliffs, NJ: Prentice Hall Regents.

Brown, J. D. 1988. *Understanding Research in Second Language Learning*. Cambridge: Cambridge University Press.

Brown, J. D. and Thom Hudson. 2002. *Criterion-Referenced Language Testing*. Cambridge: Cambridge University Press.

Byrnes, James P. 1996. *Cognitive Development and Learning in Instructional Contexts*. Needham, MA: Allyn and Bacon.

Campbell, Lyle. 1998. *Historical Linguistics: An Introduction.* Cambridge, MA: The MIT Press.

Carello, Claudia, M. T. Turvey, and Georgije Lukatela. 1992. "Can Theories of Word Recognition Remain Stubbornly Nonphonological?" In *Orthography, Phonology, Morphology, and Meaning.* Advances in Psychology Series No. 94, edited by Frost and Katz, 211-226. Amsterdam, London, New York, Tokyo: North-Holland.

Carr, Michael. 1981. "Pedagogy, Radicals, and Grapho-Semantic Fields." *Journal of the Chinese Language Teachers Association* 16, no. 3: 51-66.

Carreiras M., M. Perea M. and J. Grainger. 1997. "Effects of Orthographic Neighborhood in Visual Word Recognition: Cross-Task Comparisons." *Journal of Experimental Psychology: Learning Memory, and Cognition* 23, no. 4: 857-71.

Carroll, David W. 2004. *Psychology of Language*. 4th ed. Belmont, CA: Wadsworth/ Thomson Learning.

Catford, J. C. 1994. *A Practical Introduction to Phonetics*. Oxford: Oxford University Press.

Chao Yuen Ren. 1968a. *A Spoken Grammar of Chinese*. Berkeley: University of California Press,.

———. 1968b. *Language and Symbolic Systems*. London: Cambridge University Press.

Chen Hsuan-Chin, ed. 1986. "Component Detection in Reading Chinese Characters," in *Linguistics, Psychology, and the Chinese Language*. Centre of Asian Studies Occasional Papers and Monographs No. 67, edited by Kao and Hoosain, 1-10. Hong Kong: Centre of Asian Studies, University of Hong Kong.

———. 1997. *Cognitive Processing of Chinese and Related Asian Languages*. Hong Kong: The Chinese University Press.

Chen Hsuan-Chin and Ovid J. L. Tzeng, eds. 1992. *Language Processing in Chinese*. Amsterdam: Elsevier Science Publishers.

Cheng Chao-Ming. 1992. "Lexical Access in Chinese: Evidence from Automatic Activation of Phonological Information," in *Language Processing in Chinese*, edited by Chen and Tzeng, 67-92. Amsterdam: Elsevier Science Publishers.

Cheng Tsai-fa. 1985. *Ancient Chinese and Early Mandarin*. Journal of Chinese Linguistics, Project on Linguistic Analysis, University of Berkeley.

Child, W. Daniel. 2004. "Towards a Typology for Kanji of Japanese Origin—the Underlying Rationale in Kokuji Composition." *Selected Papers from the Eighth College-Wide Conference for Students in Languages, Linguistics, and Literature*, edited by Laura Clarito and Treela McKamey, 12-26. Honolulu: NFLRC, University of Hawaii at Manoa.

———. 2006. "Pilot Study on the Effect of Grouping and Compositional Analysis on Character Retention." *Selected Papers from the Tenth College-Wide Conference for Students in Languages, Linguistics, and Literature*, edited by Zoe Madden-Wood and Kaori Ueki, 12-22. Honolulu: NFLRC, University of Hawaii at Manoa.

Choy, Rita Mei-Wah. 1990. *Read and Write Chinese: A Simplified Guide to the Chinese Characters*. San Francisco: China West Publishing.

Chung, Karen S. 2001. "Second Review of Packard: Chinese Morphology." http://linguistlist.org/issues/12/12-844.html.

Cook, Vivian and Benedetta Bassetti. 2005. *Second Language Writing Systems*. Second Language Acquisition 11, edited by David Singleton. Cleveland, UK: Multilingual Matters Limited.

Coulmas, Florian. 1989. *The Writing Systems of the World*. Oxford: Blackwell Publishers.

Craik, Fi. I. M., and R. S. Lockhart. 1972. "Levels of Processing: A Framework for Memory Research." *Journal of Verbal Learning and Verbal Behavior* 11: 671-684.

Creel, Herrlee Glessner. 1936. "On the Nature of Chinese Ideography. *T'oung Pao* 32: 85-161.

———. 1938. "On the Ideographic Element in Ancient Chinese." *T'oung Pao* 34: 265-294.

Crowley, Terry. 1997. *An Introduction to Historical Linguistics*. 3rd ed. Auckland: Oxford University Press.

DeFrancis, John. 1984. *The Chinese Language: Fact and Fantasy*. Honolulu: University of Hawaii Press.

———. 1989. *Visible Speech: The Diverse Oneness of Writing Systems*. Honolulu: University of Hawaii Press.

de Groot, Annette M. B. 1992. "Bilingual Lexical Representation: A Closer Look at Conceptual Representations," in *Orthography, Phonology, Morphology, and Meaning*. Advances in Psychology Series No. 94, edited by Frost and Katz, 389-412. Amsterdam, London, New York, Tokyo: North-Holland.

Dew, James Erwin. 1999. *6000 Chinese Words: A Vocabulary Frequency for Chinese Language Teachers and Students*. Taipei: SMC Publishing.

Dong Tonghe 董同龢. *Hanyu yinyunxue* 漢語音韻學. Taipei: Guangwen Shuju, 1968.

Duanmu, San. 1997. "Wordhood in Chinese," in *New Approaches to Chinese Word Formation: Morphology, Phonology, and the Lexicon in Modern and Ancient Chinese*, edited by Jerome L. Packard, 135-196. Berlin: Walter de Gruyter and Co..

———. 2002. review of *The Morphology of Chinese: A Linguistic and Cognitive Approach* by Jerome L. Packard, Cambridge University Press. *Diachronica* 19, no. 1: 190-200.

Ebbinghaus, H. (1885) 1913. *Memory: A Contribution to Experimental Psychology*. Translated by H. A. Ruger and C. E. Bussenius. New York: Teachers College Press.

Eichenbaum, Howard. 2007. "Persistence: Necessary, But Not Sufficient," in *Science of Memory: Concepts*, edited by Henry L. Roediger, Yadin Dadai, and Susan M. Fitzpatrick, 193-198. New York: Oxford University Press.

Erbaugh, Mary, ed. 2002. *Difficult Characters: Interdisciplinary Studies of Chinese and Japanese Writing*. Columbus, OH: Ohio State Foreign Language Publications and Services.

———. 2002. "How the Ideographic Myth Alienates Asian Studies from Psychology and Linguistics" In *Difficult Characters: Interdisciplinary Studies of Chinese and Japanese Writing*, edited by Mary Erbaugh, 21-51. Columbus, OH: Ohio State Foreign Language Publications and Services.

Everson, Michael E. 1997. "An Inquiry into the Reading Strategies of Intermediate and Advanced Learners of Chinese as a Foreign Language." *Journal of the Chinese Language Teachers Association* 32, no. 1: 1-20.

Eysenck, Michael W. and Keane, Mark T. 1995. *Cognitive Psychology: A Student's Handbook* (3rd edition). United Kingdom: Psychology Press.

Fang Sheng-Ping, Horng Ruey-Yun, and Ovid Tzeng. 1986. "Consistency Effects in the Chinese Character and Pseudo-Character Naming Tasks," in *Linguistics, Psychology, and the Chinese Language.* Centre of Asian Studies Occasional Papers and Monographs No. 67, edited by Kao and Hoosain, 11-22. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Feldman, Laurie B. and Darinka Andjelkovic. 1992. "Morphological Analysis in Word Recognition," chap. 17 in *Orthography, Phonology, Morphology, and Meaning*. Amsterdam, London, New York, Tokyo: North-Holland.

Feng Shengli. 1997. "Prosodic Structure and Compound Words in Classical Chinese," in *New Approaches to Chinese Word Formation: Morphology, Phonology, and the Lexicon in Modern and Ancient Chinese*, edited by Jerome L. Packard, 197-260. Berlin: Walter de Gruyter and Co.

Flores d'Arcais, Giovanni B. 1992. "Graphemic, Phonological, and Semantic Activation Processes during the Recognition of Chinese Characters," in *Language Processing in Chinese*, edited by Chen and Tzeng, 37-66. Amsterdam: Elsevier Science Publishers.

Flores-d'Arcais, Giovanni .B., H. Saito, and M. Kawakami. 1995. "Phonological and Semantic Activation in Reading Kanji Characters." *Journal of Experimental Psychology: Learning, Memory and Cognition* 21, no. 1:34-42.

Fok, Angela, Ursula Bellugi, and Diane Lillo-Martin. 1986. "Remembering in Chinese Signs and Characters," in *Linguistics, Psychology, and the Chinese Language.* Centre

of Asian Studies Occasional Papers and Monographs No. 67, edited by Kao and Hoosain, 177-202. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Frost, Ram. 1998. "Toward a Strong Phonological Theory of Visual Word Recognition: True Issues and False Trails." *Psychological Bulletin* 123, no. 1 (January): 71-99.

Frost, Ram and Leonard Katz, eds. 1992. *Orthography, Phonology, Morphology, and Meaning*. Advances in Psychology Series No. 94. Amsterdam, London, New York, Tokyo: North-Holland.

Gao Hongjin 高鴻儘. 1963. *Zhongguo zi li* 中國字例. Taipei: Sanmin Shuju Yinhang.

Gathercole, Susan and Alan D. Baddeley. 1993. *Working Memory and Language.* East Sussex: Psychology Press.

Gendreau-Hétu, Jean-Pierre. 1995. "Prolégomènes a une théorie linguistique de l'écriture." Masters thesis, Department of Linguistics and Translation, University of Montreal.

Gleitman, Lila R., and Paul Rozin. 1977. "The Structure and Acquisition of Reading I: Relations between Orthographies and the Structure of Language," in *Toward a Psychology of Reading*, edited by Arthur S. Reber and Don L. Scarborough, 1-53. Hillsdale, NJ.

Goodman, Kenneth S. 1970. "Psycholinguistic Universals in the Reading Process" *Visible Language* 4, no. 2: 103-110.

Grainger, Jonathan. 1992. "Orthographic Neighborhoods and Visual Word Recognition," chap. 7 in O*rthography, Phonology, Morphology, and Meaning*. Advances in Psychology Series No. 94, edited by Frost and Katz.Amsterdam, London, New York, Tokyo: North-Holland.

Halpern, Jack. ed. *The Kodansha Kanji Learner's Dictionar*y. Tokyo: Kodansha International: 1999.

Hannas, William C. *Asia's Orthographic Dilemma*. Honolulu: University of Hawai'i Press, 1997.

Harbaugh, Rick. 1998. *Chinese Characters: A Genealogy and Dictionary*. New Haven: Yale University Press.

Hashimoto, Mantaro J. 1978. *Phonology of Ancient Chinese*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Haspelmath, Martin. 2002. *Understanding Morphology*. New York: Oxford University Press.

Hasselmo, Michael E. 2007. "Encoding: Models Linking Neural Mechanisms to Behavior," chap. 22 in *Science of Memory: Concepts*, edited by Henry L. Roediger, Yadin Dadai, and Susan M. Fitzpatrick, 123-128. New York: Oxford University Press.

Hayden, Jeffrey J. 2003. "Shocking Our Students to the Next Level: Language Loss and Some Implications for Teaching Chinese as a Foreign Language." *Journal of the Chinese Language Teachers Association* 38, no. 3: 1-20.

———. 2004. "Why Johnny Can Read Chinese: Working Memory, Cognitive Processes, and Reading Comprehension." PhD diss., University of Hawaii.

Hayes, E. B. 1987. "The Relationship Between Chinese Character Complexity and Character Recognition." *Journal of the Chinese Language Teachers Association* 22: 45-57.

Hirsh, David and Paul Nation. 1992. "What Vocabulary Size is Needed to Read Unsimplified Texts for Pleasure?" *Reading in a Foreign Language* 8, no. 2.

Hirsch, E.D. 2003. "Reading Comprehension Requires Knowledge—of Words and the World" *American Educator*. (Spring) 10-45.

Hoosain, Rumjahn. 1991. *Psycholinguistic Implications for Linguistic Relativity: A Case Study of Chinese*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

———. 1992. "Psychological Reality of the Word in Chinese," in *Language Processing in Chinese*, edited by Chen and Tzeng, 111-130. Amsterdam: Elsevier Science Publishers.

Hoosain, Rumjahn, Li Hai Tan, and Dan-ling Peng. 1995. "Role of Early Presemantic Phonological Code in Chinese Character Identification." *Journal of Experimental Psychology: Learning, Memory and Cognition* 21, no. 1:43-54.

Hsu, Kylie. 2002. *Selected Issues in Mandarin Chinese Word Structure Analysis*. Lewiston, NY: Edwin Mellen Press.

Huang Jong Tsun and Man-Yin Wang. 1992. "From Unit to Gestalt: Perceptual Dynamics in Recognizing Chinese Characters," in *Language Processing in Chinese*, edited by Chen and Tzeng, 3-36. Amsterdam: Elsevier Science Publishers.

Hummel, Kirsten M. 2010. "Translation and Short-Term L2 Vocabulary Retention: Hindrance or Help?" *Language Teaching Research* 14, no. 1: 61-74.

Hung, Daisy L., Ovid J. L. Tzeng, and Angela K.Y. Tzeng. 1992. "Automatic Activation of Linguistic Information in Chinese Character Recognition," in *Orthography, Phonology, Morphology, and Meaning.* Advances in Psychology Series No. 94, edited by Frost and Katz.Amsterdam, 119-130. London, New York, Tokyo: North-Holland.

International Chinese Language Program at National Taiwan University. 1997. *Thought and Society: An Advanced Text for Spoken Chinese*. Taipei: SMC Publishing Inc..

Jaxontov, Sergej. 1978-1979. "Old Chinese Phonology", *Early China* 4: 37-40. Jerry Norman trans.

Kano Yoshimitu 加納喜光. 2001. Kanji no naritachi jiten 漢字の成り立ち辞典. Tokyo: Tokyotou Chuban.

Kao, Henry S. R. and Rumjahn Hoosain. 1986. *Linguistics, Psychology, and the Chinese Language. Centre of Asian Studies Occasional Papers and Monographs No. 67*. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Karlgren, Bernard. (1923) 1991. *Analytic Dictionary of Chinese and Sino-Japanese*, New York: Dover Publications.

———. 1934. "Word Families in Chinese." *Bulletin of the Museum of Far Eastern Antiquities* 5: 1-112.

Karpicke, Jeffrey D. and Henry L. Roediger III. 2007. "Expanding Retrieval Practice Promotes Short-Term Retention, but Equally Spaced Retrieval Enhances Long-Term Retention." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33, no. 4: 704-719.

Katamba, Francis. 1993. *Morphology*. London: MacMillan Press.

Ke, C. 1996. "An Empirical Study on the Relationship Between Chinese Character Recognition and Production." *Modern Language Journal* 80: 340-350.

Kess, Joseph F. and Tadao Miyamoto. 1997. "Accessing the Japanese Mental Dictionary Through the Japanese Writing System." *Department of Linguistics Occasional Paper #15*, Department of Linguistics. Victoria, Canada: University of Victoria.

Kinoshita Sachiko. 1998. "The Role of Phonology in Reading Japanese: Or Why I Don't Hear Myself When Reading Japanese," in *Cognitive Processing of the Chinese and the Japanese Languages*, edited by Leong and Tamaoka, 285-302. Netherlands. Kluwer Academic Publishers.

Koda Keiko. 2004. *Insights into Second Language Reading: A Cross-Linguistic Approach*. New York: Cambridge University Press.

Kraemer, Stephen M. 1980. "Potentially Pedagogically Useful 'Phonetics' in the Chinese Script: Their Identification and Characterization." PhD diss., Rutgers University.

———. 1991. "Sound Clues in Mandarin Character Phonetic Series." Manuscript. The Ohio State University.

Kratochvil, Paul. 1968. *The Chinese Language Today.* London: Hutchinson and Co.

LeDoux, Joseph E. 2007. "Consolidation: Challenging the Traditional View," in *Science of Memory: Concepts*, edited by Henry L. Roediger, Yadin Dadai, and Susan M. Fitzpatrick, 171-176. New York: Oxford University Press.

Leong Che Kan and Tamaoka Katsuo. 1998. *Cognitive Processing of the Chinese and the Japanese Languages*. Netherlands. Kluwer Academic Publishers.

Li, Charles N. and Sandra A. Thompson. 1981. *Mandarin Chinese*. Berkeley: University of California Press.

Li Fang-kuei. 1983. "Archaic Chinese," in *The Origins of Chinese Civilization*, edited by David N. Keightley, 393-408. Berkeley: University of California Press.

Liu In-Mao. 1997. "The Issue of Prelexical Phonology in the Reading of Chinese Characters," in *Cognitive Processing of Chinese and Related Asian Languages*, edited by Hsuan-Chi Chen, 65-76. Hong Kong: Chinese University Press.

Liu, Irene. 1983. "The Learning of Characters—A Conceptual Learning Approach." *Journal of the Chinese Language Teachers Association* 18: 2.

Liu N. and I. S. P. Nation. 1985. "Factors Affecting Guessing Vocabulary in Context." *RELC Journal* 16, no. 1: 33-42.

Liu Ying and Dan-ling Peng. 1997. "Meaning Access of Chinese Compounds and Its Time Course," in *Cognitive Processing of Chinese and Related Asian Languages*, edited by Hsuan-Chi Chen, 219-232. Hong Kong: Chinese University Press.

Liu Yuehua, Tao-chung Yao, Liangyan Ge, Yea-fen Chen, Nyan-ping Bi and Xiaojun Wang. 1997. *Integrated Chinese Level I*. Boston: Cheng and Tsui.

Liu Yuehua, Tao-chung Yao, Nyan-ping Bi, and Yaohua Shi. 1997. *Integrated Chinese Level II*. Boston: Cheng and Tsui.

Loftus, Elizabeth F. 2007. "Forgetting: The Fate of Once Learned, But "Forgotten", Material," in *Science of Memory: Concepts*, edited by Henry L. Roediger, Yadin Dadai, and Susan M. Fitzpatrick, 321-323. New York: Oxford University Press.

Loftus, Geoffrey R. 1985. "Evaluating Forgetting Curves." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11, no. 2: 397-406.

Logan, Gordon D. 1988. "Toward an Instance Theory of Automatization." *Psychological Review*, 95, no. 4: 492-527.

———. 1992. "Attention and Preattention in Theories of Automaticity. *American Journal of Psychology* 105: 317-339.

———. 1997. "Automaticity and Reading: Perspectives from the Instance Theory of Automatization." *Reading and Writing Quarterly*, 13, no. 2 (April - June: 123-147.

Luo, Chun R., Reed A. Johnson, and David A. Gallo. 1998. "Automatic Activation of Phonological Information in Reading: Evidence from the Semantic Relatedness Decision Task." *Memory and Cognition* 26, no. 4: 833-843.

Mae, Chu-chang, and Donald J. Loritz. 1977. "Phonological Encoding of Chinese Ideographs in Short-Term Memory." *Language Learning* 27, no. 2: 341-351.

Mann, Virginia A. 1986. "Temporary Memory for Linguistic and Nonlinguistic Material in Relation to the Acquisition of Japanese Kana and Kanji," in *Linguistics, Psychology, and the Chinese Language.* Centre of Asian Studies Occasional Papers and Monographs No. 67, edited by Kao and Hoosain, 155-168. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Matin, Ethel. 1974. "Saccadic Suppression: A Review and an Analysis." *Psychological Bulletin* 81, no. 12: 899-917.

Matsunaga Sachiko. 2002. "Are Chinese Characters Ideographs? An Argument from the Psycholinguistic Perspective," in *Difficult Characters: Interdisciplinary Studies of Chinese and Japanese Writing*, edited by Mary Erbaugh, 75-91. Columbus, OH: Ohio State Foreign Language Publications and Services.

Mattingly, Ignatius G. 1992. "Linguistic Awareness and Orthographic Form," in *Orthography, Phonology, Morphology, and Meaning.* Advances in Psychology Series No. 94, edited by Frost and Katz, 11-26. Amsterdam, London, New York, Tokyo: North-Holland.

McBride-Chang, Catherine and Hsuan-Chih Chen, eds. 2003. *Reading Development in Chinese Children*. Westport, CT: Praeger.

McDermott, Kathleen B. 2007. "Integrative Comments—Retrieval: Varieties and Puzzles," in *Science of Memory: Concepts*, edited by Henry L. Roediger, Yadin Dudai, and Susan M. Fitzpatrick, 225-229. New York: Oxford University Press.

McEwen, Ping. 2006. "Vocabulary Acquisition in CFL (Chinese as a Foreign Language) Contexts: A Correlation of Performance and Strategy Use." MA Thesis. Center for Language Studies: Brigham Young University.

Meara, Paul. 1997. "Towards a New Approach to Modeling Vocabulary Acquisition," in *Vocabulary: Description, Acquisition, and Pedagogy*, edited by Schmitt and McCarthy, 109-121. Cambridge: Cambridge University Press.

Melka, Francine. 1997. "Receptive vs. Productive Aspects of Vocabulary," in *Vocabulary: Description, Acquisition, and Pedagogy*, edited by Schmitt and McCarthy, 84-102. Cambridge: Cambridge University Press.

Miles, Matthew B. and Huberman, Michael. 1999. *Qualitative Data Analysis*. London: Sage Publications.

Miyake, Marc. 1999. *The Phonology of Eighth Century Japanese Revisited: A Reconstruction Based upon Written Records*. PhD diss., University of Hawaii at Manoa.

Nadel, Lynn. 2007. "Consolidation: The Demise of the Fixed Trace," in *Science of Memory: Concepts*, edited by Henry L. Roediger, Yadin Dadai, and Susan M. Fitzpatrick, 177-182. New York: Oxford University Press.

Nagy, William. 1997. "On the Role of Context in First- and Second-Language Vocabulary Learning," in *Vocabulary: Description, Acquisition, and Pedagogy*, edited by Schmitt and McCarthy, 64-83. Cambridge: Cambridge University Press.

Nara, Hiroshi and Noda, Mari. 2004. *Acts of Reading: Exploring Connections in Pedagogy of Japanese*. Columbus, OH: Ohio State University Press.

Nation, Paul and Robert Waring. 2997. "Vocabulary Size, Text Coverage, and Word Lists," in *Vocabulary: Description, Acquisition, and Pedagogy*, edited by Schmitt and McCarthy, 6-19. Cambridge: Cambridge University Press.

Norman, Jerry. 1974. "The Initials of Proto-Min", *Journal of Chinese Linguistics* 2: 27-36.

———. 1988. *Chinese*. Cambridge, UK: Cambridge University Press.

Nunan, David. 1999. *Second Language Teaching and Learning*. Boston: Heinle and Heinle.

Omaggio Hadley, Alice. 2001. *Teaching Language in Context*. (3rd ed.) Boston: Heinle and Heinle.

Osherson, Daniel N and Howard Lasnik, eds. 1990. *An Invitation to Cognitive Science Vol. 1: Language*. Cambridge, MA: MIT Press.

Packard, Jerome Lee, ed. 1997. *New Approaches to Chinese Word Formation: Morphology, Phonology, and the Lexicon in Modern and Ancient Chinese*. Berlin: Walter de Gruyter and Co.

———. 2000. *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge, U.K.: New York: Cambridge University Press.

Paran, Amos. 1996. "Reading in EFL: Facts and Fictions." *ELT Journal* 50, no. 1: 25-34.

Peng Rui-Xiang and Chen Yong-Ming. 1986. "Preliminary Study of Semantic Retrieval in Memory for Chinese Language," in *Linguistics, Psychology, and the Chinese Language.* Centre of Asian Studies Occasional Papers and Monographs No. 67, edited by Kao and Hoosain, 169-176. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Perfetti, Charles A. 1999. "The Cognitive Science of Word Reading: What Has Been Learned from Comparisons across Writing Systems?" Second International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference, July (http://logos.mind.sccs.chukyo-u.ac.jp/jcss/ICCS/99/).

Perfetti, Charles A. and Li Hai Tan. 1999. "Phonological Activation in Visual Identification of Chinese Two-Character Words." *Journal of Experimental Psychology: Learning, Memory and Cognition* 25, no. 2: 382-393.

———. 1998. "The Time Course of Graphic, Phonological, and Semantic Activation in Chinese Character Identification." *Journal of Experimental Psychology: Learning, Memory and Cognition* 24, no. 1: 101-118.

Perfetti, Charles A. and Sulan Zhang. 1995. "Very Early Phonological Activation in Chinese Reading." *Journal of Experimental Psychology: Learning Memory, and Cognition* 21, no. 1 : 24-33.

Perfetti, Charles A., Sulan Zhang, and Iris Berent. 1992. "Reading in English and Chinese: Evidence for a 'Universal' Phonological Principle," in *Orthography, Phonology, Morphology, and Meaning*. Advances in Psychology Series No. 94, edited by Frost and Katz, 227-248..Amsterdam, London, New York, Tokyo: North-Holland.

Perfetti, Charles A. and Y. Liu. 2006. "Reading Chinese Characters: Orthography, Phonology, Meaning, and the Lexical Constituency Model," chap. 20 in *The Handbook of East Asian Psycholinguistics, Vol. 1: Chinese*, edited by P. Li, L. H. Tan, E. Bates, and O.J. L. Tzeng. New York: Cambridge University Press.

Perfetti, Charles A., Y. Liu, and L. H. Tan. 2002. "How the Mind Can Meet the Brain in Reading: A Comparative Writing Systems Approach," in *Cognitive Neuroscience Studies of the Chinese Language*, edited by Kao, Leong, and Gao, 36-60. Hong Kong: Hong Kong University Press.

Pulleyblank, Edwin G. 1984. *Middle Chinese: A Study in Historical Phonology*. Vancouver: University of British Columbia Press.

Qiu Xigui. 2000. *Chinese Writing*. Translated by Gilbert L. Mattos and Jerry Norman. Berkeley, CA: Society for the Study of Early China.

Ram, Ashwin and Kenneth Moorman, eds. 1999. "Towards a Theory of Reading and Understanding," in *Understanding Language Understanding: Computational Models of Reading*, edited by Ram and Moorman, 1 - 11. Boston: MIT Press.

ROC Overseas Chinese Affairs Commission 中華民國僑務委員會. 1997. *Tanjiu zhongguo wenzi jianhua wenti lunwen*. 探究中國文字簡化問題論文集. Taiwan: 101-122.

Roediger, Henry L. 2008. "Relativity of Remembering: Why the Laws of Memory Vanished." *Annual Review of Psychology* 59: 225-254.

Roediger, Henry L., Yadin Dudai, & Susan M. Fitzpatrick., eds. *2007. Science of Memory: Concepts*. New York: Oxford University Press.

Rubin, David C. 2007. "Forgetting: Its Role in the Science of Memory," in *Science of Memory: Concepts*, edited by Henry L. Roediger, Yadin Dadai, and Susan M. Fitzpatrick, 325-328. New York: Oxford University Press.

Sagart, Laurent. 1999. *The Roots of Old Chinese*. Amsterdam, Philadelphia: John Benjamins Publishing Co.

Sampson, Geoffrey. 1985. *Writing Systems*. Stanford: Stanford University Press.

Samuels, S. Jay, and Richard F. Flor. 1997. "The Importance of Automaticity for Developing Expertise in Reading." *Reading and Writing Quarterly*, 13, no. 2 (April - June): 107-123.

Schmitt, Norbert and Michael McCarthy, eds. 1997. V*ocabulary: Description, Acquisition, and Pedagogy*. Cambridge: Cambridge University Press.

Schuessler, Alex. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu. University of Hawai'i Press.

Schunk, Dale H. 2004. *Learning Theories: An Educational Perspective*. 4th ed. Upper Saddle River, NJ: Pearson Education.

Segui J, Grainger J. 1990. "Priming Word Recognition with Orthographic Neighbors: Effects of Relative Prime-Target Frequency." *Journal of Experimental Psychology: Human Perception and Performance* 16, no. 1: 65-76.

Smith, Edward E. and Stephen M. Kosslyn, eds. 2007. *Cognitive Psychology: Mind and Brain*. New Jersey: Prentice Hall.

Smith, Norval and Wang Jialing. 1997. *Studies in Chinese Phonology*. Berlin: Walter de Gruyter & Co..

Sokmen, Anita J. 1997. "Current Trends in Teaching Second Language Vocabulary," in *Vocabulary: Description, Acquisition, and Pedagogy*, edited by Schmitt and McCarthy, 237-257. Cambridge: Cambridge University Press.

Spear, Norman E. 2007. "Retrieval: Properties and Effects," in *Science of Memory: Concepts*, edited by Henry L. Roediger, Yadin Dadai, and Susan M. Fitzpatrick, 215-220. New York: Oxford University Press, 2007.

Spencer, Andrew. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar.* Cambridge, MA: Basil Blackwell Ltd.

Sproat, Richard. 2001. "Review of Packard's The Morphology of Chinese." http://linguistlist.org/issues/ 12/12-11.html.

Stanovich, K.E. 1980. "Toward an Interactive-Compensatory Model of Individual Differences in the Development of Reading Fluency." *Reading Research Quarterly* 16, no. 1: *32-71.*

———. 1987. "The Impact of Automaticity Theory." *Journal of Learning Disabilities* 20, no. 3: 167–168.

———. 1990. "Concepts in Developmental Theories of Reading Skill: Cognitive Resources, Automaticity, and Modularity." *Developmental Review* 10: 72-100.

Sydney Morning Herald, May 29, 2010. http://www.smh.com.au/opinion/society-and-culture/first-reduce-the-dropout-rate-20100528-wlbg.html.

Sydorenko, Tetyana. 2010. "Modality of Input and Vocabulary Acquisition." *Language Learning and Technology.* 14, no. 2: 50-73.

Takazawa Suiun. 高澤翠雲. 1999. *Tensho jiten* 篆書辞典. Tokyo: Seitoshobo.

Todo Akiyasu 藤堂明保. 1957. *Chugoku no on'inron* 中国の音韻論. Tokyo: Kounan shoin.

———. 1965. *Kanji gogen jiten* 漢字語源典. Tokyo: Gakutousha.

Todo Akiyasu 藤堂明保 and Kobayashi Hiroshi 小村博校本. 1971. *Incho inkyo kohon* 音注韻鏡校本. Tokyo: Mokujisha

Townsend, J. T. 1971. "A Note on the Identifiability of Parallel and Serial Processes," *Attention, Perception, and Psychophysics* 10, no. 3: 161-163.

Tsai Chih-hao. 2005. *Chih-hao Tsai's technology page*. http://technology.chtsai.org/charfreq/.

Tsao, Yao-chung, and Tsai-guey Wang. 1983. "Information Distribution in Chinese Characters." *Visible Language* 17, no. 4: 357-364.

Tseng, Ovid and Daisy L. Hung. 2002. "A Phantom of Linguistic relativity: Script, Speech, and Thought," in *Difficult Characters: Interdisciplinary Studies of Chinese and Japanese Writing*, edited by Mary Erbaugh, 52-74. Columbus, OH: Ohio State Foreign Language Publications and Services

Tseng, Ovid J. L., Daisy L. Hung, and William S-Y. Wang. 1977. "Speech Recoding in Reading Chinese Characters." *Journal of Experimental Psychology: Human Learning and Memory* 3, no. 6: 621-630.

Unger, James Marshall. 2004. *Ideogram*. Honolulu: University of Hawaii Press.

Vaezi, Shahin. "Theories of Reading." British Council. http://www.teachingenglish.org.uk/think/articles/theories-reading-1 and /theories-reading-2.

Wang Li 王力. 1982. *Tongyuan zidian* 同源字典. Beijing: Shangwu.

Wee Lee Woon. 1987. *Chinese Writing: Its Origin and Evolution.* Macau: University of East Asia

Wickelgren, W. A. 1974. "Single-Trace Fragility Theory of Memory Dynamics." *Memory and Cognition*, 2, no. 4: 775-780.

Wixted, John T. 2004. "The Psychology and Neuroscience of Forgetting." *Annual Review of Psychology* 55: 235-269.

Wixted, J. T., and S. K. Carpenter. 2007. "The Wickelgren Power Law and the Ebbinghaus Savings Function." *Psychological Science*, 18, no. 2: 133-134.

Wu, J. T. and Liu, I. M. 1997. "Phonological Activation in Pronouncing Characters," in *Cognitive Processing*. In *Cognitive Processing of Chinese and Related Asian Languages*, edited by Hsuan-Chi Chen, 47-64. Hong Kong: Chinese University Press.

Xu Jinxiong 許進雄. 1995. *Guwen xiesheng zigen* 古文諧聲字根. Taipei: Taipei Shangwu Yinshuguan.

Yonelinas, Andrew P., Leun J. Otten, Kendra N. Shaw, and Michael D. Rugg. 2005. "Separating the Brain Regions Involved in Recollection and Familiarity in Recognition Memory." *Journal of Neuroscience* 25, no. 11: 3002-3008.

Yoshimura Fumiko. 2000. "Automaticity Theory and EFL in Japan." *Literacy Across Cultures.* 4, no. 1: 1-11.

Zhang, Biyin and Danling Peng. 1992. "Decomposed Storage in the Chinese Lexicon." In *Language Processing in Chinese,* edited by Chen and Tzeng, 131-150. Amsterdam: Elsevier Science Publishers.

Zhao Cheng 趙誠. 1990. *Jiagu wenzixue gangyao* 甲古文字學綱要. Beijing: Shangwu Yinshu Chubanshe.

Zhu Jianing 竺家寧. 1991. *Shengyunxue* 聲韻學. Wunan Tushu Chuban Fufen Youxian Gongsi: Taipei.