

ABSTRACT

Title of dissertation: PHONOLOGICAL PROCESSING UNIT TRANSFER:
THE IMPACT OF FIRST LANGUAGE SYLLABLE
STRUCTURE AND ITS IMPLICATIONS FOR
PREFERRED SUBSYLLABIC DIVISION UNITS

Shih-wei Chen, Doctor of Philosophy, 2006

Dissertation directed by: Professor Mariam Jean Dreher
Department of Curriculum and Instruction

This study investigated the potential transfer of first language (L1) phonological processing unit to second language processing. English and Chinese phonology differ mainly in the complexity of their syllable structures. English phonology allows highly complex syllable structures, whereas Chinese has been characterized primarily as a core syllable language, i.e., its syllables typically consist only of a consonant and vowel (CV). This sharp contrast is hypothesized to entail different phonological processing units in the two languages, and to result in, through L1 transfer, the poor phonological awareness often observed in Chinese speakers learning English as a second language (ESL). This hypothesis was tested by examining the performance patterns of Chinese ESL fourth graders on phoneme deletion and phoneme isolation tasks. The results suggest that Chinese ESL children do seem to process an English syllable in terms of an intact core syllable plus its appendices due to L1 transfer. This gives support to a developmental account of subsyllabic division unit preference, which suggests that core syllable is universally preferred in the initial stages of language development, only after which

speakers of different languages diverge in their division unit preferences due to linguistic characteristics of their respective L1s.

The presence of transfer suggested that Chinese ESL children performed differently on two item types—core-syllable items (requiring segmentation of an element within the core syllable) and non-core-syllable items (requiring segmentation of any appendices from the core syllable). As phonological awareness involves the ability to segment cohesive sound units, it was hypothesized that only performance on core-syllable items should represent phonological awareness. This hypothesis was tested by analyzing the item types' respective contribution to decoding skills. Phonological awareness has long been established as a strong predictor of decoding skills; thus the analyses served to test the two item types' respective criterion validity in tapping phonological awareness. The results confirmed the hypothesis. This implies that, methodologically, phonological awareness of Chinese ESL children could be more reliably measured if, in future studies, only core-syllable segmentation items are employed. Educationally, instruction in phonological awareness might emphasize core-syllable segmentation, which alone appears to reflect Chinese ESL children's phonological awareness.

PHONOLOGICAL PROCESSING UNIT TRANSFER:
THE IMPACT OF FIRST LANGUAGE SYLLABLE STRUCTURE AND
ITS IMPLICATIONS FOR PREFERRED SUBSYLLABIC DIVISION UNITS

By

Shih-wei Chen

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee

Professor Mariam Jean Dreher, Chair
Professor C. Mitchell Dayton, Dean's Representative
Professor Linda Baker
Professor Ellen Broselow
Professor Marilyn J. Chambliss

© Copyright by
Shih-wei Chen
2006

Dedicated to
My Beloved **Mom** and **Dad**
in Western Pure Land of Ultimate Bliss;
My Dearest Bodhisattva-like **Wife**
and Best Friend,
Kwee-len;
and
My **Doggie Sons:**
Rambo in Pure Land
and
Gueiyifo in Kaohsiung

Acknowledgements

My wife said I was blessed to have a trustworthy advisor like Professor Mariam Jean Dreher. I couldn't agree more. I have said too many thanks to her, so I decided not to this time. Instead, I intend to express my gratitude to her by passing forward all the generosity and help she has never failed to offer me to my future students, so that they can be multiplied and amplified, all because of one origin. I want to thank Professor Ellen Broselow, who, along with Professor Christina Y. Bethin, helped me lay a solid foundation in phonology and phonetics when I was studying linguistics at SUNY, Stony Brook. I am grateful to Professor Mitchell Dayton for his always quick responses to my statistics questions. My sincere thanks also go to Professor Linda Baker for her kindly encouragement and insightful comments. I am appreciative of Professor Marilyn J. Chambliss for her practical suggestions, which may prove critical to whether I could make real interdisciplinary contribution in the future.

I do not think I can express in words my appreciation for the most important person in my life—my dearest wife and best friend, Kwee-len. Her wisdom is the power that has always lifted me up above troubled waters and carried me forward. Literally, I could not have obtained my Ph.D. without her. If wisdom were the criterion, she deserves this degree more than I do. I want to dedicate this dissertation to my late father and mother, who had been so loving to and supportive of me. How I wish they could have lived to see me receiving the degree, though I believe they are now in pure land tenderly watching me with pride.

I wish to thank also Professor Gregory Hancock, my minor advisor in Measurement, Statistics and Evaluation. He showed me the beauty of statistics and how a teacher could make a difficult subject so easily understood. My gratitude also goes to Professor Rebecca Oxford, who allowed me the freedom to follow my own academic interests. It was her kindness that had made me an interdisciplinary student.

I am indebted to the teachers and students in Kaohsiung Municipal Shin Juang Elementary School, whose assistance in providing the subjects and the library for testing was vital to this dissertation. My special thanks go to Ms. Chiou-ping Zheng, Dean of the Academic Affairs, who asked nothing for return in helping me arranging for the testing.

To me, all these gentle beings are Bodhisattva.

Table of Contents

Dedication.....	ii
Acknowledgements.....	iii
Table of Contents.....	v
Chapter I: Theoretical Rationale.....	1
Introduction.....	1
English and Chinese: a Brief Comparison.....	4
Research Questions.....	9
Chapter II: Literature Review.....	21
Introduction.....	21
Studies on Phonological Transfer and Phonological Processing Units.....	21
L1 Phonological Transfer Studies.....	21
Core Syllable as a Linguistic Universal.....	27
Syllable as a Phonological Processing Unit in Chinese.....	29
Sub-syllabic Unit Studies.....	32
Introduction.....	32
Onset-rime Studies.....	34
Body-coda Studies.....	39
Methodological Issues.....	45
Introduction.....	45
Sonority.....	47
Global Similarities.....	50
Vowel Length.....	51
Overview of the Present Study.....	53
Introduction.....	53
The Study.....	53
Use and Creation of Measures.....	53
Predictions.....	58
L1 Processing Unit Transfer.....	59
Nature of the Segmentation Performance Based on Core-syllable Awareness....	67
Chapter III: Methodology.....	69
Participants.....	69
Measures.....	69
Oral Proficiency.....	70
Nonverbal Intelligence.....	71
Decoding Skills.....	71
Phoneme Deletion Task.....	72
Phoneme Isolation Task.....	73
Procedure.....	74
Coding.....	76
Design and Data Analyses.....	78
Chapter IV: Results and Discussion.....	81
Introduction.....	81
L1 Processing Unit Transfer.....	82
Results from Phoneme Deletion Task.....	82

Results from Phoneme Isolation Task.....	84
Discussion: L1 Transfer of Phonological Processing Unit.....	86
Discussion: Implications for the Preferred Subsyllabic Division Units Debate.....	92
Implications for Future Studies.....	97
Nature of the Segmentation Performance Based on Core-syllable Awareness.....	99
Results from the Multiple Regression Analyses.....	99
Discussion.....	103
Chapter V: Conclusion.....	105
Introduction.....	105
Theoretical Implications	107
Methodological Implications.....	110
Educational Implications.....	111
Appendix A: Phoneme Deletion Items.....	114
Appendix B: Phoneme Isolation Items.....	115
Appendix C: Similarity Judgment Items.....	116
Appendix D: Prediction Tables.....	117
Appendix E: Glossary of Linguistic Terms.....	121
Appendix F: An Optional Study.....	122
References.....	126

Phonological Processing Unit Transfer:
The Impact of First Language Syllable Structure and
Its Implications for Preferred Subsyllabic Division Units

CHAPTER I: THEORETICAL RATIONALE

Introduction

It has been well established that phonological awareness is a strong predictor of success in beginning literacy in alphabetic languages (see Adams, 1994; Ioup, 1984, for a review; see also Castles & Coltheart, 2004, for critiques as well as reviews of the assumed causal links). The predictive power is not confined to phonological awareness expressed in the native language. Cross-language studies have shown that phonological awareness as measured in the first language (L1) is also predictive of reading acquisition success as well as phonological awareness in a second language (L2; e.g., Cisero & Royer, 1995; Durgunoglu, Nagy, & Hancinbhatt, 1993; Lindsey, Manis, & Bailey, 2003). Such positive transfer, however, is more often observed when both the L1 and L2 involved are alphabetic languages (e.g., Bialystok, Luk, & Kwan, 2005; Comeau, Cormier, Grandmaison, & Lacroix, 1999). Although there is research indicating that positive transfer may also occur between non-alphabetic L1 and alphabetic L2 (e.g., Gottardo, Yan, Siegel, & Wade-Woolley, 2001), the mainstream of studies showed that cross-language transfer occurred only between languages sharing similar writing systems but not between alphabetic and non-alphabetic languages (e.g., Bialystok et al., 2005). This is because, unlike alphabetic L1 speakers, non-alphabetic L1 speakers have long been found to be poor in phonological awareness.

The differences in phonological awareness, hence transfer effects, have been traditionally attributed to L1 orthographies. The writing systems of non-alphabetic languages such as Chinese provide no phonological information at the subsyllabic levels

(i.e., the phonological components of a syllable, such as phonemes or onsets and rimes), which is fundamental to the insight into the language's finer phonological structure. Knowledge of an alphabet, on the other hand, has been demonstrated to be critical to the development of phoneme awareness, the finest and final level of phonological awareness. It is generally agreed that phonemic awareness only comes with the onset of alphabetic literacy (see Ziegler & Goswami, 2005, for a review). Lack of alphabetic orthographies in non-alphabetic languages has therefore been held responsible for their speakers' poor phonological awareness, hence decoding skills, in an alphabetic L2.

More recently, however, it has been suggested that the L1 phonology underlying the particular forms of L1 writing system may deserve much greater, if not the sole, attention in accounting for the observed performance differences in L2 decoding skills as well as phonological awareness (e.g., Chen, Wang, & Cheng, 2005; Yamada, 2004; Yoon, Bolger, Kwon, & Perfetti, 2002; see also Ziegler & Goswami, 2005). Spoken languages, after all, exist before their writing systems, as there are, among natural languages, no writing systems without spoken language but there are spoken languages without writing systems. In other words, it is usually the sound system (i.e., what we hear and say) that determines the writing system (i.e., what we write and read) but not vice versa. Studies on the impact of L1 orthography thus cannot rule out the confounding variable of L1 spoken language (Wang, Koda, & Perfetti, 2004; Yamada, 2004) and certain L1 impacts may actually be attributable to L1 phonology rather than to L1 orthography (Chen et al., 2005; Yamada, 2004). In the second language acquisition (SLA) literature, indeed, the impact of L1 phonology on L2 phonology acquisition has been the least controversial compared to other areas of linguistics. The traces of L1 phonology in SLA are so

apparent that they have been accepted “even by those who doubt L1 influence in the area of syntax” (Eckman, 2004, p. 515).

The present study began as an attempt at exploring the phonological factors¹ that could underlie the limited phonological awareness of speakers of L1 Mandarin Chinese (Chinese hereafter), a non-alphabetic language, who are learning English, an alphabetic language, as a second language. It follows part of the old but enduring Contrastive Analysis Hypothesis (CAH; Lado, 1957)², which predicts that, among other things, what is different between L1 and L2 is also what poses difficulties to L2 learners. Two often studied areas of differences between Chinese and English have been their writing systems and their phonologies (i.e., sound systems). As briefly mentioned earlier, the English writing system is alphabetic, whereas Chinese is non-alphabetic and has been characterized by some as logographic for the picture-like quality of its basic writing units, i.e., characters. Studies on the impact of different L1 writing systems on the visual and phonological processing of L2 words or nonwords have been especially fruitful in the past decades (e.g., Akamatsu, 2003, 1999; Cheung & Chen, 2004; Cheung, Chen, Lai, Wong, & Hills, 2001; Holm & Dodd, 1996; Koda, 1988, 1999, 1989; Muljani, Koda, & Moates, 1998; Read, Zhang, Nie, & Ding, 1986; Wang & Geva, 2003; Wang, Koda, & Perfetti, 2003).

Research on the impact of L1 Chinese phonology on L2 English phonological awareness and decoding skills, on the other hand, has been strikingly scant, if not non-

¹ As alphabeticity of the writing systems has traditionally been considered a major influence on phonological awareness in reading literature, languages are usually classified as alphabetic or non-alphabetic. As argued earlier, however, the alphabeticity of the writing system has its root in the language’s sound systems. The concern in this study is therefore with pronunciation, not spelling.

² Though researchers have nowadays been more interested in language universals, including the popular Optimality Theory (OT), in explaining L2 phonology, such theories usually complement but not contradict CAH.

existent. One of the critical phonological differences between English and Chinese lies in their divergent syllable structures: while English phonology allows syllables of highly complex structure, Chinese has a very limited inventory of possible syllables. As will be detailed below, this sharp contrast in syllable structures is suggested to entail different phonological processing units in the two languages. Such unit differences are hypothesized to be the major phonological factor responsible for the often-observed poverty of Chinese speakers' L2 English phonological awareness, hence decoding skills. Implications of findings from this study are manifold. They concern beginning reading instruction in English as a second language (ESL), phonological awareness measurement in Chinese ESL populations, and the debate over preferred subsyllabic division units—a heated topic in current reading psychology literature. Three specific research questions concerning, respectively, phonological processing unit transfer, its theoretical implications for preferred subsyllabic units, and its practical implications for measurement and instruction are raised. As all three questions are motivated by the phonological contrasts and the resulting differences in phonological processing units between Chinese (the L1) and English (the L2), it is necessary to give a brief sketch of such contrasts before detailing the research questions and their theoretical background.

English and Chinese: A Brief Comparison

As far as phonology is concerned, English and Chinese differ mainly in the complexity of their syllable structures³. In linguistic theories, a syllable is usually

³ It is important to clarify a possible source of confusion. As we are dealing here solely with pronunciation, not spelling, we are using short forms like C (for consonant) and V (for vowel) to represent individual component **sounds**, not the spelling, of a syllable. Thus a syllable of the structure “CVC” is exemplified by words like *both* and *boat*, which are phonemically transcribed as /boθ/ and /bot/, but not words like *bee* or *she*, both of which are phonetically CV and transcribed as /bi/ and /ʃi/. Moreover, as spelling in English has been notorious for the poor correspondence between the alphabetic letters and the sounds they represent,

hierarchically represented in a binary manner. A syllable can be first subdivided into the *onset*, i.e., the optional prevocalic consonant or consonant cluster, and the *rime*⁴, i.e., the obligatory vowel plus the optional ensuing consonant or consonant cluster⁵. The rime unit can be further divided into the vowel *nucleus* and, if any, the postvocalic *coda*. Thus given a syllable of the consonant-vowel-consonant sequence (C₁VC₂, as exemplified by /boθ/, *both*, or /bot/, *boat*), for example, the first consonant C₁ (as, in the example of *both*, the sound represented by the *b*) would be analyzed as the onset and the remaining elements VC₂ (for the same example, the remaining two sounds represented by, respectively, -o- and -th) as the rime. The rime VC₂ can be further analyzed as consisting of the nucleus V (i.e., the *o*) and the coda C₂ (i.e., the *th*). One way of classifying syllable types is based on the presence or absence of the coda: a syllable is called *closed* if it has a coda (e.g., /med/, *made*); a syllable is *open* if it does not end in any consonants (e.g., /me/, *may*). As we shall see, this classification marks one important distinction between Chinese and English.

In terms of the permitted size of a syllable, English and Chinese can be said to lie near the two extremes of syllabic structure complexity. In English, consonant clusters are

phonetic scripts (which maintain a strict one-to-one correspondence between the symbols and the sounds they represent) are used throughout this paper.

⁴ Note that our use of rime here is limited to sounds only. The word “rime” is sometimes used interchangeably with “rhyme”, and the distinction between the two is not always strict. In one account, in linguistics, “Rime” and “rhyme” are variants of the same word, but the rarer form “rime” is sometimes used with the definition given [for phonology] in order to differentiate it from the concept of poetic [rhyme](#). This distinction is not made by all linguists and does not appear in most dictionaries” (Syllable rime. 2006, June 25. In *Wikipedia, The Free Encyclopedia*. Retrieved 10:18, July 9, 2006, from http://en.wikipedia.org/w/index.php?title=Syllable_rime&oldid=60453855). In reading literature, the use of the two forms is more confusing, referring sometimes to the sounds and sometimes to the orthography of words. Goswami, Gombert, and Barrera (1998), for example, appeared to use “rime” and “rhyme” interchangeably. “Rhyme,” for example, is used in the abstract: “Exp 1 compared English and French children’s reading of nonsense words that shared **rhyme orthography** with real words ... Exp 2 compared English and French children’s reading of nonsense words that shared **rhyme phonology** with real words” (p.19; the boldface is mine). Elsewhere, in the body of the paper, “rime” is used: “A third type of nonsense word, such as *faish* (English), *chenfe* (French), and *saf* (Spanish), had neither **orthographic rime** neighbors nor **phonological rime** neighbors in the mental lexicon” (p. 22; the boldface is mine).

⁵ A glossary of linguistics terms used in the present study is provided as Appendix E.

possible, and can actually be very complex, in both the onset and the coda positions (Goldsmith, 1990; Spencer, 1996). An English onset allows up to three consonants (as the beginning sound sequence of /spr/ in the word /spre/, *spray*); and an English coda, up to four consonants (as the ending consonant cluster /lpts/ in the plural-form word /skʌlpts/, *sculpts*). In contrast, a Chinese syllable allows no consonant clusters at all in either the onset or the coda position (Anderson, 1987; Hansen, 2001)⁶. The contrast appears even sharper when we consider the fact that, despite its presence in a Chinese syllable, the coda is limited to two nasal phonemes only: namely, /n/ and /ŋ/, as the respective final sound of *kin* and *king*. Nasal phonemes, along with liquids (e.g., /r/ as in *bear* and /l/ as in *tell*), glides (e.g., /w/ as in *wore* and /j/ as in *your*), and vowels, have been classified as *sonorants*, for the production of these sounds involves lesser degrees of air constriction in the vocal tract than all other consonants. They are thus more *vowel-like* and have been demonstrated to show a stronger cohesion, when in the coda position, with the preceding nucleus vowel (to form the subsyllabic unit of rime) than do *obstruents* (Geudens & Sandra, 2003; Treiman, 1995, 1984; Treiman, Zukowski, & Richmondwelty, 1995; Yavas & Core, 2001; Yavas & Gogate, 1999 for sonority effect in the onset position). In contrast to sonorants, obstruents are defined by the complete or narrow constriction of air in the vocal tract when producing the sounds. This class thus includes both *stops*

⁶ It has been generally agreed that Chinese syllable maximally consists of four elements, (C)(G)V(C), that is, an optional initial consonant, an optional glide (a glide is defined as a phonological category that resembles vowels in every way except for its position in a syllable; that is, unlike a vowel, a glide cannot form the nucleus by itself), an obligatory vowel nucleus, and an optional nasal coda. The subsyllabic division of a Chinese syllable, however, is much more controversial. Researchers following the traditional partition divide a Chinese syllable into an initial (the beginning consonant) and a final (the remaining part, including the glide). Researchers supporting a modern syllable parsing, on the other hand, argue for an onset-rime division, with the glide included as part of the onset. In either case, glide is seldom considered a phoneme by itself. Even with the modern parsing, for example, the glide has been argued to be a secondary articulation to the initial consonant, in a sense similar to the nasal quality of a consonant with nasalization. The CG sequence is thus considered one, instead of two, phonemes and usually transcribed with a superscripted G, i.e., C^G. See Li (1999) for a review.

(complete constriction phonemes), such as /p/, /t/, /k/, and *fricatives* (narrow constriction phonemes), such as /f/ and /z/, both of which are least vowel-like in terms of air constriction. As the coda in the Chinese syllable is limited in both number (allowing only two candidate phonemes) and nature (both candidate phonemes being the more vowel-like nasals), Chinese has been characterized as “a predominantly open syllable language” (Anderson, 1987, p. 280) in spite of the presence of coda, which defines a closed syllable.

Given such differences in syllable structure, it should not be surprising if native speakers of these two languages should develop divergent cognitive strategies for more efficient processing of native speech sounds. The complexity of syllabic structure itself, to begin with, can directly translate into the size of the syllable repertoire: there are, for example, only about 420 syllables in the Chinese syllable repertoire and the number is increased to only about 1300⁷ even when the permutation of Chinese tones⁸ is considered (Chen, 1999; Hoosain, 1991). Perhaps more importantly, Chinese has been characterized as *morphosyllabic* in that a Chinese morpheme is usually represented by only one syllable. This entails the abundance of homophones, for which Chinese has been notorious, since only 420 syllables are available to represent the 40,000 to 80,000 words in an ordinary Chinese dictionary (Jia & Zhang, 2001). An English morpheme, in contrast, can be multi-syllabic to the extent that there is no actual limit to how many syllables a word can consist of. The complex syllabic structure plus the possibility of multi-syllabic words thus implies that each and every English morpheme can be,

⁷ The number 420 will be used throughout the remaining paper, as syllables as the phonological processing unit of Chinese have been shown to exclude tones (e.g., Chen, 2000; Chen et al., 2002), which are attached to the rimes.

⁸ In Chinese, tones are the various pitch contours in producing a syllable, and make meaning distinctions. For example, while the syllable /ma/ in the first tone, sometimes indicated as /ma1/, means *mother*, the same syllable in the third tone, /ma3/, means *horse*.

theoretically, represented by a unique combination of speech sounds. Given the sharp contrast in the numbers of available syllables, it follows naturally that a Chinese syllable, with only 420 variations, could well be processed simply as a holistic unit (Chen, 2000; Chen, Chen, & Dell, 2002; Chen, Lin, & Ferrand, 2003), while an English syllable could be more easily processed if it is further divided into subsyllabic chunks such as onsets and rimes.

A Chinese syllable could be processed as a holistic unit because each of the 420 *syllables* has to be *recycled* in order to represent the tens of thousands of Chinese words. In other words, most Chinese syllables would enjoy extremely high frequencies of use, which may have made syllables especially salient in Chinese (Chen, Dell, & Chen, 2004, August). What is demanded of a young native speaker of Chinese in learning new words, as a result, is more often the ability to *associate* the old, familiar morph-syllables, possibly already in store in the mental syllable repertoire, with the new meaning of the new lexical items. Acquisition of vocabulary in English, on the other hand, usually requires greater sensitivity to the sound structure of spoken words (Treiman, 2000), as each and every new word is theoretically a new pattern of sounds. Recycling is also possible in English, but what is recycled is, at best, *part* of the word, such as the onsets and the rimes in the cases of alliteration and rhyming. In contrast to Chinese, therefore, it is more likely the subsyllabic units rather than the syllable in its entirety that is the psychological unit of phonology for English speakers. As a result, sensitivity to syllable subdivision may be necessary for English speakers but *redundant* for their Chinese counterparts.

Indeed, empirical studies using speech error analysis (e.g., Chen, 2000) and masked or implicit priming tasks (e.g., Chen et al., 2002; Chen et al., 2003) tapping speech production in Chinese have all supported the syllable as the most salient phonological processing unit in Mandarin Chinese. In a recent study using simple recurrent networks, a connectionist model, syllable was also found to be more salient in Chinese than in English (Chen et al., 2004, August). Given this backdrop of the processing unit differences between English and Chinese, certain specific forms of, probably negative, L1 transfer should be expected in Chinese-speaking children learning English as an L2 (Chinese ESL children, hereafter). After all, sensitivity to subsyllabic structure is not so critical to learning Chinese as it is to learning English, and lack of such sensitivity should be more of a rule than an exception in Chinese ESL children. In addressing the issues of phonological processing unit transfer, accordingly, the following three research questions can be put forth. As we shall see presently, they concern, respectively, (1) the very existence of such transfer, (2) its theoretical implications to the current debate over preferred subsyllabic division units, and (3) its educational and methodological implications.

Research Questions

Research questions one and two are closely related in that while the first concerns the presence of processing unit transfer, the other concerns its implications in terms of subsyllabic division unit preference.

(1) *Do Chinese ESL children transfer their L1 phonological processing unit in the processing of L2 spoken syllables by analyzing an English syllable as consisting of a core syllable plus its appendices?*

Given that Chinese allows no consonant clusters in either the onset or coda position, a Chinese syllable can be further characterized as consisting predominantly of *core syllables*, i.e., CV—open syllables that are made up of only one optional singleton onset plus the vowel nucleus (see Note 6 for debates over the subsyllabic division of Chinese syllables). Given CV as the predominant Chinese syllable structure, if Chinese ESL children do transfer their L1 phonological units in their processing of L2 English syllables, the following performance patterns should be observed. (A) Given a CVC syllable, to begin with, isolation or deletion of the initial consonant should be more difficult than isolation or deletion of the final non-nasal consonant. This is because while the initial consonant is *internal* to the processing unit of CV and hence more *cohesive* to the nucleus, the final consonant is *external* to it. This logic has been successfully employed to argue for the natural subsyllabic division units of onsets and rimes in native speakers of alphabetic languages like English and Dutch (e.g., Bertelson, de Gelder, & van Zon, 1997; Schreuder & van Bon, 1989; Stahl & Murray, 1994; Treiman, 1985; Treiman & Weatherston, 1992). In these studies, onsets and rimes were argued to be the preferred subsyllabic division units. Thus taking a CVC syllable for example, subjects were expected to more likely segment it as C-VC (i.e., the onset plus the rime, where the hyphen indicates the point of segmentation). However, in the current study, the same logic would predict that a CV-C segmentation would result, because here the CV is hypothesized as the more cohesive unit rather than the rime, i.e., the VC, as a result of L1 transfer.

Even if the initial consonant proves to be more difficult to isolate or delete, however, we still need to rule out another competing interpretation of a preference for the

hierarchical *body-coda* division (e.g., Derwing, Yoon, & Cho, 1993; Geudens & Sandra, 2003; Yoon et al., 2002; Yoon & Derwing, 2001), where *body* designates the prevocalic consonant or consonant cluster plus the vowel nucleus. That is, native Chinese-speaking children may alternatively demonstrate a preference to hierarchically subdivide a syllable into units of bodies and codas. This differs from the core-syllable hypothesis in which Chinese ESL learners are expected to treat the core syllable CV *in its entirety* as the processing unit, which entails that whatever comes before or after the core syllable should be processed only as linear (non-hierarchical) *appendices*. To rule out the rival interpretation of body-coda preference, more item types are needed for comparisons capable of distinguishing the two theories. Thus (B) given the isolation or deletion of the two item types CV(C) and CVC(C), where consonants in parentheses designate the sounds to be segmented, we should expect the target sound in CV(C) items to be more easily segmented than that in the CVC(C) items, if bodies and codas are what is actually preferred as subsyllabic division units. This is because, given the coda as a unit corresponding to the body, it should be easier to break up *between* the body and the coda (i.e., [CV]-[C]) than *within* the coda (i.e., [CV][C-C], where brackets indicate the hypothesized subsyllabic units of bodies and codas and hyphen, the point of segmentation). On the other hand, if it is core syllable in its entirety plus its appendices at work, we should expect little, if any at all, difference in segmentation difficulty between the two items types. This is because both target sounds in the two item types are external to the core syllable; that is, both are appendices and should be easily segmented to a similar degree. Most importantly, however, (C) given the item types of (C)VC and (C)CVC, we should expect the former to be more difficult to segment than the latter, if

the core syllable transfer hypothesis holds. Again, this is because while the target sound in the former (i.e., [(C)V][C]) is internal to the core syllable, that of the latter (i.e., [(C)][CV][C]) is external. On the other hand, if it is bodies and codas at work, we should expect the opposite results: [(C)V][C] to be more easily segmented than [(C)CV][C].

Results obtained from tasks examining processing unit transfer can be further examined against the larger theoretical background of preferred subsyllabic division unit debate, which is expressed in research question two:

(2) Are the findings consistent with the developmental account of preferred subsyllabic unit division?

Both traditional parsing in theoretical linguistics and research in emergent literacy skills with alphabetic languages have suggested onsets and rimes to be the preferred subsyllabic division units across languages (e.g., Bradley & Bryant, 1983; Gombert, 1996; Kirtley, Bryant, Maclean, & Bradley, 1989; Schreuder & van Bon, 1989; Treiman, 1986, 1995, 1983, 1985; Treiman & Danis, 1988; Treiman & Kessler, 1995; Uhry & Ehri, 1999). Evidence supporting onset-rime segmentation is especially abundant in research with English-speaking subjects, as attested by various tasks: word games (e.g., Treiman, 1983, 1985, Exp. 1; Treiman, 1986), similarity judgment (e.g., Treiman, 1991), odd man out (e.g., Bradley & Bryant, 1983; Kirtley et al., 1989), phoneme isolation (e.g., Stahl & Murray, 1994; Treiman, 1985; Treiman & Weatherston, 1992), spelling errors (e.g., Treiman, 1991), short-term memory errors (e.g., Treiman, 1995), and syllabification studies (e.g., Treiman, Bowey, & Bourassa, 2002). In addition to the psycholinguistic reality, rime as a natural subsyllabic unit has also found its basis in the statistical properties of the languages in question as well. Frequency studies of phoneme

collocation, for example, point to the rime as the subsyllabic units as it has, compared to other rival combinations, the most phonological neighbors. Such dominance of rime as attested by statistical properties has been borne out not only in English (e.g., De Cara & Goswami, 2002; Ziegler & Goswami, 2005) but in German, Dutch, and French as well (see Ziegler & Goswami, 2005, for a review).

Recent studies in Dutch and Korean, however, have cast doubt on this universality implication. Some studies with Dutch speakers (e.g., Bus, 1985; Geudens & Sandra, 2003) and all studies with Korean speakers (e.g., Derwing et al., 1993; Yoon et al., 2002; Yoon, 2001) have found that a CV-C segmentation is easier than a C-VC (i.e., onset-rime) segmentation in subdividing a syllable. An important question immediately follows with the non-onset-rime division findings: What is the nature or structure of such a CV-C division tendency? As the items used in the related studies were restricted to syllables of simple structures only, i.e., CVC, CV, or VC, the two rival interpretations—core-syllable vs. body-coda division preferences, as discussed earlier—are equally possible. More specifically, the final C in either the VC or CVC syllable could be either a hierarchical subsyllabic unit by itself (i.e., the coda for a body-coda hypothesis) or simply a sound linearly attached to the core syllable (i.e., an appendix for the core-syllable hypothesis).

While little has been elaborated in related literature on the construction of the body-coda division hypothesis, except for its different division tendency from that of onset-rime hypothesis, literature consistent with the hypothesized preference for core syllables has been abundant, usually as an integral part of a developmental account. Geudens and Sandra (2003), for example, suggest that CV-C division and onset-rime division represent the preferred divisions of different developmental stages. CV syllables

are preferred at a younger age because the vocabulary acquired by very young children is comprised mostly of core syllables (i.e., CV). With development, as in the case of English speakers, the more complicated syllable structures coming along with the increased vocabulary size may force the children to restructure their syllables so as to make onsets and rimes their preferred units in reflection of the statistical properties of the language (De Cara & Goswami, 2003; Walley, 1993). Following the same reasoning, then, Chinese speakers should have core syllables as their preferred syllable structure, since the language itself does not provide instances of more complex syllables to motivate the syllable restructuring as Geudens and Sandra (2003) argued for Dutch children.

In alphabetic languages with complex syllable structure, however, whether the more ready CV-C segmentation reflects a preference for core syllables or for body-coda division is hard to distinguish in practice. For one, as several prior studies have shown (e.g., Geudens & Sandra, 2003), tasks used for tapping division preference are often too highly demanding cognitively. As items required to disambiguate the preference nature are necessarily more complex linguistically (involving at least 4-phoneme instead of the 2- or 3-phoneme syllables used in the prior studies), hence demanding even higher cognitive abilities, it is very difficult to administer such tasks on young children—even if we are able to judge when a child would begin his/her syllable restructuring, which is itself yet another experimental obstacle to overcome. Chinese ESL children, on the other hand, due to the predominantly core syllable structure of their L1 and the complex syllable structure of their L2, may serve as an approximate equivalent of young speakers of alphabetic languages. That is, given the dominance of core syllables in Chinese,

Chinese-speaking children's syllable structure should remain little changed even with growth. Their core-syllable structure being certain, the more cognitively demanding tasks can then be administered on older children to test for syllable subdivision preference.

Preliminary support of the core-syllable hypothesis explored here comes from an analysis of the phoneme deletion data from the author's prior study (Chen et al., 2005). The study was conducted with 4th and 6th graders from both Taipei and Beijing. One of the tasks, phoneme deletion, comprised 20 items of three- or four-phoneme monosyllabic nonwords, which were categorized as three combination types: CVC, CCVC, and CVCC. It was found that, given the same CVC type nonwords, the deletion of the ending consonant, i.e., CV(C),⁹ is much easier (with a mean accuracy rate of 86%) than deletion of the beginning sound, i.e., (C)VC (mean accuracy 64%). This ruled out the possibility of an onset-rime division preference. When nonwords involving consonant clusters (i.e., CVCC and CCVC) were taken into account, furthermore, it was found that while deletion of the target sound in a singleton-consonant onset (i.e., (C)VC; mean accuracy 64%) is more difficult than deletion of the target sound in a clustered-consonant onset (i.e., (C)CVC; mean accuracy 86%), a corresponding contrast in the coda position shows that the CV(C) items (86%) were not more easily segmented than the CVC(C) items (93%). Both results would argue against a body-coda preference, which would predict totally different comparison results. Similar patterns re-emerge when separate analyses in terms of both age (4th and 6th graders) and region (Taipei and Beijing) were conducted. The core syllable transfer hypothesis thus appears tenable.

⁹ There were two CV(C) items in the task. As one of them ended in a nasal, which has been considered to be more cohesive than obstruents to the preceding vowel, only the other item is considered.

However, as the items used in the Chen et al. (2005) study consist mainly of obstruent codas, it would be desirable to see if the same patterns would also occur on items with sonorant codas. *Sonority* (the degree of resonance in the vocal tract) has been shown to confound interpretations of subsyllabic division unit preferences in earlier studies (Geudens & Sandra, 2003). In general, in producing a speech sound, the lesser the air constriction in the vocal tract, the higher the sound's degree of sonority. Thus, according to Selkirk (1982), sound categories can be rank ordered in terms of sonority: vowels > glides > liquids > nasals > fricatives > stops. In dealing with the subsyllabic division units, it has been found that the degrees of difficulties in segmenting a coda consonant also depend on the sonority of the consonant in question (e.g., Geudens & Sandra, 2003; Treiman, 1984, 1992; Treiman et al., 2002; Treiman et al., 1995). This is especially true when the consonant is in the coda position (e.g., Treiman, 1984; Treiman & Cassar, 1997). Generally, sonorant codas (e.g., the *l* sound of the word /dɛl/, *dell*) are usually more cohesive to the preceding vowel than obstruents (e.g., the *t* sound of the word /dɛt/, *debt*), and are thus more difficult to segment. Geudens and Sandra (2003), for example, have shown that the conflicting results of two previous studies, i.e., Bus, 1985 and Schreuder & van Bon, 1989, may have actually lain in the differences in coda sonority of the items employed. In the present study, sonority of the test items concerning subsyllabic units is, therefore, controlled for by including an additional within-subject factor of sonority with four levels: liquids, L1 nasals, non-L1 nasals, and obstruents. The further division of nasals into L1 nasal and non-L1 nasal is because Chinese syllables do have nasal codas. Thus the expected greater degrees of segmentation difficulties with nasals could be further confounded by the existence of the

two possible nasal codas in Chinese. After all, as part of the intact syllable, the two possible nasal codas should be expected to be at least as difficult as the onset in terms of segmentation. By subdividing nasal items in terms of their legitimate presence in L1 Chinese, it was hoped that the sonority effect can be distinguished from L1 effect.

If Chinese ESL children do transfer their L1 processing unit in their processing of phonological awareness tasks, a further methodologically and educationally important question should be asked:

(3A) *Does segmentation performance as a result of L1 processing unit transfer represent true phonological awareness by making significant contribution to decoding skills?*

If, as we have expected, the L1 transfer results in the processing of L2 English syllables in terms of a core syllable plus its appendices, performance in segmenting the core syllable from its appendices (*non-core-syllable segmentation* hereafter) should not reflect true segmentation ability, as the performance in C-V segmentation (henceforth *core-syllable segmentation*) does. If this prediction proves true, we should expect performance on non-core-syllable segmentation items to make no contribution to English decoding skills. However, if the results do not support the prediction, it may be that any sensitivity to sound segmentation would contribute to initial reading skills. Results of the author's prior study (Chen et al., 2005), however, appear to favor the non-influential expectation, as accuracy rates for non-core deletion items are close to ceiling: 86% for (C)[CV][C] items, 90% for [CV][C][(C)] items, and 86% for [CV][(C)] items. The relatively low accuracy rate (73%) of the [CV][(C)][C] items may appear as an exception. This, however, may very likely reflect the effect of its non-marginal position, which requires more complex cognitive manipulation compared to the other three.

A similar question can be asked, if the results do not support the core-syllable preference hypothesis but instead favor a preferred body-coda division unit interpretation.

(3B) *Does segmentation performance as a result of a preferred body-coda division tendency represent true phonological awareness by making significant contribution to decoding skills?*

Even with body-coda results, this question would still be methodologically and educationally important in deciding whether body-coda division, if any, reflects a similar ability as that of phonological awareness and whether it is conducive to or impeding reading acquisition. This question, however, has never been brought up in related studies.

In order to answer the above three research questions, Chinese-speaking fourth graders in Kaohsiung, Taiwan were recruited. Two considerations informed the subject selection. In second language acquisition theory, L1 transfer effect is most obvious for beginning L2 learners (e.g., Major, 2001). This is especially true to L2 as a foreign language environment, where exposure to the L2 is extremely limited. As Lado (1957) put it, “we have ample evidence that when learning a foreign language we tend to transfer our entire language system in the process” (p.11). A transfer effect is thus more likely to be observed with young children in an English-as-a-foreign-language environment. However, the subjects cannot be too young, either. They must be old enough to meet the high cognitive demands of the phonological tests just described on the one hand and, on the other, they must be old enough to have been sufficiently exposed to English even if in a classroom setting. As students in Taiwan begin English learning between first and third grade, depending on policies of local governments, fourth graders appear to better meet the conditions considered above.

One potential confound with recruiting elementary students in Taiwan, however, is that they also learn to read Zhuyin, a phonetic script system taught to help with smooth transition to Chinese character reading, in the very first semester of their elementary years. Given the predominantly CV structure of Chinese words, which are always monosyllabic, Zhuyin is taught necessarily by segmenting a spoken word into its onset and rime (i.e., C-V); its instruction is thus a potential confound. To rule out this potential confound, an optional study had also been prepared in the initial construction of this study, just in case such confound should be suspected of any significant impact. The optional study, as originally planned, would have been conducted on kindergarteners to rule out the impact of Zhuyin instruction. As the results turned out to accord with the predictions made for the present study, the optional study was not needed and hence not conducted. It is nevertheless described in Appendix F.

In summary, based upon the phonological contrasts between the alphabetic English and the non-alphabetic Chinese, three specific research questions are raised concerning, respectively, the presence of L1 transfer of phonological processing unit, its theoretical implications for preferred subsyllabic division units, and the relationship between the resulting segmentation performance and phonological awareness proper. The contribution of this dissertation study is three fold: educational, methodological, and theoretical. Educationally, the transfer of intact core syllable, if any, would suggest the pedagogical need to sensitize learners to the sound structure specific to the English language, such as the preferred onset-rime segmentation, in ESL literacy instruction. Reading-related skills assessment could employ tasks with different syllable types for diagnosis of subsyllabic division tendency, i.e., the educationally desirable onset-rime

division or the problematic core-syllable preference, which would then serve as index to the necessary educational remedy or reinforcement. Methodologically, tasks tapping phonological awareness could be more precise, hence more reliable, by excluding items tapping non-phonological-awareness segmentation skills, such as core-appendices segmentation. This is especially important with studies involving subjects speaking predominantly core-syllable L1s, such as Chinese and Japanese or even Korean, in an L2 reading acquisition context. Theoretically, the results could serve to confirm or disconfirm the core-syllable hypothesis, hence the developmental account. Core syllable, that is, may be the universally preferred phonological units early in development, but the road divides after the early stage. For speakers of languages whose syllable structures are extremely simple, such as Chinese, core syllables may remain their preferred syllable type. For speakers of other languages whose statistical properties inherent in their phonological structures encourage an onset-rime division, the chunking of a speech sound stream into onsets and rimes may be natural. Results of the current study can provide empirical evidence for or against the developmental account.

CHAPTER II: LITERATURE REVIEW

Introduction

This chapter reviews research related to the issues examined here and the literature important to experimental controls. It begins with a discussion of the L1 phonological transfer research, plus a review of studies concerning core syllables, both as a linguistic universal and as the phonological processing unit in Chinese. The debate over the preferred subsyllabic division units and the critical arguments of each camp are then reviewed. As some of the conflicting results have been shown to have resulted from failure to control for certain linguistic properties of the items used, literature concerning methodological issues, including the potential confounding factors of sonority, global similarities, and vowel length to be incorporated into the present study, is then examined for its importance in ruling out rival interpretations. The chapter ends with an overview of the present study.

Studies on Phonological Transfer and Phonological Processing Units

L1 Phonological Transfer Studies

One of the earliest theories about L1 impact is based on contrastive studies of L1-L2 similarities and differences, as expressed by Lado's (1957) often-cited Contrastive Analysis Hypothesis (CAH):

We assume that the student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be simple for him, and those elements that are different will be difficult. (p. 2)

CAH was proposed to explain the systematic pronunciation errors often observed in L2 learners speaking a given L1, and part of the reason for its popularity, as it seems, is the apparent L1 influences felt in L2 pronunciation. Indeed, while L1 impact has been

questioned in other areas of linguistics such as syntax, its influence on L2 phonology “has never been seriously contested by researchers in SLA [second language acquisition] theory” (Eckman, 2004, p. 515). Ioup (1984), for example, presented two groups of native speakers of English with a composition written by either Spanish or Hebrew ESL learners or a recorded speech taped by two other groups of Korean or Arabic ESL learners. Subjects were then asked to identify which of the texts or speeches belonged to authors/speakers of the same language group. As the results showed, the L1 English judges could identify people from the same language group only when phonological evidence was available.

The predictions of CAH, however, are much less reliable when the elements are similar in the two languages, which according to CAH should be easier. Later research on L1 impact has found that errors can occur even where L1 and L2 are similar (e.g., Flege, 1987; Wode, 1978). Flege (1987), for example, examined French and English speakers for their pronunciation of sounds that are either similar (but not identical) or different in the two languages. It was found that the subjects had more difficulties correctly pronouncing sounds that are similar in the L1 and the L2. The pronunciation of L2 sounds that are different from the L1, on the other hand, is more accurate. Obviously, similar sounds may actually be more confusing than different sounds and cause more difficulties in native-like pronunciation. Flege argued that this is because a new category is usually created for L2 sounds that are absent in the L1 so that less confusion may result. Similar but not identical sounds in L1 and L2, in contrast, tend to be perceived as belonging in the same phoneme category and their native-like pronunciations are thus more difficult to acquire.

While prediction of learning difficulties (or easiness in this case) based on L1-L2 similarities has been shown to be problematic, L1 influence based on L1-L2 differences has been found to be far more reliable (see Eckman, 2004, for a review). This is especially true with L1 transfer based on syllable structure differences between L1 and L2 (e.g., Benson, 1988; Broselow, 1984, 1983; Hansen, 2001; Major, 1987; Osburne, 1996; Sato, 1984; Tarone, 1987; Weinberger, 1987). Broselow (1983), for example, examined speakers of Egyptian Arabic and speakers of Iraqi Arabic for their pronunciation errors in speaking L2 English. Both groups demonstrate a tendency to simplify English onset clusters and a preference for the same simplification strategy of *epenthesis*, i.e., the insertion of a vowel into a consonant cluster, in their error patterns. A critical difference remained nevertheless: while the Egyptian Arabic speakers tended to insert a vowel *between* the consonants, the Iraqi Arabic speakers tended to insert a vowel *before* the consonant cluster. Such differences were found to be attributable to the epenthesis rules in their individual L1s.

As fundamental as difference-based contrastive analysis has been in predicting L2 learning difficulties, its limitations have also been well recognized. There were certain patterns of speech errors that had bothered L2 researchers simply because they could not be solely attributed to either the L1 or the L2. Anderson (1987), for example, found that Chinese speakers used both epenthesis and deletion in simplifying complex coda clusters in English. However, “Since epenthesis does not occur at all and since deletion does not occur widely in either dialect of Chinese [Mandarin and Amoy Chinese], contrastive analysis cannot predict which simplification process would be the dominant one for the Chinese group” (p. 288). Phonological universals, therefore, must exist to explain these

phenomena (see Eckman, 2004, for a review). Various models of phonological universals have been proposed, including models based on *typological markedness* and those based on *universal constraints*.

Markedness indicates the relative degrees of (un)naturalness of a certain linguistic structure. Structures that are *unmarked* are more basic, thus more natural and better preferred, both within and across languages. Open syllables (e.g., /ti/, *tea*), for example, are typologically unmarked relative to closed syllables (e.g., /tiz/, *tease*) because, while every language has open syllables, not all languages have closed syllables (e.g. Hawaiian). The relative markedness of closed syllables to open syllables has been tested for its ability to predict L2 errors, but with limited effect (e.g., Benson, 1988). L1 transfer appears to endure as a strong predictor of L2 learning difficulties (see Eckman, 2004; Hansen, 2001, for reviews). In fact, new models were usually brought up to complement, rather than to replace, the L1 transfer account. For example, the Markedness Differential Hypothesis (MDH) as proposed in Eckman (1977), while attempting to explain certain L2 error patterns in terms of markedness, includes nevertheless L1-L2 differences in two of his three criteria for predicting learning difficulties. Similarly, the constraint-based *Optimality Theory* (OT), which we now turn to, also has to take L1 transfer into account (e.g., Broselow, Chen, & Wang, 1998; Lombardi, 2003).

OT is a theory of Universal Grammar (UG). It differs from traditional UG in that while traditional grammar stipulates non-violable rules, OT is based upon groups of *violable universal constraints*. According to OT, all languages share the same set of universal constraints, which sometimes contradict each other, and they differ only in

terms of rankings of the constraints (Prince & Smolensky, 2004). For a given language, the higher a certain constraint is in the ranking, the higher its priority is in being satisfied. Most importantly, the lower-ranked constraints can be violated in order to meet the high-ranking constraints. Constraints being violable, OT appears especially capable of capturing developmental changes (e.g., Levelt, Schiller, & Levelt, 1999) and L2 error patterns (e.g., Broselow et al., 1998; Lombardi, 2003). More specifically, given a target sound or series of sounds to be pronounced (the input), the actual pronunciation (the output) is determined by the ranking of the constraints; and the ranking of the constraints can be decided by developmental stages or, in L2 learning's case, by L1 rankings. In other words, changes can result from the re-ranking of the same set of universal constraints; and a particular ranking may signify a particular developmental stage or a particular L1.

Lombardi (2003) illustrates how the different output sounds by speakers of different L1s can be interpreted in terms of OT. Given the same L2 English inter-dental fricatives (e.g., the *th* in *thank*) that are absent in all the L1s in question, it has been found that the speakers of some languages, such as Thai, Russian, and Hungarian, tend to replace it with the stop /t/, while speakers of other languages, like Japanese, German, and Egyptian Arabic, tend to replace it with the fricative /s/. Ioup and Weinberger (1987) interpreted this discrepancy solely in terms of L1 transfer, as the substitution patterns are consistent for any given language. However, as all these languages do not have the target sound in their L1s, it would be problematic to imagine how transfer can be possible. By means of the re-ranking of universal constraints, Lombardi was able to demonstrate that the different outputs result from the different constraint rankings of these speakers'

respective L1s. Interestingly, although Lombardi argued that while one group demonstrates a “more explicit transfer effect” (p. 246) but the other, a “universal ranking,” the comparisons are based on the rankings of respective L1s. In other words, L1 is still implied as the source of differences. L1 transfer appears to be indispensable.

Although the impact of L1 can be felt almost everywhere in L2 phonology studies, the degree of L1 transfer effect depends largely on the stages of L2 acquisition (e.g., Major, 1987). Major’s well-supported Ontogeny Model suggests that L1 transfer is at its strongest in the early stages of L2 acquisition. Its impact decreases with the increase in learners’ L2 proficiency, and its prominent role is gradually replaced by the impact of *L1 developmental process*, which consists of the natural stages of acquisition if the target language is the learner’s mother tongue. In other words, upon reaching a more advanced level of L2 acquisition, the learner’s L2 development begins to follow the same developmental stages as native speakers of the L2 (for the learner in question) do. The natural, native-like developmental process of the target language also decreases as the learner approaches L2-like proficiency. Although three stages are suggested, not every L2 learner can reach the final or even the second stage. This is especially true when the L2 is being learned as a foreign language, where the L2 is learned as a subject, not very different from math or history, rather than as a tool of communication. The second stage actually poses far fewer problems than the first, as the L1 developmental process is a natural process regulated more by the L2 than the L1. Educationally, therefore, L1 transfer appears to be the major obstacle to overcome especially in the initial stages of L2 acquisition.

Core Syllable as a Linguistic Universal

Core syllables (CV) have been described as the “optimal syllable” (Jakobson & Halle, 1956) and have been well accepted as one of the language universals because all languages have core syllables but not necessarily other types of syllables. In *second language acquisition* (SLA) research, Tarone (1976) was among the first to propose the universal preference for core syllables as a factor independent of L1 transfer on SLA. Examining production data from 6 subjects speaking three different L1s—Cantonese, Korean, and Portuguese—Tarone (1980) argued that, while some of the errors produced can not be accounted for on the basis of L1 transfer, the subjects’ simplification of codas appears to suggest preference for open syllables¹⁰. Ensuing studies investigating the influence of core syllable preference, however, have shown the effect to be far more limited compared to L1 influence (e.g., Benson, 1988; Hodne, 1985; Sato, 1984). Sato (1984), for example, examined two Vietnamese ESL learners at three time points over a ten-month period. A preference for closed, instead of open, syllables was observed. Similarly, using error analysis on the L2 data in contrast to the L1 data, Benson (1988) also found the effect of open syllable preference to be minor, compared to L1 transfer. The impact of core syllable *as a universal preference* in L2 acquisition therefore appears to be limited.

Core syllables are perhaps more important as the first acquired and the most frequently heard syllable type in language acquisition than as a universal syllable type for

¹⁰ Core syllables (CV) appear to have been used interchangeably with open syllable (e.g., Tarone, 1976; 1980; Benson, 1988), although open syllables in general definition can include but are not limited to core syllables (e.g., CCV). Eckman (2004), for example, refers to the core syllable preference in the following way “All languages appear to have syllables consisting of a single onset consonant followed by a vowel (open syllables)” (p. 527).

reasons detailed below. Levelt et al. (2000) mapped out the acquisition sequence of different syllable types of native Dutch-speaking children and found CV to be the first acquired syllable type. The study was based on the developmental data of 12 Dutch-speaking children of age 1 to 2 at the beginning of the study; spontaneous speech was recorded every next week for a period of 6 to 13 months. A Guttman scale was applied to the data collected at different time points. As the scale for the initial recordings showed, CV is the first syllable type acquired, compared to other syllable types that appeared in the acquisition order of CVC, V, VC, CVCC, VCC, CCV, CCVC, and CCVCC (In English, such structures would be exemplified by, respectively, *lead*, /lid/; the letter *e*, /i/; *eat*, /it/; *least*, /list/; *east*, /ist/; *flea*, /fli/; and *fleece*, /flis/). In addition to developmental order, the prominence of CV also emerged from the syllable type frequencies. In either the CELEX Database or the Corpus Van de Weijer cited in Levelt et al. (2000), CV remains the most frequently occurring syllable type in Dutch. Speech syllables of the CV structure were heard at least one third of the time (36% - 42%); and this was true with both child-directed and adult-directed speech data (p.259-261). Compared to CV, the syllable type VC ranked only 4th in the order of acquisition and its frequency is not even half of that of CV (11% - 15%). These are the potential reasons why Geudens and Sandra (2003) should find the Dutch children preferred a body unit (CV) instead of a rime unit (VC), although Dutch, like English, has more rime neighbors than body neighbors¹¹ (Ziegler & Goswami, 2005).

¹¹ Neighbors refer to other words that share the same phonological unit in question (e.g., rime or body). For the word /lid/, *lead*, for example, /lis/, *lease*, would be a body neighbor (sharing the body /li/); and /hid/, *heed*, would be a rime neighbor (sharing the rime /id/).

Syllable as a Phonological Processing Unit in Chinese

Syllables in Chinese have been suggested to play a unique role in speech production, even independent of tones (e.g., Chen, 2000; Chen et al., 2002; Chen et al., 2003; Chen et al., 2004, August). This special status of syllables has been supported by various studies with various tasks including speech error analyses (e.g., Chen, 2000), masked priming (e.g., Chen et al., 2003, where word naming was found faster if the target word was preceded by another word, which acted as a masked prime, that shared with it the same syllable), implicit priming (e.g., Chen et al., 2002, where subjects were asked to remember cue-target word pairs and to recall later the target when prompted with the cue; the cue thus served here as an implicit prime), and simple recurrent network modeling (e.g., Chen et al., 2004, August). The last of these also includes a direct comparison between Chinese and English and demonstrates the relatively greater salience of the syllable in Chinese than in English. Part of the plausible reasons appear to be the limited syllable repertoire of Chinese as discussed in the previous chapter (see also Chen et al., 2004, August, for a similar interpretation), which necessarily increases the frequencies of most, if not all, possible Chinese syllables. In what follows, three studies with different experimental designs will be reviewed for the particular status of syllables in morphosyllabic languages like Chinese.

Based on a corpus of 960,000 words collected from 120 recorded radio programs, Jenn-Yeu Chen (2000) conducted a computation of pure syllable movement errors. A pure syllable movement error was determined “(1) if the target and the source character [syllable] do not share the tone or any syllable constituents and (2) if the source character [syllable] moves without carrying its tone with it” (p. 16). As it turns out, the error rate

was low: only 10 such errors were obtained. However, the probability of syllable slipping (i.e., slip of tongue with syllable as a holistic unit in the error) far exceeds that of the subsyllabic elements simultaneously and independently slipping. Chen thus argued that “syllables are represented and processed as planning units in Mandarin Chinese” (p. 19).

Using an implicit priming paradigm, Jenn-Yeu Chen and her colleagues (2002) replicated and expanded the findings of her previous study. In the implicit priming task, sets of word pairs were given to undergraduate subjects in a Taiwan’s university for memorization and later recall. Each set consisted of 6 to 8 pairs (a cue word that served as the implicit prime plus a target word) of Chinese words, which were further manipulated in terms of two conditions: homogeneous condition and heterogeneous condition. Word pairs in the homogenous condition contained words that shared a certain examined unit such as syllable, tone, or onset; word pairs in the heterogeneous condition contained words that shared neither tones nor syllable constituents. A homogeneous pair that shared the same syllable can be exemplified by the two words /ma1/, *mother*, and /ma3/, *horse*, which share the same syllable /ma/ but differ in tones (first tone for the former and third for the latter; see note 8 for details on tones). A heterogeneous pair, on the other hand, can be exemplified by the word pair /ma1/, *mother*, and /bi3/, *pen*, which share neither syllabic elements nor tone. The subjects were asked to supply the target word when prompted with a cue word. As the results showed, while tone alone (i.e., the target words shared the same tone but not the same syllable) produced no priming effect, syllable alone (with different tones) produced a significant implicit priming effect. However, as tone is attached to the rime but did not contribute to the

variance, it is possible that it was onset, rather than the syllable as a whole, that contributed to the effect. An additional experiment was conducted to examine the rival interpretation, and it was found that onset alone did not play a part, either. Thus the priming effect can be attributed to the syllable in its entirety. Syllables in Chinese thus “act as a ... planning unit” (p.762).

While previous studies have attested to the status of syllables in Chinese as a phonological processing unit, it is Train-Min Chen et al. (2004) who have made a direct comparison of Chinese with English in terms of the syllable salience within a connectionist model, i.e., the simple recurrent network. The network, in the form of a computer program, is useful as a model “of how people implicitly learn the structure of sequences” (p. 2). Chen et al. (2004) applied the network to model the learning of sound sequences in Chinese and English. Ten short English-Mandarin Chinese bilingual children’s stories were entered to train the network, which was then made to predict a sound in a sequence. It was hypothesized that if a unit was salient in a language, the sounds that began the unit should be more difficult to predict than sounds situated elsewhere within the unit. This is because, given a unit, the foregoing elements within that unit should facilitate the prediction of the incoming elements, if the unit occurs as a whole frequently enough. Given words as the unit and given the word /le/, *late*, for example, the preceding occurrence of /le/ should narrow the possible candidates for the following sound, thus contributing to the higher prediction rate of the final sound /t/. Sounds that begin the word, i.e., /l/ in this example, on the other hand, would be more difficult to predict, as the sound preceding it belongs to another unit/word, say the /u/ of the word /tu/, *too*. More specifically, the word *too* can be followed by sounds of a much

greater variety, as the initial sounds of the second word in the following phrases: *too good*, *too happy*, *too naive*, etc., which all compete with the /l/ in *too late*. As the results showed, while sounds that begin a syllable (syllable initial items) in Chinese were more difficult to predict than non-initial, within-syllable sounds (within syllable items), they were no more difficult than sounds that begins a word (word initial items). In contrast, word initial items were more difficult to predict than both syllable initial and within syllable items. Thus, in Chinese, “the predictability of a sound was almost entirely determined by whether or not it is at a syllable boundary,” since “word-initial sounds themselves are also syllable-initial, and the syllable unit alone could explain the pattern without postulating a word level” (p. 3). Syllables as processing units were further confirmed.

Sub-syllabic Unit Studies

Introduction

In addition to the hypothesized L1 transfer of phonological processing units, another potential contribution of the present study is the potential insights it provides to the preferred subsyllabic division debate. As mentioned briefly in the previous chapter, while studies in the fields of theoretical linguistics and reading acquisition have pointed to onsets and rimes as the natural subsyllabic division units, recent studies with both Korean and Dutch speakers have suggested the onset-rime division as language-specific. In these studies, contrary to the usual predictions in earlier literature, the immediately prevocalic consonant (e.g., the first sound /d/ in /did/, *deed*) is found to be more difficult, rather than easier, to separate from the nucleus vowel than its postvocalic counterpart (e.g., the second /d/ of /did/, *deed*), given test items of the syllable types CV, VC, and

CVC. Two rival hypotheses compete to interpret the results. Like the onset-rime division, the results may indicate a similarly hierarchical division into bodies and codas (CV+C, where the first constituent, CV, is considered as a body unit). Alternatively, the results may pertain to a linguistic development that favors a core syllable interpretation:

[T]he stronger cohesion between the phonemes within a CV may be related to two observations in non-experimental settings: the preferred phonological structure of a syllable in languages across the world, and the preferred syllable type in child language. First, linguists have claimed that the CV is the universal core syllable ... There are no languages in which the CV syllable is not found. There are also no languages in which the only syllable types are V and VC... Second, in first language acquisition the CV is the first syllable type to be acquired... It makes sense that a sequencing of phonemes which is preferred across languages and considered to be natural in infants' spoken language, is difficult to take apart in a task at the level of explicit phonological awareness. (Geudens & Sandra, 2003, p. 174)

At later stages of linguistic development, the young speakers are argued to restructure their syllable structure with increased vocabulary to accord with the statistical properties of their native languages. Chinese beginning ESL students appear to be an ideal alternative population for investigating the rival interpretations, for core syllables as a salient phonological processing unit in Chinese is well established on the one hand and, on the other, L1 transfer effect is at its strongest in the initial stages of L2 acquisition. Chinese beginning ESL students thus function as equivalents of the, say, young Dutch-speaking children who are just in their initial stages of native language acquisition and thus have just acquired the universal core syllable. Before advancing my hypotheses and predictions, however, I will first review literature involved in the discussion of preferred subsyllabic division units in the following sections.

Onset-rime Studies

In addition to the segmentation tasks discussed in the previous chapter, the strongest support for the preferred status of onsets and rimes comes from two strands of research: short-term memory error studies (Treiman, 1995; Treiman & Danis, 1988) and studies on the statistical properties of the languages in question, including distributional studies (e.g., Blevins, 1996; Kessler & Treiman, 1997), and neighborhood density studies (e.g., De Cara & Goswami, 2002; Ziegler & Goswami, 2005). The first strand of studies, to begin with, examined the patterns of recall errors and found that the most frequent type of errors usually demonstrated an onset-rime division. Treiman (1995), for example, examined subjects across a wide range of age groups, including kindergarteners, third graders, sixth graders, and undergraduates, for the types of recall errors they tended to make. For each list of three nonsense syllables, the subjects first heard each syllable with a one-second interval and, after hearing the third, were asked to repeat all the three syllables in the order they had been given. Results were analyzed and categorized in terms of the errors and their types. It was found that, across all age groups, recall errors that combined the onset of one to-be-remembered nonsense syllable with the rime of another to-be-remembered syllable were most frequent. The author thus argued for the psychological reality of onsets and rimes. This study replicated the findings of Treiman and Danis (1988), which studied only adults, and thus established the same psycholinguistic reality also in children.

The other strand of onset-rime research focuses on the statistical properties of phonological syllable structures. De Cara and Goswami (2002), for example, examined English monosyllabic words in the Celex database for their *phonological neighborhood*

densities, i.e., the frequencies of occurrence of words with the same unit of interest.

Three possible types of psycholinguistic neighbors are distinguished for their densities, including rime neighbors, which differ from the target words only in onset; consonant neighbors, which differ from the target words only in the nucleus; and lead neighbors (i.e., body neighbors), which differ from the target words only in coda. For the target word /hæt/, *hat*, for example, its rime neighbors would include /væt/, *vat*, /ræt/, *rat*, /tæt/, *tat*, etc.; its consonant neighbors would include /hit/, *heat*, /het/, *hate*, etc.; and its body neighbors would be /hæv/, *have*, /hæp/, *hap*, etc. Based on an inventory of about four thousand monosyllabic spoken words adapted from the CELEX database (Baayen, Piepenbrock, & Guilikers, 1995, cited in De Cara & Goswami, 2002, p.417), the numbers of phonological rime neighbors, consonant neighbors, and lead neighbors as well as their corresponding proportions (against the total number of neighbors that include all three types) were calculated. As the analysis showed, the proportion of phonological rime neighbors was much larger than either of the other two types of neighbors both by type (54.2%, compared to the 17% of consonant neighbors and the 28.9% of lead neighbors) and by token (56%, compared to the 24.2% of consonant neighbors and the 19.8% of lead neighbors).

Similar analyses were also conducted of German, Dutch (based on the same CELEX database), and French (based on the BRULEX database of Content, Mousty, & Raadeau, 1990, cited in Ziegler & Goswami, 2005, p.8). Like those of De Cara and Goswami (2002), the results showed a similarly dominant status for rimes, though the relative neighborhood densities for the other two neighbor types, i.e., consonant neighbors and lead neighbors, vary among the three languages. To test for the

psychological reality of rimes as the predominant subsyllabic units, De Cara and Goswami (2003) conducted a further study on English pre-reading children. Given a rhyme oddity task, where subjects were asked to pick the word that sounded “odd” relative to other two sounds from triples of words, pre-readers with high vocabulary, but not those with low vocabulary, were found to demonstrate the neighborhood density effects. The authors thus concluded that neighborhood density and vocabulary contribute, directly and indirectly, to the development of phonological awareness.

Such a statistically prominent status of rime is present also in the spelling-sound correspondence in languages with the so-called deep orthographies (see below), such as English (but not those with shallow orthographies), which has led to, among other things, the psycholinguistic grain size theory (for a review, see Ziegler & Goswami, 2005). Psycholinguistic grain-size models (Goswami, Ziegler, Dalton, & Schneider, 2001, 2003; Ziegler & Goswami, 2005; Ziegler, Perry, Jacobs, & Braun, 2001) can be viewed as an extension of the forerunning Orthographic Depth Hypothesis (ODH; Katz & Feldman, 1983; Katz & Frost, 1992). In terms of consistency differences in letter-sound correspondence among alphabetic orthographies, the ODH postulates that readers would develop different processing strategies according to the orthographic depth of the language they are learning (Katz & Feldman, 1983). In shallow orthographies, such as German, Spanish, and Italian, for example, letters or letter clusters (graphemes) constantly represent the same phonemes with rare exceptions. In deep orthographies such as English, in contrast, grapheme-phoneme correspondence (GPC) is so varied as to be far less reliable. In English, for example, the sound /e/ can be spelled as different as *a_e*, as in *cake*; *ai*, as in *mail*; *ay*, as in *day*; or *ei*, as in *eight*. Conversely, the same letter can

represent sounds as varied as /æ/, as in *am*; /ɛ/, as in *many*; or /a/, as in *are*. The unreliability at the GPC level in English, however, is usually reduced at higher levels with larger units such as orthographic rimes. Thus the orthographic rime *-ar* is consistently pronounced as /ar/, as in *arm*, *large*, *car*, *mar*, etc; and *-are* is consistently read as /ɛ/, as in *care*, *mare*, *hare*, *pare*, etc.

Such small vs. large grain sizes of psycholinguistic units have been argued for both their statistical and psychological realities. Treiman, Mullennix, Bijeljac-Babic, and Richmond-Welty (1995) examined the letter-sound correspondence consistency of English vowels that differ in the presence/absence of and/or combination order with an accompanying consonant. The results indicated that with CVC words, the VC₂ clusters have far more predictable, hence reliable, pronunciations than the C₁V cluster or the vowel (V) alone. These statistical properties indicated a prominent role of the orthographic rime unit (i.e., VC₂) and were further examined for their psycholinguistic impact. Both adults and children were examined in both mega-studies and small-scale factorial studies for psychological realization of these statistical characteristics of English. The results supported the psychological reality of these statistical characteristics of English, and pointed to the possible coexistence of small-unit (GPC) and large-unit orthographic decoding.

Based on the varying degrees of orthography-phonology correspondence consistency of different writing systems, the psycholinguistic grain-size hypothesis proper further postulates that the grain sizes of psycholinguistic units are a function of such consistency of the language in question (Goswami et al., 2001, 2003; Ziegler & Goswami, in press; Ziegler et al., 2001). Small grain size is naturally adopted as the

psycholinguistic unit in a shallow language, in which the GPC is highly reliable and small-unit decoding (based on GPC) thus becomes the dominant method of orthographic processing in reading new words. In a deep orthography, on the other hand, the reliability of letter-sound correspondence varies with the level of orthographic representation (usually greater at the higher levels and less at the lower levels) and readers of such an orthography “have no choice but to rely on a variety of psycholinguistic grain sizes, including the whole word phonology and orthographic units corresponding to rhymes, in addition to grapheme-phoneme correspondences” (Goswami et al., 2003, p.237). Similar to the psycholinguistic grain-size hypothesis, the flexible-unit-size hypothesis (Brown & Ellis, 1994) also points to the possible parallel development of both grain sizes of phonological representation, with a special emphasis on the development being dependent on the “informational demands of the precise task they are required to perform” (Browns & Deavers, 1999, p.210).

Thus in addition to the statistical basis for the onset-rime division in the spoken languages, an analogous statistical basis can also be found for the preferred status of rime in the orthography of a language with poor GPC, though not those with consistent GPC. Statistical properties, however, have their application limitations, especially in regard to developmental issues. That is, statistical properties imply a relatively more mature lexicon, orthographic or phonological, which is usually possible only with older children or adults. For young children with insufficient vocabulary, as in the low vocabulary group of De Cara (2003), it is plausible that the younger they are, the less sensitive they are to such properties, as their limited lexicon might limit the sample sizes (i.e., the number of words they know) from which such statistical properties can be reliably

inferred. Onset-rime as the subsyllabic division unit, therefore, may not be inborn. Indeed, based on the linguistic properties of different languages (e.g., Ziegler & Goswami, 2005) and certain linguistic universals such as the unmarked open syllable (e.g., Benson, 1988), bodies and codas, instead of onsets and rimes, have been argued to be the preferred subsyllabic units in other languages such as Korean (e.g., Derwing et al., 1993; Yoon et al., 2002; Yoon & Derwing, 2001).

Body-coda Studies

Yoon et al. (2002), for example, demonstrated in four experiments that bodies and codas appeared to be the preferred subsyllabic division units for native speakers of Korean. Exp. 1 of Yoon et al. (2002) is basically a replication of its first author's 1997 dissertation study. However, in order to rule out the competing factor of task demands, Yoon and her colleagues opted for the analogy task used in Goswami (1993), where an onset-rime division preference was found for native English speakers, instead of the Grapheme Substitution task the first author used for her dissertation, where a body-coda division preference was obtained. Three phases were involved for each subject (28 kindergarteners) in the task. In the pre-reading phase, test nonwords were given to the subject to test the novelty of the nonword items. Nonwords were replaced with new ones if the subject was able to read them. In the word learning phase, the subject was given a written clue word along with its pronunciation; the subject was then encouraged to try to read the word. In the final, analogy phase, the subject was asked to read the test items (the target words) with the help of the clue words they had already learned. The clue words were so manipulated that they shared (1) the rime, i.e., VC; the (2) body, i.e., CV; (3) the nucleus, i.e., V; or (4) all the syllable consonants, i.e., C_C, with the target words.

The results showed, among other things, that children scored a significantly higher accuracy rate when the clue nonwords shared syllable body with the target word than in all other three conditions. The results thus bore out Korean students' body-coda preference when task demands were held constant. One thing worthy of note is that, though the nature of the task used by Yoon and her colleagues was basically the same as that used in Goswami (1993), the former nevertheless used nonwords in contrast to the real words used in the latter. In both studies, moreover, vowel length was not controlled. For example, though the actual test items were not provided, all the nonword examples given in Exp. 1 of Yoon et al. (2002) were of the syllable pattern CGVC (e.g., /sjap/). As a glide plus a vowel constitute a rising diphthong, in contrast to a falling diphthong where the vowel precedes the glide as in the word /aj/, *eye*, the vowel length effect (Treiman et al., 2002) might confound which of the two division types are preferred.

One possible explanation for such body-coda preference could reside in the tendency in Hangul, the square-block-shaped Korean alphabet, to group onset-vowel together more often than vowel-coda grouping. To investigate whether such orthographic characteristics could contribute to the different subsyllabic segmentation preferences as observed in, respectively, Exp. 1 and Goswami (1993), two additional experiments were conducted to examine the orthographic effect. A Grapheme Substitution task was employed in both Exp. 2 and Exp. 3. Similar to the analogy task, the subjects were first given an explanation of the basic structure of the writing system in question (English alphabet in Exp. 2 and Korean Hangul in Exp. 3). Next, during the learning phase, subjects were taught to read a clue word or nonword. In the final, testing phase, subjects were first presented with a clue word written on a card along with its pronunciation.

With the card in sight, the target word was then read to the subjects, who were required to say which part of the clue word had to be changed in order to produce the target word. In Exp. 2, English words and nonwords were tested on Korean-speaking kindergarteners in the Grapheme substitution task, while in Exp. 3, artificial Korean syllables were tested on English-speaking undergraduates. The results suggested that, despite orthographic differences, the division preferences appear to be language-specific. While Korean kindergarteners demonstrated a body-coda preference even working with the English alphabet, American college students demonstrated an onset-rime preference even working with the Korean orthography. The sources of the preference differences therefore are unlikely to be found in the writing system.

With orthography failing to demonstrate an impact, linguistic factors become the most likely candidates, which is expressed in the Linguistic Hypothesis Yoon and her colleagues proposed: “subsyllabic pattern preferences observed in reading arise from language differences, rather than script differences or other factors” (p.156). To examine the hypothesis, Exp. 4 was conducted with a similarity judgment task on both Korean and American undergraduates. Though items of both CVC and CVCVC syllable patterns were examined, Yoon et al. (2002) reported only results concerning the former, and so only results of the monosyllabic CVC items are summarized here. In this experiment, both word and nonword items were used, and non/word pairs of the syllable pattern CVC were created that varied in terms of number of shared phonemes. The number ranged from 0 to 3; and for pairs of non/words that shared two phonemes, three subtypes were distinguished: those sharing CV, VC, and C_C. Such pairs of non/words were presented in their spoken form to the subjects for judgment of similarity on a scale of (and

ascending from) 0 to 6. Again, language-specific preferences were found: while Korean college students tended to judge pairs of non/words as more similar if they shared the CV units than if they shared VC or C_C units, English undergraduates tended to judge non/word pairs as similar if they shared VC units than otherwise. As the impact of literacy could not be ruled out, moreover, the authors suggested that further studies be conducted with pre-readers to further substantiate the Linguistic Hypothesis (p.157). However, the Linguistic Hypothesis was defined at a very general level without specifying which aspects of linguistic structure contributed to the results. One of the few more specific speculations arising from the Hypothesis was the contrast in syllabic complexity between Korean and English. As we shall see in the following, however, syllabic complexity in and of itself may not be the sole underlying factor, as speakers of Dutch, a language similar to English in syllable complexity, are also found to favor a CV-C segmentation (e.g., Geudens & Sandra, 2003).

Geudens and Sandra (2003) conducted four experiments with Dutch-speaking children with a view to examining whether onset-rime preference was universal. Unlike previous studies, Geudens and Sandra tried to control most potential competing factors that included phoneme components of the test items (so that the C of the CV is the same as the C of the VC, and so is the V). In addition, a distinction was also made between implicit and explicit tasks on the basis of whether the skills involved were conscious or unconscious to the subjects. Finally, only long vowels (such as the vowel /e/ in *lake*) and diphthongs were used in the test items, thus keeping vowel length a constant. In Exp. 1, kindergarteners were tested on pairs of CV vs. VC items. Children first heard a two-phoneme syllable and were then asked to segment it. Both division types and sonority

were significant. Division type was significant because VC items were found to be easier to segment than CV items. The sonority effect was a bit more complex. Items containing stops and fricatives, both being obstruents, did not differ from each other; neither did items containing liquids and nasals, both being sonorants. Comparisons across the two overarching categories, i.e., obstruents and sonorants, however, were all significant. This is consistent with Treiman et al.'s (1995) findings, where differences in vowel-consonant cohesion were also observed on an obstruent vs. sonorant basis, but not on the further subcategories of, for example, stops vs. fricatives. More importantly, VCs were easier to segment than CVs regardless of whether the consonants were sonorants or obstruents, though the division preference effect was larger for the latter than the former.

As the segmentation task was too hard for the kindergarteners in Exp. 1, Geudens and Sandra further recruited older (about one year older than those in Exp.1) but nonetheless pre-reading children for the same test in Exp. 2. An additional difference distinguished Exp. 2 from Exp. 1. In Exp. 1 the children received only one of two different lists of items, and the items were so manipulated that the presence of a particular CV item would exclude the presence of a VC with identical phoneme components. Pre-readers in Exp. 2, in contrast, received items including pairs of exact reversals, i.e., pairs of CV and VC that share the same phoneme components. This ensured a full within-subject design. Again, as the results showed, VC syllables were easier to segment than CV ones; and obstruents were easier to segment than sonorants from the nuclei. Two additional interesting results were also obtained. The subjects tended to misperceive or even failed to repeat sonorant sounds (i.e., liquids and nasals). When subjects responded erroneously with only one single phoneme (instead of the two phonemes that constituted

the test items), such single phonemes tended to be the initial sound rather than the final sound. That is, when an incomplete one-phoneme answer was given, it was more likely to be the first sound regardless of whether the item presented was of CV or VC type. Uhry and Ehri (1999) also ascribed the easier segmentation of VC than that of CV found in their kindergartener subjects to such a positional effect.

To determine whether the results obtained in Exp. 1 and 2 were task-specific and to confirm the position effect, Exp. 3 adopted a substitution task in which the subjects were supplied with a single phoneme to substitute one of the two phonemes in the CV or VC items. The substitution task was used because one potential problem with segmentation tasks was the inequality of perceptual salience implied in the different orders of CV vs. VC. It was suspected that while there was no interruption in articulating a CV syllable, the interruption was obvious in pronouncing the other. With the substitution task given to some of the same kindergarteners in Exp. 1, it was found that division unit was no longer significant. On the other hand, the position effect was again significant. It thus appeared that the easier segmentation of VC than of CV could have been confounded by phonetic factors. However, even with the phonetic factors controlled, the traditionally assumed onset-rime preference still failed to find support in results of all three experiments. As test results with two-phoneme items (i.e., CV vs. VC) were different from those with three-phoneme items (i.e., CVC) in Uhry and Ehri (1999), a further experiment was conducted to see whether syllable complexity would contribute to different results. As the number of phonemes was expected to be related to item difficulties (i.e., CVC items should be more difficult than either CV or VC items), the subsyllabic division tendency was measured by analyzing error types. That is,

incomplete segmentation of CVC items was classified according to whether it was CV or VC that was left unsegmented by the children. Again, the results showed a comparative easiness in segmenting bodies from codas than segmenting onsets from rimes, as more CVs were left unsegmented than VCs.

Despite all the controls adopted, one potential problem remains with Geudens and Sandra's (2003) study: their selection of only long vowels or diphthongs as the syllable nuclei. The problem arises because of vowel quality (i.e., length) effect as demonstrated in Treiman et al. (2002), in which subjects responded differently to syllables with short vs. long vowels. It appears that an intervocalic consonant in a disyllabic word (e.g., the second C in a CVCVC word, such as the consonant /b/ in *habit* or *rabbit*) is more likely to be syllabified with the first syllable if the vowel of the first syllable is a short vowel than if it is a long vowel. The body-coda preference therefore may have been confounded by the vowel length effect, which along with the other controls adopted in Geudens and Sandra (2002) will be reviewed in the next section.

Methodological Issues

Introduction

The experimental controls adopted in this study can be classified as of two types: general controls and linguistic (i.e., item) controls. General controls refer to the inclusion of more traditional confounding variables that are related to the characteristics of the participants involved such as oral vocabulary and nonverbal intelligence. They are incorporated so that their impact on individual performance can be ruled out as rival interpretations. As they are traditional, moreover, they are more likely to have established, standardized measures, such as Peabody Picture Vocabulary Test, Third

Edition (PPVT-III henceforth), which measures English oral vocabulary, and Matrix Analogy Test (MAT), which measures nonverbal intelligence. In this section, therefore, the focus will be placed on the linguistic controls. This is because, for one, the standardized measures usually have better established validity and reliability, and, for another, linguistic controls have been argued only more recently to be important. This section, therefore, will be devoted to the importance of such linguistic controls as argued in prior research.

Although the contribution of phonological awareness to beginning reading has been well established, phonological awareness as a homogeneous concept has not gone unchallenged (e.g., Stahl & Murray, 1994; Yopp, 1988). Yopp (1988), using principal factor analysis, for example, obtained two factors from measures of ten phonological awareness tasks. Yopp's findings, however, have been argued by Stahl and Murray (1994) to have been confounded by linguistic complexity. Based on a series of linguistics-based studies she and her colleagues had conducted for more than a decade, Treiman (2000) also pointed out that "phonological awareness is not a single homogeneous skill. Some linguistic structures are more difficult than others, and these difficulties can cause specific errors in spelling and reading" (p.90). Given the significant impact of linguistic features, the importance of linguistic control in test items has received more and more attention in phonological awareness related tasks, which include syllable recombination (e.g., Treiman, 1984, 1986, where subjects were required to recombine two syllables into one), phoneme segmentation (e.g., Geudens & Sandra, 2003), phoneme substitution (e.g., Geudens & Sandra, 2003), phoneme counting (Treiman et al., 1995), alliteration/rime matching (e.g., Carroll & Snowling, 2001),

resyllabification (Treiman et al., 2002), and spelling (e.g., Treiman et al., 1995). These controls can be roughly classified as of three types: sonority (e.g., Geudens & Sandra, 2003; Treiman, 1984; Treiman et al., 1995; Treiman et al., 2002); vowel length (e.g., Geudens & Sandra, 2003; Treiman et al., 2002), and global similarities (e.g., Carroll & Snowling, 2001). In the present study, these three types of linguistic controls are all implemented in the items of the experimental measures to allow for clean and clear interpretations of the results. Their importance is reviewed as follows.

Sonority

As discussed in Chapter I, sonority has been defined mainly by the degree of resonance in the vocal tract, i.e., the degree of vowel-likeness, and speech sounds so categorized by sonority can be of great avail in explaining phonological phenomena (Spencer, 1996). The importance of such categorization can be exemplified by sonority ranking in its ability to describe one important formation principle of a syllable, which Selkirk (1984) termed the “Sonority Sequencing Generalization” (SSG): “In any syllable, there is a segment constituting a sonority peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values” (p.116). Following this generalization, we are more likely to see an obstruent precede a sonorant or a glide in an onset cluster, as in *please* (where the more sonorant liquid /l/ is closer to the nucleus vowel than the less sonorant stop /p/), but vice versa in a coda cluster, as in *guilt* (where, again, the more sonorant liquid /l/ is closer to the nucleus vowel than the less sonorant stop /t/). For reading acquisition, particularly for phonological awareness, the implication of this sonority dimension is that the farther apart two neighboring phonemes are in sonority, the more distinct the two sounds become from each other. On the other hand,

the closer two phonemes are, the more difficult it will be to identify or separate the two sounds as distinct. As arguments over the preferred subsyllabic division units have been based on performance in speech sound segmentation, control for sonority in phonological awareness tasks appears necessary, as it would impact how cohesive two sounds are, hence how difficult it is to segment them.

In reading literature, indeed, sonority has been found to play a major role in how subsyllabic units are perceived (e.g., Geudens & Sandra, 2003; Treiman, 1984, 1992; Treiman et al., 1995; Treiman et al., 2002). Treiman (1984) asked college students to blend two VC₁C₂ nonwords into one, while manipulating the sonority of the first consonant of the coda consonant cluster (i.e., the C₁). It was found that the subjects recombined the nonwords differently according to whether the manipulated phoneme was a liquid (as the /r/ in /ar^z/ and /l/ in /l^d/), a nasal (as the /n/ in /~~Λ~~ns/ and the /m/ in /æmd/), or an obstruent (as the /p/ in /i^pt/ and the /k/ in /u^ks/). The subjects tended to blend liquid pairs by combining the VC₁ of the first syllable and the C₂ of the second (i.e., a VC/C blending). Thus the liquid pair /ar^z/ and /l^d/ were more likely recombined as /ard/ rather than /ald/, giving a response type proportion of, respectively, 49% and 21%. For an obstruent pair, the blending tendency is reversed, with a V/CC blending more likely (.58) than a VC/C blending (20%). The obstruent pair /i^pt/ and /u^ks/ was thus more often combined into /iks/ than /ips/. The different degrees of cohesion strength between the vowel nucleus and its neighboring coda consonant in terms of the consonant's sonority were also demonstrated by children's tendency to spell syllables containing a postvocalic liquid or nasal differently from those including obstruents (Treiman et al., 1995). Thus the four-phoneme syllable /vans/ with an immediately

postvocalic nasal /n/ is more likely spelled as a three-phoneme one like “vos” (66%) than its obstruent counterpart /dʌsp/ being spelled as a three-phoneme one such as “dup” (23%).

Perhaps the best example of the confounding of sonority is given in Geudens and Sandra (2003), where the conflicting results among different studies were argued to have resulted from the failure to control for sonority in the studies compared. Bus (1985, cited in Geudens & Sandra, 2003), to begin with, found that for Dutch kindergarteners VC syllables are easier to segment than CV ones, which is consistent with the body-coda account. An opposite result, however, was reported in Schreuder and Van Bon (1989), where it was CV that was found to be the easier of the two to break up for Dutch first graders. On further examination of the items, it was shown that Schreuder and Van Bon (1989) had test items containing mainly sonorants for the postvocalic consonants. Thus VC was harder to break up than CV due the former’s stronger cohesion between the nucleus vowel and the coda consonant. In Bus (1985), in contrast, three fourths of the items involved postvocalic obstruents. Geudens and Sandra (2003) thus suggested that the conflicting results among different studies could have originated in the failure to control for sonority in the related studies. As similar effects of sonority was also observed in their own study, Geudens and Sandra thus concluded that “The methodological implication is that in experiments on explicit phonological awareness, the researcher should control for the consonant’s sonority [and] use identical phoneme material in the critical condition” (p.173). The spirit of the additional control of “identical phoneme material” is nicely reflected in the control of global similarities, which we now turn to.

Global Similarities

Global similarity refers to the degrees of similarity in phonological properties between two phonological units (e.g., phoneme, rime, syllable, or word) and was operationally defined in Treiman and Breaux (1982) as the sum of the ratings of similarity between each phoneme of one unit and its corresponding phoneme in the other unit. Based on the ratings of similarities established in earlier research, Treiman and Breaux (1982) were able to calculate the overall similarity of each pair of items they examined. The ratings were given on a scale from 0 to 7, with 0 representing identical phonemes and 7, maximally dissimilar phonemic pairs. The overall similarity ratings of a pair of syllables are calculated by summing up the ratings of each phoneme pair. The smaller the ratings, the higher the overall similarities are between two units of the same number of phonemes. Take the two three-membered syllables, /bɪs/ and /bun/, for example. Their overall similarity is the sum of the similarity ratings of each of the three corresponding phoneme pairs: /b/ and /b/, /ɪ/ and /u/, and /s/ and /n/. The first pair was given a rating of 0, as they are the same phoneme, the second, 5.1, and the third, 5.0, and thus the overall similarity is 10.1 ($= 0+5.1+5.0$). Treiman and Breaux (1982) found that pre-reading children are more sensitive to *global similarities*, as operationally defined by rating calculations above, than to the *number of phonemes shared* in pairs of test items, whereas adults are more sensitive to the latter than the former. Methodologically, this implies that in tasks tapping *implicit* phonological awareness, i.e., tasks that require only a vague sensitivity to speech sound structure, young children may rely more on their sensitivity to global phonological similarities than on their actual phonological awareness, the latter of which is usually measured by the number of phonemes shared, instead of the

overall impression of similarity, between two phonological units in implicit tasks such as oddity task and matching task.

Results from related studies have supported the effect of global similarities and suggested the necessity of global similarity control in phonological awareness tasks, especially with young children (e.g., Byrne & Fielding-Barnsley, 1993; Cardoso-Martins, 1994; Carroll & Snowling, 2001). Carroll and Snowling (2001), for example, tested pre-literate children with rhyme and alliteration matching tasks that varied in terms of three distractor types: those unrelated either phonologically or semantically to the cue, those semantically related to the cue, and those matched for global similarity with the cue. Significant main effect of distractor types was obtained in the ANOVAs of both alliteration and rime matching tasks, and post hoc analyses showed that the effect was due mainly to the distractors that were matched for global similarities. Items with such “global distractors” were found to be more difficult than those with either unrelated or semantically related distractors. The authors thus concluded that “The present study ... confirmed the suitability of using phonological awareness tasks, controlled for global similarity, with preschool children” (p.338). In the present study, global similarities are controlled for whenever possible.

Vowel Length

Treiman et al. (2002) investigated orthographic and phonological influences on a wide range of age groups, including first graders (for Experiment 1 only), second graders, sixth graders, and adults. While orthographic impact was found only with older children and adults, thus suggesting the importance of familiarity with words’ spellings in order for orthographic impact to take place, the impact of phonological factors was significant

across all age groups (first graders not included). The phonological factors examined were two: sonority and vowel length. In Experiment 2, for example, where phonological impact was investigated, bi-syllabic spoken words with singleton intervocalic consonant and with stress on the first syllable, i.e., (C)(C)VCV(C)(C), were presented to the subjects for syllabification (into two syllables). The main responses categories of interest were classified according to whether the intervocalic consonant is syllabified with (1) the *first syllable only* (i.e., VC + V), (2) the *second syllable only* (i.e., V + CV), or (3) *both syllables* (i.e., VC + CV). The stressed vowel varied by *length*, i.e., a long or a short vowel, and the intervocalic consonant varied by sonority, i.e., sonorants or obstruents. The results showed that the main effects of both sonority and length were significant. The sonority effect was significant because intervocalic singleton sonorants were more likely to be syllabified as belonging with the first syllable only than intervocalic obstruents were. Length effect is also significant because, for both sixth graders and college students, the intervocalic singleton consonants in words with stressed long vowels were more likely to be assigned to second syllable only than to first syllable only or to both syllables. Thus the /m/ in the word *demon*, whose stressed (first) vowel is long, was more likely to be syllabified as *de* + *mon*. On the other hand, a stressed-short-vowel word was more likely to retain the intervocalic consonant in the first syllable, with or without its repetition in the second syllable. For example, *lemon* was more likely to be syllabified as either *lem* + *on* or *lem* + *mon* than as *le* + *mon*, which was true across all age groups. In addition to sonority, vowel length thus appears to have an additional impact on the cohesiveness of the postvocalic consonant with its preceding vowel. Applying the findings to the examination of subsyllabic division unit preference, it would

mean that vowel length would at least in part determine how easily the coda of a CVC syllable could be segmented, and its effect obviously needs to be controlled for as well.

Overview of the Present Study

Introduction

Two studies were originally planned for this dissertation: the study and one optional, backup study, as discussed in Chapter 1. The optional study as a backup plan was intended to rule out the potential confound of Zhuyin instruction. As the results turned out as expected, the impact of Zhuyin instruction, if any, seems ignorable and the optional study is thus no longer required. As it was part of the theorizing process for this dissertation, a discussion of the optional study is nevertheless retained as Appendix F.

The Study

Two experimental tasks were tailored to the research questions that have motivated the present study. Items created for the two tasks, with major confounding variables controlled for, are central to the statistical analyses of the results. Given the significant role these experimental tasks assume in the present study, the rationale and the constraints behind the item creation will be elaborated. Predictions based on all possible theories, instead of only the core-syllable awareness hypothesized here, will then be made as an attempt to make interpretation of the results as clear-cut and comprehensive as possible.

Use and Creation of Measures

Three standardized measures were adapted and two experimental measures were created specifically for the issues addressed in the present study. The two experimental measures were a phoneme deletion task and a phoneme isolation task, and the three

standardized measures were (a) Peabody Picture Vocabulary Test—Third Edition (PPVT-III), which measures oral vocabulary, (b) the Pattern Completion subtest of Matrix Analogies Test (MAT), which measures nonverbal intelligence, and (c) the subtest Word Attack of Woodcock-Johnson Tests of Achievement, Third Edition (WJ-III), which measures decoding skills. Among the three standardized tasks, PPVT-III and MAT subtests were employed to control for the potential confounding effects of the abilities they measure. Measures taken from Word Attack, on the other hand, functioned as the criterion variable to gauge the predictive power of the theoretically different segmentation abilities (i.e., core-syllable vs. non-core-syllable segmentation abilities). The considerations and rationale involved in the generation of the experimental items are given below, beginning with the deletion task.

In the phoneme deletion items, linguistic complexity were so controlled that the to-be-deleted (target) phoneme was always part of a consonant cluster and that all items contained exactly four phonemes and were all closed syllables. This resulted in only two syllable types: CCVC and CVCC. These items were then manipulated to reflect three within-subject factors, including two experimental factors, i.e., position and marginality, and one control factor, i.e., sonority. *Position* refers to whether the consonant to be deleted was situated before or after the vowel: an item was *prevocalic* if the target, to-be-deleted sound preceded the vowel, as in the case of (C)CVC or C(C)VC, where the parentheses indicate the target sound; an item was *postvocalic* if its target sound followed the vowel. *Marginality* refers to the target sound's position within the syllable: an item was *marginal* if the target sound either began or ended the syllable, as in the item types (C)CVC and CVC(C). An item was *non-marginal*, if vice versa, as in the item types

C(C)VC and CV(C)C. Thus an item of the type (C)CVC, for example, was a prevocalic, marginal item, whereas an item of the type CV(C)C was a postvocalic, non-marginal item. Finally, all the consonantal items were further classified, mainly in terms of sonority, as one of the following categories: liquid, L1 nasal, non-L1 nasal, and obstruent. This control was implemented to ensure that whatever subsyllabic division patterns were obtained, they were not confounded by sonority effect. It is important to note that the sonority control was limited to consonants immediately preceding or following the nuclei only. This is because both the Sonority Sequencing Generalization in general and *collocation constraints* in English allow neither liquids nor nasals to begin a consonant cluster in the onset position (Spencer, 1996). For example, while in English words can begin with obstruent-sonorant sequences such as /sm/ of the word /smir/, *smear*, or /fl/ of the word /flaj/, *fly*, the opposite sequences, i.e., */ms/- or */lf/- (where asterisks indicate unacceptable forms or usage), are not allowed to begin a syllable.

Sonority constraints of this sort seriously limited marginality as a full-fledged factor for analysis. For example, while an obstruent may exist, with some constraints, in both marginal and non-marginal positions in a consonant cluster, as in the two postvocalic items /kus(p)/ (marginal) and /ku(p)s/ (non-marginal), a nasal or a liquid cannot, as we have /ku(m)p/ (non-marginal) but not */kup(m)/ (marginal). In cases like the latter, where nasal and liquid consonants can be found only in non-marginal items, the target sounds in marginal items must be replaced with obstruents, thus losing contrasts in marginality. Confinement imposed by the sonority constraints on marginality was found on linguistic complexity as well. In the Design and Data Analyses section in Chapter III, marginality and linguistic complexity were thus similarly excluded from the analyses of

variance, as they were not fully cross-sectioned with other two factors, namely, sonority and position, in either phoneme isolation or phoneme deletion tasks.

In addition to sonority, the other two confounding variables reviewed in the previous section were also controlled here. Vowel length was controlled by the inclusion of only short vowels in all the items in both the phoneme deletion and the phoneme isolation tasks (see Appendix A for all deletion items and Appendix B for all isolation items). As codas have been shown to be more cohesive to short vowels than long vowels, any bias resulting from using only short vowels should favor an onset-rime segmentation. For example, the likelihood is higher for the word /tɛk/, *tech* (where the vowel is short), to be segmented as /t/ + /ɛk/ than for the word /tek/, *take* (where the vowel is long), to be segmented as /t/ + /ek/. If, as our hypothesis would predict, a non-onset-rime segmentation still turned out to be preferred, this would only make even more rigorous the core-syllable hypothesis argued here. A global similarity effect, on the other hand, is also controlled for by attempts to match, whenever possible, the linguistic features of the items involved in contrasting conditions. In extreme cases, this would mean contrasting items being the total reversal of one another. Given liquid items contrasting in position, for example, we have the nonword items /(f)rep/ and /pɛr(f)/ that are the total reversal of each other and that contrast only in position.

However, there were certain constraints preventing all items from being so manipulated. For example, with the exception of liquid items (e.g., /flaj/, *fly*), a clustered onset can only begin with the phoneme /s/ (as in /spɛr/, *spare*; /slaj/, *sly*; /smajl/, *smile*; etc.). This meant that all non-liquid, postvocalic items, to be the total reversal of their corresponding prevocalic items, would have to end in the phoneme /s/. However, this

manipulation would create a potential confounding variable, as the fricative /s/ also serves as an *allomorph* (i.e., the various forms of the same morpheme) of the morpheme of plurality. As a morphological marker, /s/ is reasonably easier to segment than other sounds in the same position. Therefore, while total reversals remain the ideal, other phonemes as close as possible in linguistic features were used in place of the phoneme /s/. The reversal of the prevocalic item /(s)mɛf/, for example, would be /fɛm(p)/ instead of /fɛm(s)/. In this example, in terms of linguistic similarity, it would seem that the alveolar stop /t/ would be a better substitute than the labial /p/, as the replaced fricative /s/ is also alveolar. Unfortunately, just like the morphological marker of plurality /s/, the phoneme /t/ as an allomorph of the past tense marker (as in *passed*) also threatens to confound the results in a similar manner. Morphological as well collocational constraints like these thus greatly limited the extent to which controls of global similarities were possible.

The creation of items became even more complicated when more than one constraint was present. Taking the creation of the nonword item /(s)kɛf/ for example, we had expected its reversal to sound like /fɛk(s)/. As the final /s/ has to be ruled out, we need to replace it with other obstruents. As the preceding consonant /k/ can be followed only by either /s/ or /t/, the collocation constraint thus leaves us with the only choice /t/. As discussed earlier, however, the stop /t/ happens to be a morphological marker, too, and substituting it for control of the confounding global similarity effect would actually create yet another confounding variable. To solve this problem, positions in the coda cluster were switched to produce a final product such as /fɛs(k)/. Adjustments like these, admittedly, might result in some imbalance between the prevocalic and postvocalic items.

However, as the cohesion between vowel and consonant has been generally found to vary only between obstruents and sonorants but not among obstruents themselves, the impact of such adjustment, if any, should be insignificant.

Predictions

The limitations of, as well as the controls incorporated into, the items having been described, this section relates the tasks to the research questions earlier raised and makes predictions about possible as well as expected results. Being the main measures, the phoneme deletion task and the phoneme isolation task have been designed to cross-validate the results of each other. Both research question one and question two, i.e., the presence of processing unit transfer and the preferred subsyllabic division unit, can be answered by either of the two tasks. It is to be noted, however, that while the presence or absence of processing unit transfer can be directly answered by the results, the interpretations concerning the preferred subsyllabic division unit will be theoretical in nature. That is, the conclusion to be made about division unit preference will draw on findings from other studies as well as those obtained in the present study. As discussed in Chapter I, this is because the complexity and the cognitive demands inherent in the tasks aimed to answer the preference question are too high to be practically implemented in a purely native language study whose target languages are likely to motivate restructuring of syllable structure. The well-recognized presence of L1 phonological transfer along with the predominantly core-syllable structure of Mandarin Chinese appears to be an alternative in offering an indirect solution. Being a transfer effect, the interpretations are necessarily inferential notwithstanding.

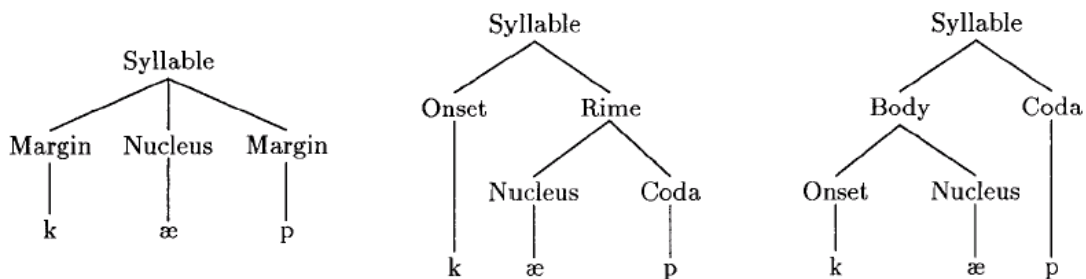
The following predictions are arranged according to the research questions asked. As question one and question two can be addressed, directly or indirectly, by the same tasks of phoneme deletion and phoneme isolation, predictions based on them were made under the same section titled *L1 Processing Unit Transfer*, which covers possible results of the planned t-tests that examine the predictions given in Appendix D. Predictions for analyses designed to answer question three, namely, multiple regression analyses, are given in the section titled *Nature of the Segmentation Performance Based on Core-syllable Awareness*. Predictions originally made for the optional study follow these two sections.

L1 Processing Unit Transfer

In addition to the three theories of subsyllabic unit preferences so far discussed, namely, onset-rime division, body-coda division, and core-syllable-plus-appendices hypothesis, there is yet another, less known theory of subsyllabic structure, the flat theory, which simply denies any subsyllabic levels intermediate between the syllable and its component phonemes (see Fig. 1, from Kessler & Treiman, 1997, for a comparison of the flat theory and the other two traditional theories). It is worthy of note that while both the

Figure 1

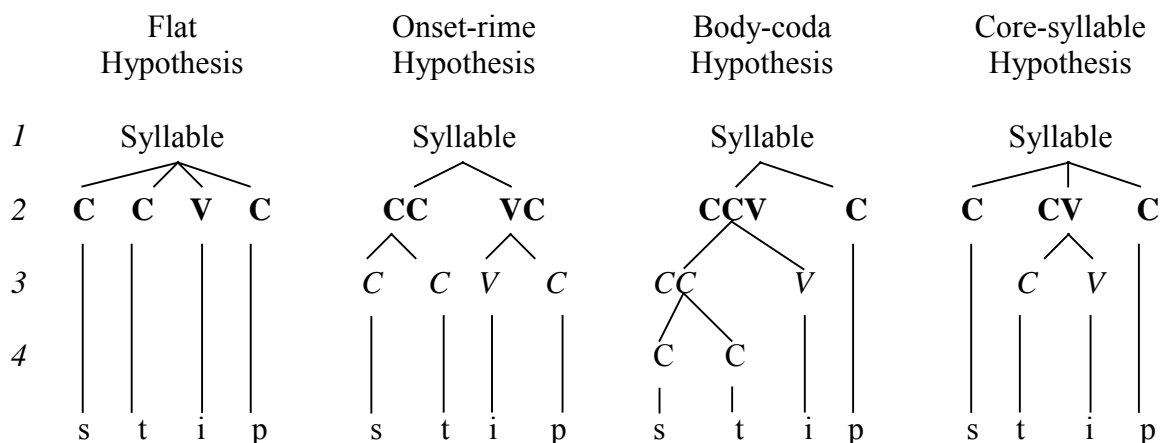
Traditional theories of syllable structure expressed in tree diagrams, taking the word *cap* for example (from Kessler and Treiman, 1997, p. 297)



onset-rime division and the body-coda division are hierarchically dichotomous down to the level *below* that of the subsyllabic units (e.g., to the coda level of the onset-rime theory or the onset level of the body-coda theory), the flat theory is non-dichotomous. That is, the syllable from the flat perspective is not necessarily branching in a binary manner, the way a syllable would be divided, in the onset-rime theory for example, into the binary onsets and rimes and rimes further into nucleus and coda (see Fig. 2). As the core-syllable hypothesis, like the flat theory, also assumes no binary branching in its syllable structure, we can categorize the four theories further into two general types in terms of branching nature: *the binary theories*, which include the onset-rime and the body-coda division theories, and *the non-binary theories*, which include the other two. As we shall see presently, this marks an interesting distinction in the criteria on which the predictions will be based.

Figure 2

Tree diagram expressions of the four hypotheses examined in the present study, taking the word *steep* (of the CCVC syllable type) for example. The numbers to the left indicates the four hierarchical levels, namely, 1. syllabic, 2. subsyllabic unit, 3. sub-unit, and 4. phonemic level.



As there are four theories, including the one presented here, of how a syllable may be sub-syllabically divided, one should also be able to predict four different outcome patterns based on the theoretical division units of each of the four candidate theories in the tasks measuring subsyllabic division preference. Four tables of prediction corresponding to the four theories are made for the obstruent items of the phoneme deletion task (see Tables A-1 to A-4, Appendix D) and for both the obstruent items (see Tables B-1 to B-4, Appendix D) and the L1-nasal items (see Tables C-1 to C-4, Appendix D) of the phoneme isolation task. Predictions, hence prediction tables, for the other items in either task were not made because of the presence of more factors than the two used in the above-mentioned predictions. However, as the above predictions were themselves capable of discerning the underlying division preference at work, they should suffice for the purpose of the present study. The performance patterns found on the items not included in the three prediction groups, moreover, might additionally shed light on the viability of certain confounding factors for future research.

Two assumptions were made as the criteria for making the predictions: (1) the *level assumption*: a target sound as a unit by itself (a between-unit segmentation) is easier to segment than a target sound which is part of the same unit (a within-unit segmentation); and (2) the *marginality assumption*: given sounds at the same level, a target sound in the marginal position is easier to segment than one in a non-marginal position. An example of the first, if the onset-rime division theory is assumed, is the easier initial-sound segmentation of words with singleton-consonant onset such as bar than that of words with complex onset such as brow (Treiman & Weatherston, 1992). An example of the latter is the assumed easier isolation of the same phoneme /r/ in marginal items like ray

than in non-marginal items like *pray*, again if onsets and rimes are the presumed phonological structure. The marginality criterion assumes that, given units of comparison at the same level, the difference in item difficulty can be determined by difference in marginality. In the example, assuming that /r/ and /p/ are units of the same level, the target sound /r/ is more easily segmented in the word *ray* than in the word *pray* because the segmentation of the latter involves more steps of cognitive manipulation than that involved in segmenting the former.

The reason for the first assumption to be labeled the “level” assumption, instead of “within- vs. between-unit” assumption, moreover, is because the concept of levels can capture the hierarchical characteristics of each theory the other cannot. Taking the CCVC syllable type for example, as Fig. 2 shows, there could be four hierarchical levels involved in the comparisons: 1. syllabic, 2. sub-syllabic unit, 3. sub-unit, and 4. phonemic level. Under the body-coda hypothesis, to illustrate, it was predicted (1) that CV(C) items would be easier to segment than CVC(C) items and (2) that (C)VC items would be easier to segment than (C)CVC items (see Table B-3). In terms of within- vs. between-unit contrast, CV(C) items should be easier because the target sound is the coda itself (as shown in the analysis [CV][(C)]) and hence its segmentation involves between-unit segmentation; whereas the target sound of the CVC(C) items is part of the coda (i.e., [CV][C(C)]) and hence it is within-unit segmentation. Comparisons in terms of the same between- vs. within-unit contrast can be made in a similar manner for the other pair as well.

The between- vs. within-unit distinction, however, cannot capture the fact that the levels involved are not the same in the two examples. In the first comparison, the target

sound of CV(C) is a unit at the 2nd level, i.e., at the sub-syllabic level (again, [₂CV][₂(C)]), whereas that of CVC(C) is a unit at the 3rd level, i.e., at the sub-unit or phoneme level (i.e., [₂CV][₂[₃C][₃(C)]]), since there are no sub-categories for the coda in the body-coda theory. The levels involved in the other example, however, are the 3rd level for the (C)VC items ([₂[₃(C)][₃V]][₂C]) but the 4th level for the (C)CVC items ([₂[₃[₄(C)][₄C]][₃V]][₂C]). The word “level” thus better captures the same contrasts than the between- vs. within-unit distinction does. It is important to note that the level differences are hypothesis-specific. That is, the levels involved depend on under which theory the predictions are made.

It is also worthy of note that, given the binary nature of the binary theories, the criterion of marginality applies only to the predictions under non-binary theories. This is because, for the relatively simple monosyllabic items adopted here, any units comprising more than two phonemes can be further subdivided into binary units until it reaches the phonemic level under binary theories (see the tree diagram for the body-coda hypothesis in Fig. 2 for example). Since marginality is meaningful only when there are more than two units at the same level and their comparisons remain at that level, marginality under the binary theories is out of question. This further simplifies the analyses in the predictions as differences in difficulties in the binary theories (i.e., the onset-rime and the body-coda theories) can always be predicted on the basis of levels alone. Taking the syllable type CV(C)C under the onset-rime division theory for example (i.e., [C][V(C)C], where brackets indicate division unit boundaries), the target sound may appear to be in a non-marginal position. As the rime (i.e., V(C)C) can be further divided into the nucleus and the coda (i.e., [V][C(C)]), the target sound is actually in a marginal, rather than non-

marginal, position as seen at the sub-unit level (i.e., [(C)C]). The marginality effect thus is irrelevant under the framework of binary theories.

Based mainly on the two assumptions discussed above, the three groups of tables (A1~A4, B1~B4, C1~C4, Appendix D) illustrate the respective predictions under the four competing theories. As some individual comparison patterns are predicted by more than one theory, only distinctive comparisons need to be discussed. An additional consideration merits discussion before the predictions are presented. While linguistically subsyllabic structure can extend to levels a step farther than that of the immediately subsyllabic units (i.e., the sub-unit levels, such as the coda level in the onset-rime theory, as shown in Fig. 2), whether the same complex levels of syllabic structure can be also applied to phonological awareness is questionable. At least in current reading literature, the only sub-unit levels (i.e., levels below the onset-rime or body-coda divisions) are phonemes themselves. Fortunately, with both the complex division of linguistic theories and the simple division available in reading literature, similar predictions are obtained concerning distinctive comparisons. The predictions reported below nevertheless represent the distinctive comparisons obtained in both contexts. Please refer to the attached tables of predictions for the following discussions.

(1) Phoneme Deletion Task: Obstruent Items

For the obstruent items of the phoneme deletion task, it is predicted that, for the C(C)VC and CV(C)C item types, (1) there should be no performance difference if the flat hypothesis is true; (2) the former should be easier than the latter if the onset-rime hypothesis is true; (3) if the latter turns out to be easier, however, both the body-coda hypothesis and the core-syllable hypothesis could be true. The underlying subsyllabic

division preferences can be readily determined if the results of (2) are obtained. However, if those of (3) are found, more comparisons have to be made that contrast the body-coda and the core-syllable hypotheses. One such comparison can be found between CVC(C) items and CV(C)C items. No differences are expected if the body-coda preference is true. If core-syllable is the underlying processing unit, on the other hand, the former should be found to be easier than the latter. Finally, if prediction (1) is true, more evidence will be needed due to the limitations of statistical inferences. That is, a non-significant result does not translate directly into a no-difference interpretation. The data from the author's prior study (Chen et al., 2005), in any case, appear to support a core-syllable hypothesis. CV(C)C items (accuracy rate 72%) were easier than the C(C)VC items (32%) on the one hand, and CVC(C) (90%) items were also found easier than the CV(C)C items (72%).

(2) Phoneme Isolation Task: Obstruent Items

Similar judgments can be made on the performance comparisons on obstruent items in phoneme isolation task. It is expected that, for the (C)VC and CV(C) comparison, (1) there should be no performance difference if the flat theory is true; (2) the former should be easier to isolate than the latter if the onset-rime account holds; and (3) if the former is more difficult than the latter, both the body-coda and the core-syllable hypotheses could be confirmed. Additional distinctive comparisons may be required. If the core-syllable hypothesis holds, (C)VC items should be more difficult than (C)CVC items, but CV(C) items should not be different from the CVC(C) items. Again, the results from the author's prior study appear to favor the core-syllable hypothesis. Although the tasks are different, phoneme deletion in the former study but phoneme isolation in this comparison, the results of the prior study should still provide some rough

reference, since phoneme deletion and phoneme isolation have always been shown to load on the same factor. Similar to the prediction for core-syllable hypothesis here, the (C)VC (mean accuracy 64%) items in the prior study were more difficult than the CV(C) items (86%). Moreover, the (C)VC items were also more difficult than the (C)CVC items (85%), but the performance on the CV(C) and the CVC(C) items were pretty similar.

(3) Phoneme Isolation Task: L1-nasal Items

Finally, for the L1-nasal items of the phoneme isolation task, the following predictions can be made. The impact of L1 nasal items on predictions occurs mainly under the core-syllable hypothesis, as now the nasal coda is considered as part of the core-syllable, but remain structurally the same under other theories. It is expected that, for the (C)VC and (C) CVC comparison, (1) there should be no performance difference if the flat theory is correct; (2) the former should be more difficult to isolate than the latter if the core-syllable account is correct; and (3) if the former is easier than the latter, both the onset-rime and the body-coda hypotheses could be correct. To distinguish the latter two, the comparisons between CVC(C) and (C)CVC items can be distinctive. It is predicted that CVC(C) items should be more difficult than (C)CVC items if the onset-rime hypothesis is true. If the body-coda hypothesis is true, the reversed pattern, i.e., CVC(C) items easier than (C)CVC items, should be observed. As there was only one item involving a nasal coda in the author's prior study, predictions on the L1-nasal items are purely theoretical. However, as earlier predictions have been pretty well supported by the data from the author's prior study, there is good reason to believe in the predictions with the L1-nasal items. Most importantly, success with L1-nasal item

predictions would greatly strengthen the core-syllable hypothesis, as it would imply that the differences are based not on subsyllabic structures but on the entire syllable as a processing unit.

Nature of the Segmentation Performance Based on Core-syllable Awareness

If the response patterns reflecting an L1 core-syllable transfer effect are obtained, it will be interesting then to further ask whether all the items measure the same ability. That is, as the non-core-syllable elements have been hypothesized as mere appendices, i.e., something attached to the core syllable rather than an integrated part, the ability to detach such elements from the core syllable should differ from the ability to segment the core syllable itself. The core-syllable items were hypothesized to reflect true phonological awareness, whereas the non-core-syllable items are not. Should this hypothesis hold, it is predicted that while core-syllable items should contribute to decoding skills, their non-core-syllable counterpart should not. This hypothesis can be translated into the following research question: Do the performance on non-core-syllable segmentation items and that on core-syllable segmentation items contribute, respectively, to decoding skills in L2 English, just as onset-rime awareness has been demonstrated to?

To answer this question, sets of multiple regression analyses were planned, with decoding skills (as reflected in nonword reading performance) as the criterion and the two contrasting variables (i.e., core-syllable and non-core-syllable items) as the predictors. To partial out the impact of vocabulary power and nonverbal intelligence, scores on PPVT-III and MAT will be first entered before all other variables. It is predicted that while core-syllable segmentation items are significant predictors, non-core-syllable segmentation items are not.

In summary, with experimental items specifically designed for the three research questions addressed here and the major confounding variables carefully controlled for, making interpretations as well as predictions of results clean and clear, we should expect a lucid picture of the L1 phonological unit transfer effect and its methodological and instructional implications. This clear picture, however, does not address research question two, which concerns the preferred subsyllabic division unit. As emphasized earlier, the results serve only to approximate the syllable division tendency of alphabetic language speakers in their early stage of language development. This limitation recognized, the empirical data of the study should nevertheless point to a clear direction where future studies in this regard might go. The following chapter on methodology describes how the data were collected and analyzed.

CHAPTER III: METHODOLOGY

Participants

Three elementary schools in Kaohsiung, Taiwan, were contacted for willingness to participate. Of the three, one consented to take part. Informed consent forms were given to teachers who volunteered to offer assistance by distributing the forms to the students for their parents' consent. Out of the 200 copies of consent forms sent out, about 80 were returned with parents' signatures of consent. Based on the available time for testing, 48 fourth-grade students were first scheduled. As the testers were inexperienced in the beginning, however, data from many subjects were either missing or contaminated by, for example, environmental noises including those caused by other students doing the morning cleaning chore required by the participating school. After some communication, the author secured the librarian's cooperation to keep the noises of the library, where the testing was conducted, to their minimum. Contaminated data, i.e., data for 15 subjects, nevertheless had to be discarded. They included, among other things, cases where obvious noises were present, those in which a wrong testing order was given, those where the subjects had an obvious cold, and those where the data were missing because the testers forgot to press the recording button for some of the tasks that required recording. Data from 48 uncontaminated subjects ($N = 48$) were collected, counterbalanced in 16 different testing orders.

Measures

Five measures were used in the study. Among the five used, three were adapted from or were subtests of standardized measures, namely, the first two sets of Peacock Picture Vocabulary Test, Third Edition (PPVT-III), for oral vocabulary; the Pattern

Completion subtest of Matrix Analogies Test for nonverbal intelligence, and the Word Attack subtest of Woodcock-Johnson III Tests of Achievement for decoding skills. The other two are experimental measures created specifically for the present study, i.e., phoneme deletion task and phoneme isolation task. Task reliabilities will be reported at the end of each measurement description; interrater reliabilities will be given in the coding section.

Oral Proficiency

An adapted version of the PPVT-III (Dunn & Dunn, 1997) was used to tap subjects' English receptive vocabulary. This adaptation involved using the first two sets of vocabulary items of the original measure as in the author's prior study (i.e., Chen et al., 2005), on which the current study has been based. In that previous study, this adaptation was shown to have sufficient variability as well as reliability. Subjects heard a test word, which had been pre-recorded by a native speaker of English, and were simultaneously presented with four pictures on a piece of A4 paper. They were asked to give the number of the picture that they thought best represented the word they had heard. Twenty-four items from the first two (out of a total of 17) sets of PPVT-III comprised the task. As mentioned before, exposure to English is very limited in an English-as-a-foreign-language environment, which has enabled these basic items to show sufficient variability. PPVT-III was used because there was no established oral vocabulary instrument tailored specifically for Chinese ESL population. The subjects obtained an average score of 14.67 out of the 24 items ($SD = 4.33$). The reliability of this test is satisfactory, with a reliability coefficient (Cronbach's α) of .78.

Nonverbal Intelligence

The Pattern Completion subtest of the Matrix Analogies Test (Naglieri, 1985) was employed as the measure of non-verbal intelligence. The subjects were shown a big picture of certain geometric forms with a missing part; under the big picture are several small pictures serving as the potential candidates for the missing part. The participant was required to select from the 5 to 6 candidate small pictures one that best completed the big picture in terms of both colors and patterns. There are a total of 16 items. An average score of 13.44 out of the 16 items has been obtained ($SD = 1.54$). The reliability of this task is surprisingly low, with a reliability coefficient (Cronbach's alpha) of .32. This is probably due to ceiling effects from two-thirds of the items: out of the 16 items, there are 3 items with a mean accuracy of 1 (i.e., all subjects answered correctly), 6 items with mean scores above .90, and 2 items with means scores above .85.

Decoding Skills

To measure the subjects' decoding skills, twenty English nonword items were taken mainly from the first twenty items of the subtest Word Attack of Woodcock-Johnson III Tests of Achievement (McGrew & Woodcock, 2001). The decision to use only twenty items was made on two grounds: the time allotted to the task were limited due to the many tasks to take in a session, and, given the participants' limited exposure to English, the inclusion of the remaining 25 Word Attack items (which are more difficult than the first 20 items) might pose great difficulties to them. Three out of the twenty items were taken out of the other 25 items, however, to avoid potential confusion. The nonword "*weat*," for example, might be read differently, depending on whether the students analogize it to the word *sweat* or to the word *heat*.

Participants were asked to read as fast and as correctly as possible the nonwords presented to them, one by one, on A-4 sheets of paper. Their responses were recorded on a digital recorder for later coding. Results of this task, indexing the subjects' decoding skills, serve as the criterion variable in the multiple regression analyses. An average score of 13.37 was obtained from a total score of 20 ($SD = 4.56$). The reliability of this task is high, with a reliability coefficient (Cronbach's alpha) of .93.

Phoneme Deletion Task

The phoneme deletion task in the current study consisted of 32 items (see Appendix A for all items). In its testing design, detailed introduction in Chinese is first given and then followed by 4 practice items before the formal items are administered. Each of the four practice items is repeated up to three times if not answered correctly. For the formal items, subjects are asked to first repeat the nonwords, e.g., “/kɪft/,” and were then asked, in Chinese, “How would you say it without pronouncing the [phoneme to be deleted; e.g., “/f/” in this example]?” The 32 items consist of four-phoneme closed syllables. They have been designed to vary along three dimensions: *sonority*, *marginality*, and *position*. While sonority is a control factor, whose 4 levels were introduced in the Overview of the Present Study section of Chapter II, marginality is defined by the to-be-deleted sound's position in terms of syllable margins, and position is categorized in reference to the syllable nuclei. There are two levels for the variable position: *prevocalic*, including item types (C)CVC and C(C)VC, if the target sound precedes the vowel nucleus, and *postvocalic*, including item types CVC(C) and CV(C)C, if the target sound follows the nucleus. Similarly, the two levels of marginality are classified according to whether the target sounds are found at the syllable margins: *marginal*, including item

types (C)CVC and CVC(C), if the target sound is in the syllable margins, or *non-marginal*, including item types C(C)VC and CV(C)C, if elsewhere. The reliability of the phoneme deletion task is satisfactory, with a reliability coefficient (Cronbach's alpha) of .74.

Phoneme Isolation Task

The phoneme isolation task follows a format similar to that given in Stahl and Murray (1994). Two major differences remain, however: while the items in their study were real words, items in the present study are nonwords controlled for sonority; and while no repetition of the target was required in the former, participants in the latter are required, following Geudens and Sandra's (2003) suggestion, to repeat the target before they attempt the answer. In this task, the participants first heard and then repeated the target nonword and were only then asked to isolate either the initial or the final sound. Each participant would, for example, first hear and then repeat the nonword /pros/ "praws", and then be asked to produce its initial sound.

Thirty-two items (see Appendix B for all items) were created along three dimensions: *sonority*, (linguistic) *complexity*, and *position*. The levels of sonority are exactly the same as those described in the previous task. The items are further classified according to *position*, i.e., whether the to-be-isolated, target sound is located before (*prevocalic*) or after (*postvocalic*) the vowel nucleus; and linguistic *complexity*, i.e., whether the target sound forms part of a consonant cluster or stands alone as the onset or the coda. Items with target sounds as singleton consonants are referred to as *simple* items; these are all three-phoneme items, including item types (C)VC and CV(C). Those with target sounds found in consonant clusters, thus forming four-phoneme syllables, are

complex items, including the item types (C)CVC and CVC(C). Phoneme isolation is administered following the same counterbalanced item orders as in phoneme deletion task. The reliability of the phoneme isolation task is also satisfactory, with a reliability coefficient (Cronbach's alpha) of .73.

In addition to serving the purpose of answering the research questions put forth in this study, all items types in both the phoneme deletion and the phoneme isolation tasks are used also because of their representativeness: the three syllable types CVC, CVCC, and CCVC included here represent 80% of all English monosyllables (De Cara & Goswami, 2002, p.417).

Procedure

As the time available for testing was limited to mainly the first hour of a typical school day and the expected subject number was 48, several testers were required. Eight undergraduates, four freshmen and four sophomores (all females), from a local language college were recruited and trained as testers. They were divided into two groups and members of each group received a training that lasted for two hours. During the training session, each was explained the tasks they were going to administer to the subjects; rationale, principles, and procedure were all given in detail for them to follow; they were encouraged to imagine what might really happen when giving the real testing and raise any question they may feel concerned about. The testers were then paired, one as the tester and the other the subject, to practice the procedure in its entirety. After the first round of practice, members of the same pair changed their roles, and went on with the second round of practice. Members of each group were further assigned the roles of one

coordinator and three testers, the duty of the former being coordinating among testers and controlling environmental noises.

The use of multiple testers necessitated an examination of its potential confounding impact. To this end, the three main variables—English nonword reading, phoneme deletion, and phoneme isolation—were respectively entered as the dependent variable and analyzed for any potential tester effects on them. The results showed no tester effects on either of the three: for nonword reading, $F(8, 39) = .830, p = .582$; for core-syllable items, $F(8, 39) = 1.284, p = .280$; and for non-core-syllable items, $F(8, 39) = .668, p = .716$.

All tasks were individually administered. There were from two to three children taking the tests simultaneously at each session. Before the testing, a coordinator would pick up the students in the classroom, escorting them back to the library, while the testers were setting up the testing equipment. The testers then went over the task with the subjects, and asked whether he or she was willing to take part in the experiment. If the subject said yes, they would provide him or her the assent form for their signature. Each subject was tested in a session of about 40 minutes with a break in-between.

Experimental tasks were given before control tasks, in the order of phoneme isolation, phoneme deletion, English receptive vocabulary (tapped by adapted PPVT-III), nonverbal intelligence (tapped by MAT: Pattern Completion), and nonword reading. Test items for the phoneme isolation task, the phoneme deletion task, and PPVT had been pre-recorded by a native speaker of English and saved as audio files in MP3 format. For the phoneme isolation and the phoneme deletion tasks, the recording was then edited, using the Adobe Audition audio editing software, to form the 16 counterbalanced orders

of testing items. Items in each one of the 16 testing orders were later administered to 3 subjects on an MP3 player. Subjects' responses to phoneme isolation task, phoneme deletion task, and nonword reading were recorded digitally for later coding. To reduce environmental noise, headphones were provided for both the tester and the testee in all tasks requiring responses to pre-recorded instruction. Abundant practice, as described earlier in the measurement section, was given before test items were administered in both experimental tasks.

Similar to the use of multiple testers, the use of different orders has also make necessary an examination of its potential confounding effects. To this end, the three main variables—English nonword reading, phoneme deletion, and phoneme isolation—were respectively entered as the dependent variable and analyzed for any potential order effects on them. The results showed no order effects on either one of the three: for nonword reading, $F(15, 32) = .686, p = .779$; for core-syllable items, $F(15, 32) = 0.464, p = .942$; and for non-core-syllable items, $F(15, 32) = .837, p = .632$.

Coding

Data collected from the two experimental tasks, i.e., phoneme deletion and phoneme isolation tasks, and the English nonword reading were coded by two native speakers of English teaching English in Taiwan. Differences were resolved by employing a third coder, who was responsible for coding only the items differently coded by the two. The third coder, however, speaks English as a second language but has an MA in linguistics from University of Texas at Arlington. Hiring a non-native speaker as the third coder was decided mainly on practical grounds. For one, it is very difficult to find English native speakers in Taiwan willing to do coding work. For another, perhaps most

importantly, as the inter-rater reliabilities for all three tasks were high (.90 for phoneme deletion; .98 for phoneme isolation; and .98 for nonword reading), any impact would be insignificant.

Coding schemes for the three tasks had been devised for all three coders to follow. The coding scheme for nonword reading is straightforward relative to that of the experimental tasks. Response to an item was scored as 1 if the whole nonword was correctly pronounced. Response to an item was scored as .5 if half or more than half of the phonemes were correctly pronounced. Response to an item was scored as 0 if less than half of the phonemes were correctly pronounced. The rationale behind this partial coding is again a result of the participants' limited exposure to English. More specifically, their partial answer was thought to reflect their partial knowledge of the letter-sound correspondence. An all-or-none coding was considered to risk underestimation of their decoding skills. Taking the bi-syllabic item, *vunhip*, for example, scoring as zero for partial answers like *vun-* may overlook the participants' decoding ability as demonstrated in reading the first part of the nonword. Taking the monosyllabic item, *raff*, for another, the decoding skill behind a partial answer like *ra-* may be better represented by partial coding than a score of zero.

Coding schemes for the phoneme isolation and phoneme deletion tasks are more complex due to the nature of L2 learning. For both tasks, the subjects received a score of 1 for each item answered in one of the following ways: (1) item repeated correctly and answered with the correct target sound segmented (i.e., isolated in isolation task or deleted in deletion task); (2) items repeated incorrectly with a non-target sound substituted for by another but answered with the correct target sound segmented (e.g.,

/s(m)ɪk/, in phoneme deletion, repeated as /s(m)ɛk/ and answered as /sɛk/); (3) items repeated incorrectly with the target sound substituted for by another sound of the same sonority category and answered with the substituted sound segmented (e.g., /s(m)ʊk/ repeated as /s(n)ʊk/ and answered as /sʊk/); (4) items repeated incorrectly with a non-contiguous consonant omitted (e.g., /s(m)ɪk/ repeated as /s(m)ɪ/, where the omitted consonant /k/ has no impact on either marginality or position, two of the three factors examined in this study; in contrast to contiguous omission, e.g., /s(m)ɪk/ repeated as /(m)ɪk/, where the missing consonant /s/ changes the non-marginal item into a marginal one) and answered with the target sound segmented. The subject received a score of 0 otherwise. As misperception and mispronunciation are all common to ESL speakers, especially in an English-as-a-foreign-language environment, both substitution, as in (2) and (3), and omission, as in (4), had been expected. Conditions (2) to (4) were considered to sufficiently reflect the subject's ability to segment the sound type under investigation and were therefore scored as correct.

Design and Data Analyses

As both phoneme deletion and phoneme isolation tasks are repeated measures, counterbalancing of the presentation of treatments was conducted to avoid any potential effect of practice, fatigue, or carry-over. The three factors with 2 x 2 x 4 levels result in a total of 16 cells. Given that 16 cells would yield tens of trillions of different presentation orders (= 16 x 15 x 14... x 1), incomplete, instead of complete, counterbalancing has been adopted. Using an even number algorithm (e.g., Girden, 1992), which follows the formula of 1, 2, n, 3, n-1, 4, n-2, ... for the first presentation order and 1 added to each of

the above numbers for the next presentation order, 16 presentation orders were obtained. As counterbalancing requires that each order be given to the same number of subjects, the sample size has to be the multiples of 16. In the present study, a sample size of 48 was used. That is, each presentation order was followed by exactly three subjects (i.e., 16 x 3).

Three groups of statistical analyses with five measures were planned to investigate the research questions raised in the Introduction. They are (1) two 4 x 2 (sonority x position) repeated-measure analyses of variance (ANOVA), one for phoneme isolation task and the other for phoneme deletion task; (2) planned t-tests corresponding to the predictions (please see Appendix D: Prediction Tables); (3) fixed-order multiple regression analyses. Marginality in phoneme deletion and linguistic complexity in phoneme isolation were not included (by excluding marginal items of the deletion task and complex items of the isolation task) in the said ANOVAs not only because they are of little theoretical interest in the present study, though both also serve as controls, but because it had been impossible to incorporate them systematically into the items due to sonority sequencing principle. For detailed discussion, please see the *Use and Creation of Measures* under the *Overview of the Present Study* in Chapter II.

Phoneme isolation and phoneme deletion tasks were designed to answer research questions one and two, which concern, respectively, L1 transfer of phonological processing units and preferred subsyllabic division units. As these two questions are examined based upon the predicted patterns as expressed in the relative accuracy among different item types, sets of planned t-tests corresponding to the predictions were thus core to the questions addressed. ANOVAs were used to test the effect reported in earlier

literature (e.g., Geudens & Sandra, 2003); that is, position effect varied with items of different sonority. As shown in the prior studies, in phoneme deletion and phoneme isolation alike, both position (i.e., onset or coda) and sonority have an effect. Position effect, however, is usually found to occur with obstruent items but not with sonorant items. Both main effects of Sonority and Position and interaction effect were, therefore, expected. It is worthy of note that results of each task alone are able to provide empirical support for the hypotheses concerning the research questions two and three. Results from both tasks thus serve to cross-validate the hypotheses.

Multiple regression analyses were planned to address research question three, namely, the nature of Chinese ESL children's segmentation performance and its contribution to their ESL decoding skills. Three sets of multiple regression analyses were planned, all with decoding skills as the criterion variable and nonverbal intelligence and oral English oral vocabulary as the control variables. That is, with decoding skills as the dependent variable, nonverbal intelligence and English oral vocabulary were entered at the first step in all three sets of analyses. In the first set of analyses, all items in both isolation and deletion tasks were entered at the second step to see if all the segmentation items as a whole contribute to decoding skills. In the second set of analyses, non-core-syllable segmentation items were entered at the second step, whereas the core-syllable segmentation items, at the third. Finally, the item types entered at the second and the third steps were reversed in their entry order in the third set of analyses, i.e., core-syllable items entered at the second step and non-core-syllable items were entered at the third. It was predicted that only core-syllable segmentation items would contribute to decoding skills, if L1 phonological processing unit transfer should occur.

CHAPTER IV: RESULTS AND DISCUSSION

Introduction

Three groups of statistical analyses were planned to address the three research questions motivated by the phonological contrasts between Chinese and English in syllable structures, namely, the presence of L1 transfer of phonological processing unit, its theoretical implications for preferred subsyllabic division units, and the relationship between the resulting segmentation performance and phonological awareness proper. Among the three groups of planned statistical analyses, the repeated-measure ANOVAs and planned t-tests for both phoneme deletion and phoneme isolation tasks were designed to address research question one and two, whereas the fixed-order multiple regression analyses were planned to address question three. The results are reported and discussed in two sections, following the same format given in *Predictions* under the section *Overview of the Present Study* in Chapter II. In the first section, *L1 Processing Unit Transfer*, results of the ANOVAs and planned t-tests are first described (with those for phoneme deletion task coming before those for phoneme isolation task), followed by discussion of research question one and two in light of the statistical analysis results. Results from the multiple regression analyses are then reviewed, followed by discussion of research question three, in the next section, *Nature of the Segmentation Performance Based on Core-syllable Awareness*.

Before the planned statistical results are reported, I will start with an examination of potential confounding effects (i.e., tester effect, gender effect, and test order effect) involved in the conducting of the experiment to ensure a comprehensive interpretation of the main results. As the research design is not completely crossed among (since it was

not designed to examine) the three potential confounding factors, a second-order effect (i.e., tester x gender x order) is not available. Without the second-order effect, it is difficult to decide whether an interaction effect (e.g., tester x gender) is dependent on another interaction effect (e.g., tester x order). However, with components in all their different combinations in the different models examined, all the second-order interaction effects as well as main effects were non-significant. Similar results were obtained with all main variables examined here: nonword reading, core syllable items, and non-core syllable items. It is therefore safe to state that the main results to be reported in the remaining sections were not confounded by tester effect, gender effect, or test order effect. Taking nonword reading as the dependent variable for example, as shown in one of the models examined, the main effects of tester, subjects' gender, and test order are all non-significant: $F(8, 23) = .66, p = .72$ for tester effect; $F(1, 23) = .07, p = .79$ for gender effect; and $F(15, 23) = .60, p = .84$ for test order effect. All second-order interactions are also non-significant: $F(5, 33) = 1.77, p = .15$ for tester x gender interaction; $F(18, 6) = .67, p = .77$ for tester x order interaction; and $F(11, 20) = .90, p = .56$ for order x gender interaction.

L1 Processing Unit Transfer

Results from Phoneme Deletion Task

A repeated measures analysis of variance for the 4 x 2 (sonority x position) design with non-marginal items (i.e., C(C)VC and CV(C)C items) of the phoneme deletion task showed a main effect of sonority, $F(3, 141) = 13.30, p < .001$; a main effect of position $F(1, 47) = 16.12, p < .001$; and an interaction effect $F(3, 141) = 21.95, p < .001$. The interaction occurred because, on the one hand, the position effect (that prevocalic items

were more difficult to segment than postvocalic ones) was significant only with obstruent items ($p < .001$) but not with items of other sonority levels; on the other hand, the sonority effect (obstruent items being more easily segmented than items of other three sonority levels) was significant ($p < .001$) only between obstruent items and items of each of the other three sonority levels, whereas sonorant items (i.e., liquid, L1-nasal, and non-L1 nasal items) did not differ from one another. These are consistent with findings from earlier studies (e.g., Geudens & Sandra, 2003; Yoon et al., 2002).

While the results from the ANOVA were consistent with those of earlier studies, the results of the planned t-tests are of central interest here. Table 1 shows accuracy rates for all item types classified according to sonority, marginality, and position. The results were consistent with all the four predictions made for obstruent items under the core-syllable hypothesis (please see Table A-4 in Appendix D). (1) There was no significant difference between marginal prevocalic items, i.e., (C)CVC items, and marginal postvocalic items, i.e., CVC(C) items ($p = .32$). (2) Non-marginal prevocalic items, i.e., C(C)VC items, were more difficult to segment than non-marginal postvocalic items, i.e., CV(C)C items, $t(1, 47) = -7.71, p < .001$. (3) Marginal prevocalic items, i.e., (C)CVC

Table 1

Mean accuracy rates for item types classified by sonority, marginality, and position in **Phoneme Deletion** task. Numbers in parentheses indicates standard deviation; mean accuracy rates related to predictions are given in boldface.

Marginality	Marginal		Non-marginal	
<i>Position</i>	Prevocalic	Postvocalic	Prevocalic	Postvocalic
<i>Sonority</i>	(C)CVC	CVC(C)	C(C)VC	CV(C)C
Liquid	.47 (.39)	.75 (.37)	.24 (.33)	.25 (.31)
L1 Nasal	.94 (.20)	.78 (.29)	.17 (.28)	.17 (.32)
Non-L1 Nasal	.95 (.19)	.69 (.39)	.12 (.24)	.17 (.32)
Obstruent	.93 (.21)	.88 (.26)	.09 (.27)	.63 (.41)

items, were easier to segment than non-marginal prevocalic items, i.e, C(C)VC items, $t(1, 47) = 18.32, p < .001$. (4) Marginal postvocalic items, i.e., CVC(C) items, were easier to segment than non-marginal postvocalic items, i.e, CV(C)C items, $t(1, 47) = 3.67, p = .001$. Results (2), that C(C)VC items were more difficult to segment than CV(C)C items, ruled out the tenability of both the flat hypothesis and the onset-rime hypothesis, since CV appears to be more cohesive than VC. Both results (3) and (4), that (C)CVC items were easier to segment than C(C)VC items and that CVC(C) items were easier to segment than CV(C)C items, ruled out a body-coda interpretation, since there should be no differences in either comparison if the body-coda hypothesis were correct. The results from the four planned pairwise comparisons thus met all the predictions made for the core-syllable hypothesis.

Results from Phoneme Isolation Task

The repeated measures analysis of variance for the 4 x 2 (sonority x position) design with simple items (i.e., (C)VC and CV(C) items) of the phoneme isolation task showed a main effect of sonority, $F(3, 141) = 19.11, p < .001$. The main effect of position was not observed ($p = .45$), but the interaction effect was significant, $F(3, 141) = 3.39, p < .05$. Similar to the ANOVA for the deletion task, the interaction occurred because, on the one hand, the position effect (that prevocalic items are more difficult to segment than postvocalic ones) was significant only with obstruent items ($p < .005$) but not with items of other sonority levels and, on the other, the sonority effect (obstruent items being more easily segmented than items of other three sonority levels) was significant ($p < .001$) only between obstruent items and items of each of the other three sonority levels, the latter of which did not differ among themselves. Despite the non-significant results of position

effect, therefore, results of phoneme isolation showed a similar pattern to that of phoneme deletion task.

For the planned t-tests, Table 2 shows accuracy rates for all item types classified according to sonority, (linguistic) complexity, and position. The results met all the predictions made for both the obstruent items (see Table B-4 in Appendix D) and the L1-nasal items under the core-syllable hypothesis (see Table C-4, Appendix D).

For the obstruent items, (1) simple prevocalic items were more difficult to isolate than simple postvocalic items (i.e., (C)VC items < CV(C) items, where the mathematical sign “<” indicates relative easiness), $t(1, 47) = -3.16, p < .005$; (2) there was no significant difference between complex prevocalic items and complex postvocalic items (i.e., (C)CVC items \approx CVC(C) items), $p = .25$; (3) simple prevocalic items were more difficult to isolate than complex prevocalic items (i.e., (C)VC items < (C)CVC items), $t(1, 47) = -3.52, p = .001$; and (4) there was no significant difference between simple postvocalic items and complex postvocalic items (i.e., CV(C) items \approx CVC(C) items), $p = .38$.

Results (1), that (C)VC items were more difficult to isolate than CV(C) items, ruled out the possibilities of both flat hypothesis and onset-rime hypothesis, since CV appeared to

Table 2

Mean accuracy rates for item types classified by sonority, linguistic complexity, and position in **Phoneme Isolation** task. Numbers in parentheses indicates standard deviation; mean accuracy rates related to predictions are given in boldface.

<i>Complexity</i>	Simple		Complex	
<i>Position</i>	Prevocalic	Postvocalic	Prevocalic	Postvocalic
<i>Sonority</i>	(C)VC	CV(C)	(C)CVC	CVC(C)
Liquid	.60 (.42)	.57 (.31)	.83 (.26)	.94 (.17)
L1 Nasal	.48 (.42)	.52 (.45)	.96 (.14)	.96 (.17)
Non-L1 Nasal	.63 (.39)	.56 (.42)	.94 (.20)	.94 (.20)
Obstruent	.74 (.36)	.94 (.20)	.95 (.19)	.90 (.23)

be more cohesive than VC. Results (3), that (C)VC items were more difficult to isolate than (C)CVC items, ruled out a body-coda interpretation, since the opposite results should obtain if the body-coda hypothesis were correct. The overall patterns as shown by the four planned pairwise comparisons, on the other hand, are consistent with all the predictions made for the core-syllable hypothesis.

For the L1 nasal items, there were no significant differences between prevocalic and postvocalic items, simple and complex items alike ($-1 < ts < 1$). However, simple items in both positions, i.e., prevocalic and postvocalic, were all more difficult to isolate than the linguistically complex items: for prevocalic items, (C)VC < (C)CVC, $t(1, 47) = -7.82, p < .001$; for postvocalic items, CV(C) < CVC(C), $t(1, 47) = -6.05, p < .001$. While the comparison results between simple and complex prevocalic items ruled out all the other rival hypotheses, all comparison results are consistent with predictions made for the core-syllable hypothesis.

Discussion: L1 Transfer of Phonological Processing Unit

All the possible predictions (i.e., predictions not confounded by factors in opposite directions) made for the core-syllable hypothesis were confirmed by the statistical analyses, indicating that the Chinese ESL children did process English syllables in terms of a core syllable plus appendices. Although no significant differences were observed between L1 nasal items and non-L1 nasal items, indicating that the transfer effect may have been confounded by the sonority effect at the nasal level, the transfer effect of L1 phonological processing unit remains strong. Given the dominance of core syllables in Chinese, this result should not be surprising. When Chinese is the only language acquired, in fact, phoneme division either at the subsyllabic level or the finer

phonemic level may be simply irrelevant, since there are only about 400 syllable types to acquire. A similar comment has been made about Japanese, a language whose syllable structure is similar to that of Chinese: “it is clear that an onset-rime division of the syllable appears irrelevant to Japanese” (Ziegler & Goswami, 2005, p. 7). The demand for sensitivity to syllable structure seems to come to Chinese children only with the learning of a second language with complex phonological structure, such as English.

The position effect on the part of obstruent items (i.e., the immediately prevocalic items being more difficult to segment from the nucleus vowel than the immediately postvocalic ones), however, may be explainable on phonetic grounds (e.g., Browman & Goldstein, 1989, 1988). Browman and Goldstein (1988), for example, argue that, at least in English, prevocalic

consonant clusters ... are almost completely overlapped by the vocalic gesture... [whereas postvocalic] consonants ... are aligned [in time] so that the target of the consonant is first attained just as the vocalic gesture is turned off. Thus, the target portions of final consonants are produced in their own time frame, while the target portions of initial consonants overlap the time frame for the vowel. (p. 98).

That is, prevocalic consonants appear to behave differently from postvocalic ones.

Brown and Goldstein further observe that the duration of the body unit (i.e., the prevocalic consonant(s) plus the nucleus vowel) appears to be a constant if it is measured not from any of the individual prevocalic consonants but from the global C-center (“C” for consonant) that is derived from computations taking all prevocalic consonants into account. Prevocalic consonants thus behave as a unit, in contrast to their postvocalic counterparts which “do not act as a unit” (Browman & Goldstein, 1988, p. 99). In a nutshell, the easier segmentation of postvocalic consonants may have resulted from their clearer perceptual boundary from the vowel.

While the acoustic explanation can account for why postvocalic consonants are perceptually salient as individual sounds, hence behaving as appendices to the core syllables, it does not explain why complex prevocalic obstruents (as in (C)CVC items) should be more easily isolated than simple prevocalic obstruents (as in (C)VC items) as observed in Chinese ESL children. If prevocalic consonants should act as a unit, (C)CVC items should be more difficult, instead of easier, to isolate than (C)VC items, since the former involves intra-unit segmentation, whereas the latter involves between-unit segmentation. Indeed, this prediction of prevocalic consonants as a unit has found support elsewhere with languages with more complex syllable structures. Using a deletion task with Dutch children, for example, Bertelson et al. (1997) found that the initial sound of (C)VCC items were more easily deleted than that of (C)CVC items. Similar results were obtained with studies using similar items on Czech-speaking and Anglophone children as well (e.g., Caravolas & Bruck, 1993). Note that the (C)VCC items used in their study are, in theory, similar to the (C)VC items tested here, though the tasks were different. These two item types in the two studies are comparable because the final C in their (C)VCC items had no impact on the status of the initial C as the onset itself—in (C)VC items used here, the initial C is also the onset itself. Moreover, phoneme deletion and phoneme isolation have long been found to measure the same factor, i.e., phonological awareness. Granted the comparability, the discrepancy in the applicability of the phonetic account—that it successfully predicts the Dutch data but not the Chinese ESL children data—further indicates that transfer effect of L1 phonological processing unit did occur and override the bondage among prevocalic consonants.

The different behaviors as observed in the tautosyllabic initial and final consonant clusters are also noticed by other linguists, most notably Fujimura and his colleagues (e.g., Fujimura, 1979, 1988, 1975; Fujimura & Lovins, 1978). Fujimura and Lovins (1978) illustrated that, at least in English, the same phonemic segments can show different, ad hoc phonetic properties in syllable initial and final positions. Taking the production of the voiceless stops /p, t, k/ for example, when in syllable-initial position, the glottis always remains wide-open *before* the articulatory release of the blocked air. The status of the glottis *at the time of* air release, however, differs a great deal (closed, somewhat open, or open), depending on whether the stops are preceded by the phoneme /s/ (with the glottis nearly closed, as in producing the unaspirated [p] in *speak*); begin a stressed syllable (with glottis open, as in producing the aspirated [p^h] ¹² in *peak*); and/or are followed by liquids (i.e., /l/ and /r/, in which case the glottis is relatively more open). In contrast to their context-sensitive variations in initial position, the same voiceless stops in syllable-final position are accompanied by glottis that “may or may not open for articulatory release” (p. 110). Similarly, the distinctive feature *voicing* that distinguishes voiced consonants (e.g., /b, d, g/) from their voiceless counterparts (e.g., /p, t, k/) is also almost non-distinctive when the stops are found syllable-finally.

Based on such observations, Fujimura and Lovins (1978) argued for a syllable structure that consists of, at least for English, a syllable core and phonetic affix(es), with the syllable core further divisible into two demisyllables (i.e., the initial and the final

¹² The superscript “^h” indicates aspiration—the blowing of an extra puff of air because of the open glottis. The extra puff of air can be felt if one places a piece of paper in front of one’s mouth when speaking words that begin with an aspirated stop, e.g., *peak*. When preceded by the phoneme /s/, /p/ becomes unaspirated when produced, and thus lack the extra puff of air observed in producing an aspirated [p^h].

demisyllables) with a cut in the nucleus vowel¹³. In his analyses of English syllables, Fujimura (1979) found that some apical consonants (i.e., consonants produced with the tip of the tongue, such as /t, d, s, z/) in syllable-final position appear to be more stable and independent, thus more readily detachable, than others. He termed such consonants *affixes* and the body they affix themselves to the *syllable core*. For an apical consonant to be an affix, at least three conditions must be met: (1) it must be linked to the (syllable) core-final elements; (2) agree in voicing with the core-final elements; and (3) be an apical obstruent. Thus, for example, while the final phoneme /d/ in *tend* is an affix, its counterpart /t/ in *tent* is not (since the voiceless /t/ does not agree with the voiced, core-final element /n/ in voicing) and belongs to the syllable core. The distinction is made on phonetic grounds: the articulations of *ten-* with or without the *affix* /d/ are very similar, whereas those of *ten-* with or without the *core-final element* /t/ are quite different—the vowel in the latter case is about one-third shorter in duration and the nasal murmur (i.e., the acoustic presence of the nasal sound) is almost absent (but perceivable in the nasality of the nasalized vowel) with the presence of /t/ than without it. More precisely, while the phoneme /d/ as an affix has little impact on how the syllable core is pronounced, the phoneme /t/ as part of the syllable core influences how other parts of the syllable core would be articulated and thus is an integral part of the syllable core.

The further division of the syllable core into the initial and the final demisyllables is considered mainly from the perspectives of speech synthesis and speech recognition. For synthesized speech to be natural, concatenative phonetic units (i.e., units used as the basis in sequencing sound segments into speech) must be relatively stable in their

¹³ The division of the two demisyllables during the vowel has the minimal effects of coarticulation, hence variation. Taking the syllable core /tent/ for example, its two demisyllables, initial and final respectively, are /tɛ/ and /ɛnt/, with the nucleus vowel /ɛ/ belonging to both demisyllables.

variations under different phonological contexts. The phonological unit phoneme as a concatenative phonetic unit has been found to be too unstable because of its ad hoc variations discussed earlier. Syllables instead appear to be a better candidate. Syllable as a concatenative unit, however, is not without a problem, given the huge size of syllable inventory in English (more than 10,000 items), hence potential difficulties in speech processing. Nevertheless, when only the syllable core is considered, the inventory is reduced to contain less than 5,000 items. Most importantly, when the syllable core is further divided into the two demisyllables, the number is further reduced to 1,000 or fewer items (see p. 475, Fujimura, 1979). Demisyllables thus appear to be the most efficient unit in speech synthesis or recognition.

If the English syllable consists of a syllable core and phonetic affixes, it is possible that the results obtained in this study may be confounded by test items that contain affixal elements, whose relative readiness for detachment may account for some of the patterns observed. An examination of the items, however, excludes such a possibility. As discussed in the measure section, items in both phoneme deletion and isolation tasks can be classified as of four sonority levels, namely, liquid, L1 nasal, non-L1 nasal, and obstruent. Due to sonority sequencing constraint, the sonority levels of the items are defined at the immediately prevocalic or postvocalic element (except for those at the obstruent level, to which we will return shortly; see discussions in the measure section). Taking the liquid items of the $C_1VC_2(C_3)$ type for example, the sonority level of liquid refers to the immediately postvocalic consonant (i.e., C_2), instead of the target (i.e., C_3). Since all liquid and nasal sounds are necessarily voiced but the target sounds in this study have all been limited to voiceless consonants, the final sounds of non-obstruent

items cannot be affixes, which by definition must agree with the core-final element in voicing. Thus at least for non-obstruent items, the core-affix distinction can not have confounded the results.

Obstruent items are more complicated in this respect. All the relevant obstruent items used here have been limited to those ending either in /sk/ or /sp/ for reasons discussed in the measure section. It is possible therefore that such syllable-final elements may function as affixes. This possibility, however, is also ruled out on at least two grounds. For one, both /k/ and /p/ are not apical consonants that partially define affixes. For another, the behaviors of consonant pairs that begins with the phoneme /s/ and ends with voiceless stops (i.e., /sp/, /st/, /sk/) have long been found to be unique in English. At least in terms of speech synthesis and recognition, it is implied that such consonant clusters appear to function as a unit both syllable-initially and syllable-finally (Fujimura & Lovins, 1978) in such a way that, “[i]n effect, /p/, /sp/, /b/ form a three-way distinction” (Fujimura, 1979, p. 474). As /sp/ and /sk/ are considered as single units in both syllable-initial and syllable-final positions, it appear unlikely that they should lead to different degrees of difficulties in segmentation, unless other factors come in, such as the L1 transfer effect hypothesized here.

Discussion: Implications for the Preferred Subsyllabic Division Units Debate

Given the demonstrated core syllable transfer, the next issue to tackle is how the results bear on the debate over preferred subsyllabic division units as reviewed, for example, by Ziegler and Goswami (2005). CV as an intact unit, whether interpreted as body or core syllable preference, has been found in much research studying languages of simpler syllable structure, such as Japanese and Korean (e.g., Kubozono, 1996; Yoon et

al., 2002). However, the observed preference is not limited to simple-syllable languages, as similar results have also obtained in studies with many European languages, including English, Dutch, and French (e.g., Johnston, Anderson, & Holligan, 1996; Geudens & Sandra, 2003; Martinot & Gombert, 1996), whose syllable structures are complex. Martinot and Gombert (1996), for example, examined French preliterate children aged 4 to 6 years old on their free segmentation of CVC words. Their results showed that only about 4% of the children produced a –VC (initial phoneme deletion) response, whereas as high as 71% of the children responded with a CV– (final phoneme deletion) answer. Using a deletion task that included both syllable and phoneme deletion, moreover, Johnston and her colleagues (1996) examined English-speaking four-year-old non-readers and obtained a similar result: final phonemes (16%) were far more easily deleted than initial ones (2%).

It thus seems that, in addition to syllable structure complexity, age also plays an important part in deciding how a syllable is subdivided. In the studies discussed above, subjects demonstrating a tendency to treat CV as an intact unit were either speakers of simple-syllable languages or young pre-literate children. As Geudens and Sandra (2003) observed, having reviewed related literature, “It is notable that these children and our participants, showing a CV preference in explicit awareness tasks, had less experience with the teaching of reading and writing than participants showing a rime preference” (p. 175). Several factors related to age, as they went on, may change children’s preferences: literacy, emphasis on particular types of phonological units (such as onsets and rimes) in the curriculum, and the statistical properties of the language acquired. These all favor a developmental account of preferred subsyllabic division units, where a learner begins

with core syllable as the preferred syllable structure and restructures his phonological syllable only later when factors associated with literacy and vocabulary come in.

Seen from the broader framework of cross-language studies, then, a developmental account would place all languages on a continuum of syllable complexity, with core syllables only on the one end and the most complex syllable structure on the other. For languages closer to the complex syllable end, syllable restructuring is more likely to occur with language development for reasons just described, which may result in a preference for onset-rime division as demonstrated in the majority of studies with languages of complex syllable structures. For languages closer to the core-syllable end, in contrast, their simple syllables (often core syllables) would remain the phonological processing unit in actual use even after their speakers have reached maturity in language development. Even with the acquisition of a second language whose syllable structure is complex, the force of the intact core syllable may, depending on L2 proficiency and exposure, remain strong and result in transfer, as evidenced in the present study as well as other studies with languages of simple syllable structures.

A similar account has also been found predominant in related phonological studies following the optimality theory (OT) tradition, which stipulates that while there is a universal set of violable constraints shared by all languages of the world, languages differ in the orders such constraints are ranked in terms of their priority in being met (see Prince & Smolensky, 2004). As discussed in Chapter II, such constraint-based theory is especially powerful in accounting for developmental or second language acquisition phenomena, as both of which can be simply interpreted in terms of changes or transitions from one rank order of constraints to another. Levelt et al. (2000), for example,

demonstrated how stages could be discerned from developmental data by re-ranking the constraints at each stage. Of particular interest to the present study is the argument that the initial state of syllable structure in young children is characterized by the high ranking of three structure constraints—(1) that syllable should have an onset, (2) that syllables should not have a coda, and (3) that syllables should not have complex onsets or codas. In the initial stage of phonological development of syllable structure, such constraints must be met even at the expense of faithfulness to the input. Thus, for example, upon hearing a monosyllabic word of the phonological structure CVC , , children in the initial stage of language development would respond with CV when asked to reproduce the word. This is because, at this developmental stage, the structure constraint that syllables should not have a coda outranks the faithfulness constraint that the input (i.e., what has been heard) should correspond to the output (i.e., the reproduced syllable). In later stages, development takes the form of re-ranking of the constraints. In the final stage for example, the faithfulness constraint that requires input-output correspondence would outrank the structure constraints; speakers at this stage would thus produce (i.e., the output) exactly the sounds they have heard (i.e., the input), despite violations of the said structure constraints.

Note that meeting all the three high-ranking structure constraints in the initial stage results in only one possible syllable type, i.e., the core syllable, CV. The acquisition of syllable structure thus begins with the universal, unmarked (i.e., the most natural) core syllable and proceeds toward language-specific structures, which, being particular to the language being acquired, are less natural and take longer to acquire. According to Levelt et al. (2000), moreover, re-ranking of the constraints starts with

comparing one's own language with the language surrounding one, and stops when no differences are found between the two. Given the simplicity of Chinese syllable structure, which is the environment with which a beginning learner compares his language, it follows naturally that Chinese speakers do not have to go through as many stages of acquisition as speakers of languages with more complex syllable structures do, and the constraint ranking in Chinese speakers should demonstrate a more unmarked one—that is, it should place avoidance of onsetless syllables, complex syllables, and syllables with coda (which amounts to the core syllable) high on the ranking. From the perspective of optimality theory, then, the L1 transfer of core syllable preference evidenced in this study can also be interpreted as the transfer of the high ranking of the unmarked syllable structure constraints. This should apply not only to Chinese, as demonstrated in the present study through L1 transfer, but also to other languages with simple syllable structures such as Japanese and, of particular interest here, Korean.

In contrast to the large body of literature consistent with a developmental account, studies evidencing a body-coda account appear scant. In Levelt et al.'s work (2000), the syllable type CVC was found to develop much earlier than the CCVC type. If, as in Dutch, the syllable type CVC also develops earlier than CCVC type in Korean and development stops when there exist no differences between the language input from the environment and the output one actually produces, there seems no reason why the relatively more intact unit of CV should be interpreted as the unit of body rather than core syllable. This is because the syllable development should have stopped once the CVC syllable type has been acquired; and since there is no CCVC type in the Korean-speaking environment, there appears no conditions under which a body unit (which, again, should

involve not only CV, but CCV, CCCV, etc) should develop. Core syllable preference is thus a far more plausible cause of the results found in earlier studies where CV appeared to be more cohesive than VC.

Implications for Future Studies

If the developmental account should hold, Korean should also demonstrate the same core syllable preference instead of the body-coda preference as claimed in previous studies (e.g., Derwing et al., 1993 ;Yoon et al., 2002; Yoon & Derwing, 2001). Unlike Chinese, however, the observed CV preference in Korean may not result from its syllable complexity but from morphophonemic factors. Korean shares a similar syllable structure to that of Chinese, i.e., (C)(G)V(C) (Korean: Kabak, 2003; Sohn, 1999; Chinese: Li, 1999), except for possible members of the coda. That is, while the coda in a Chinese syllable is limited to two nasals, the Korean coda includes both sonorants and obstruents (/m, n, ŋ, l, p, t, k, s/). Given the nature and great number of the Korean coda, preference for CV-C or C-VC division is equally possible. The instability of Korean syllable codas, however, may have motivated the demonstrated preference for CV-C division.

The instability of the Korean syllable coda comes from an interaction of two factors, morphophonemics and agglutinative morphology. In Korean, morphophonemic representations of a syllable are not necessarily identical to its phonetic representations (Sohn, 1999). The single-syllable morpheme *kaps* “price” is one example. While morphophonemically coda clusters such as the “ps” are allowed in its morpheme root form, in actual speech a coda cluster is never realized. That is, in actual speech, one never speaks or hears any syllable with a coda cluster. The morpheme-final consonant is either re-syllabified as the onset of the suffix syllable when followed by a derivational or

inflectional suffix that begins with a vowel or, in all other contexts, simply "erased" (the so-called "stray erasure;" Sohn, 1999, p.170). Thus, for example, the morpheme-final morphophonemic element "s" in *kaps* is re-syllabified as the onset of the inflectional suffix *ɪl* (object) to form a bisyllabic word with the re-syllabified pronunciation [kap.s'ɪl] "price (object)", where the dot "." signifies the syllable boundary. On the other hand, when standing alone or followed by another word, *kaps* is simply pronounced with the morpheme-final "s" deleted: [kap]. In Korean, there are many such phonological rules that govern the phonetic realization of morphophonemic coda consonants (see Sohn, 1999, p.163-195), and the sound to be altered is not limited to the last of the coda cluster, that is, given the same example of *kaps*, it is sometimes the "p" that is subject to change.

The rule-governed nature of the coda in a Korean syllable is thus likely to confer a distinctive status on the coda in Korean's syllable structure, thus making an onset-rime division unfavorable. Added to this instability of the Korean coda, moreover, is Korean's "agglutinative" morphology, where "a long chain of particles or suffixes with constant form and meaning may be attached to nominals... or predicate...stems... as in *ka-si-ess-keyss-sup-ni-ta* '(a respectable person) may have gone'" (Sohn, 1999, p.15). The agglutinative nature of Korean morphology apparently would require high-frequency use of morphophonemic rules, and the Korean coda as a distinctive entity can only be strengthened.

The special status of coda in Korean does not automatically translate into a preference for body-coda division, however. The theoretical arguments for the hierarchical division into bodies and codas (e.g., Yoon & Derwing, 2001) have been insufficiently supported, as the empirical evidence provided could be interpreted as a

preference for either body-coda division or simply for core syllables. Despite the differences in coda members, Chinese and Korean remain very similar in their composition of syllables aside from the coda, i.e., C(G)V. Given the lack of empirical evidence for the body-coda division preference in Korean and given the demonstrated core syllable tendency in Chinese and the similarity between the two languages in terms of its CV structure, it would be theoretically interesting to conduct a study on Korean syllable structure, with a design similar to the present study, to investigate the actual nature of subsyllabic division unit preference in Korean. Such a study would potentially offer more conclusive evidence on the nature of the preferred subsyllabic division units.

Nature of the Segmentation Performance Based on Core-syllable Awareness

The demonstrated core syllable preference in the Chinese ESL children makes even more meaningful the third research question put forth in the present study. More specifically, if syllable components other than the core syllable are merely appendices, segmentation of them should not reflect a true phonological awareness. If this is true, only performance on segmenting the core syllable but not that on segmenting other elements should contribute to decoding skills. Performance on non-core-syllable segmentation, that is, should have no impact on decoding skills. This hypothesis was tested with multiple regression analyses, whose results are reported below.

Results from the Multiple Regression Analyses

Three sets of fixed-order regression analyses were performed to compare the relative contribution, or non-contribution, of the hypothesized core-syllable and non-core-syllable items to decoding skills (please see Table 3 below). All items from both isolation and deletion tasks were included. For each set, PPVT and Matrix scores were

first entered to partial out the impact of receptive vocabulary and non-verbal intelligence. The normal probability plot (Figure 3) shows that normality of the residual distribution can be assumed. An examination of the residual plot (Figure 4) also indicates that the assumptions of linearity, homoscedasticity, random and independent errors, and a mean of zero for errors conditioned on the predictors seem to hold across all their values. Variance inflation factors (VIF) for all predictors are equal or less than 2, indicating that multicollinearity should not be a concern, either.

Figure 3

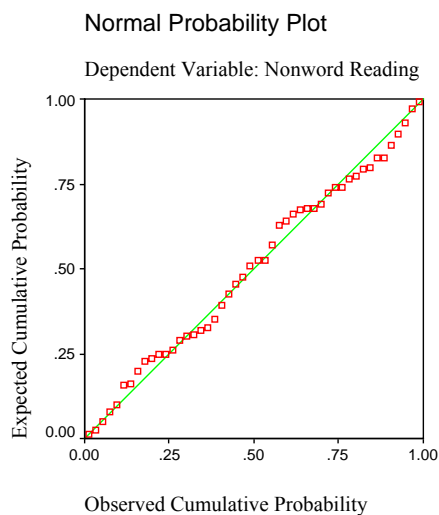


Figure 4

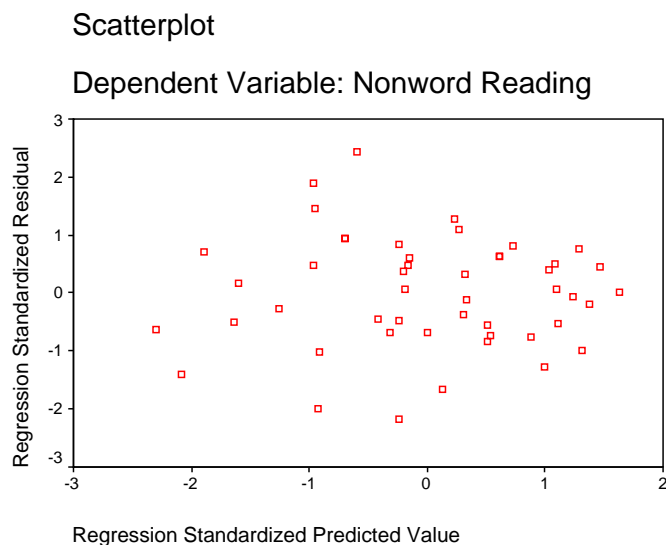


Table 3 shows the results of multiple regression analyses. As shown in the table, oral vocabulary and non-verbal intelligence as represented, respectively, by PPVT and Matrix, alone accounted for 31% of the total variance in English nonword reading. The three sets differ in the item types entered at step 2 and, if any, step 3. In the first set, both core syllable and non-core syllable item totals were entered at step 2 to examine if all the items, regardless of their types, as a whole contribute to English nonword reading. As the results showed, the *R squared change* was significant, $F(2, 43) = 6.011, p < .01$, showing that performance on phoneme deletion and isolation as a whole does seem to make unique contribution to decoding skills.

Based on the prediction that only core-syllable items would contribute to English nonword reading, two more sets of analyses were performed to examine the separate contributions of core-syllable and non-core-syllable items. In set two, non-core-syllable items were entered at step 2 and core-syllable items, at step 3. The results showed that

Table 3

Fixed-order regression analyses for assessing the relative contributions of core-syllable and non-core-syllable segmentation items (from both deletion and isolation tasks) to English decoding skills.

Predictors of English Decoding Skills						
Order of Entry	Variables	R^2	Adjusted R^2	R^2 change	F change	* <i>p</i> -value
Step 1	PPVT and Matrix	.314	.284	.314	10.320	<.001
Step 2	Both Item types	.464	.414	.150	6.011	<.01
Step 1	PPVT and Matrix	.314	.284	.314	10.320	<.001
Step 2	Non-core-syllable Total	.383	.341	.069	4.899	<.05
Step 3	Core-syllable Total	.464	.414	.081	6.511	<.05
Step 1	PPVT and Matrix	.314	.284	.314	10.320	<.001
Step 2	Core-syllable Total	.458	.421	.143	11.623	<.002
Step 3	Non-core-syllable Total	.464	.414	.007	.525	>.470

**p*-value here refers to the change statistics (i.e., F change).

the *R squared change* was significant for both non-core-syllable items, $F(1, 44) = 4.899$, $p = .032$, and core-syllable items, $F(1, 43) = 6.511$, $p = .014$. While the results do show that core-syllable items make unique contribution to decoding skills beyond and above that contributed by non-core-syllable items, oral vocabulary, and non-verbal intelligence, it is not certain if non-core-syllable items alone also make unique contribution, as their effect could have been mediated by that of core-syllable items.

To examine whether the significant result of non-core-syllable items is unique or mediated by core-syllable items, the variable entry order was reversed in set 3. That is, core-syllable items were entered at step 2 and non-core-syllable items, at step 3. As the results show, once the impact of core-syllable items is partialled out, the contribution of non-core-syllable items is no longer significant, $F(1, 43) = .525$, $p = .473$. In fact, the adjusted R square decreases at step 3, from .421 at step 2 to .414 at step 3. This indicates that the significant effect on non-core-syllable items as observed in set 2 was actually mediated by core-syllable items; non-core-syllable items themselves did not contribute to decoding skills. Core-syllable items, on the other hand, alone accounted for 14 percents of the total variance in decoding skills, $F(1, 44) = 11.623$, $p = .001$. The results are thus consistent with the hypothesis that core-syllable items alone represent true phonological awareness, whereas non-core-syllable items do not.

If non-core-syllable items do not belong in the same construct of phonological awareness as their core-syllable counterparts do, this should have further implications for test reliabilities as well. In other words, measurement errors should be abundant for non-core-syllable items but much less for core-syllable ones. Further analyses of reliabilities were conducted to examine this hypothesis. As Table 4 shows, when reliability

coefficients were calculated for each item type (classified by position (2 levels) and marginality for deletion task or complexity for isolation task (2 levels), ignoring sonority levels), only the reliabilities for core-syllable items are satisfactory (.72 for deletion task and .74 for isolation task). Reliabilities for non-core-syllable items, on the other hand, appear substantially lower, ranging from .38 to .60.

Discussion

Results from both multiple regression analyses and reliabilities have shown a picture consistent with the predictions made for the implications of the core syllable hypothesis. That is, while performance on items requiring core-syllable segmentation reflects true phonological awareness, performance on non-core-syllable items does not seem to belong in the same construct. This has both educational and methodological implications. Educationally, it suggests that emphasis should be placed on the segmentation of core syllables, if Chinese ESL students are to be trained in their sensitivity to phonological structure. Moreover, the subsyllabic units of onsets and rimes need to be stressed as the more natural subsyllabic division, as the sound-letter correspondence at the levels of onsets and rimes have been shown to be more consistent than correspondence at the grapheme-phoneme level. Methodologically, test items should be carefully selected if phonological awareness is to be measured on Chinese EFL

Table 4

Reliabilities by item types; item types in boldface and italicized indicates core-syllable segmentation items (item number = 4, which is also the 4 levels of sonority)

Task	Phoneme Deletion				Phoneme Isolation			
Item Types	(C)CVC	CVC(C)	<i>C(C)VC</i>	CV(C)C	<i>(C)VC</i>	CV(C)	(C)CVC	CVC(C)
Alpha	.60	.45	.72	.47	.74	.44	.38	.58

students. At least, core-syllable versus non-core-syllable items should be distinguished to examine if the L1 transfer effect is present.

CHAPTER V: CONCLUSION

Introduction

This dissertation study investigated the potential L1 transfer of phonological processing units and examined its theoretical, methodological, and educational implications. The results revealed a picture very much like what had been predicted according to the core-syllable transfer hypothesized here. That is, Chinese ESL students do seem to process an English syllable in terms of a core syllable plus its appendices. Probably due to the inherent simple structure of Chinese syllables, there has been little research that investigates the nature of Chinese speakers' sensitivity to complex syllables. Among the few studies that have, usually with other theoretical foci, similar response patterns have been reported, though interpreted differently.

Huang and Hanley (1995), for example, found that "English subjects were superior to children from Hong Kong when asked to delete the first sound from a CVCC word, whereas Hong Kong children were significantly better than English subjects on first sound deletion from words of the CCVC type" (p. 95). Note that the British subjects in Huang and Hanley's study were much better than their Hong Kong peers in both English reading (34.84 compared to 16.98, out of 100) and receptive vocabulary (17.64 compared to 5.02, out of 32), whereas their performance on phoneme deletion was, in contrast, puzzlingly close (34.09 vs. 32.40). The picture drawn from these figures is obviously incongruent with that drawn from the mainstream literature; and the incongruence appears to lie in the divergent performance of the Hong Kong subjects.

When the performance of the Hong Kong subjects on phoneme deletion was examined alone, the response patterns were identical to those of phoneme isolation

reported here (the item types of phoneme deletion task in Huang & Hanley, 1995, were similar to those of phoneme isolation task used here). That is, (C)VCC items were more difficult to delete (4.59 out of 10) than both CCV(C) items (9.60) and (C)CVC items (8.62). In regard to the somewhat surprising easiness of initial sound deletion from (C)CVC items for the Hong Kong students (which for their British peers were more difficult), Huang and Hanley suggested that

[t]he Chinese tend to introduce a vowel after the initial consonant in English words of this kind [i.e., CCV]. If the Hong Kong children are implicitly doing something similar when they hear such words, then the phoneme deletion task in reality becomes a syllable deletion task (p. 95-96).

In other words, the (C)CVC items were more easily segmented because its initial consonant was deleted as if it was a syllable. If this was the case, we should expect Chinese ESL children to pronounce the initial C of (C)CVC items as a syllable in an *isolation* task. Note that there was no way to find out if this is true in a deletion task, as used in Huang and Hanley's study. The voice responses to the isolation items in the present study, however, did not appear to support the interpretation. That is, the initial consonant of the (C)CVC items were generally pronounced by the Taiwanese subjects here as singleton consonants rather than as syllables. In fact, even if the Hong Kong subjects did add a vowel after the initial consonant as Huang and Hanley hypothesized, it would still be consistent with the core-syllable hypothesis. This is because the initial, non-core consonant would be treated as a separate unit, whose phonological structure would also be core syllable in nature. The comparative easiness of (C)CVC items in contrast to (C)VC items reported in Huang and Hanley's study thus should reflect the same core-syllable transfer demonstrated by the Taiwan children here.

Theoretical Implications

The first important implication from the observed L1 transfer effect concerns the debate over the preferred subsyllabic division units. Although the preference for onset-rime division has been shown to hold for a great many alphabetic languages, its universality has also been challenged on solid empirical grounds (e.g., Derwing et al., 1993; Geudens & Sandra, 2003; Yoon et al., 2002; Yoon & Derwing, 2001; see Ziegler & Goswami, 2005 for a review). At least for Korean, given a CVC syllable, the initial consonant has been repeatedly found to be more difficult to segment than the final one. This CV-C division tendency was usually treated as demonstrating a body-coda division preference in the hierarchical fashion that onset-rime division had been theorized in linguistics. Yoon and Derwing (1993), for example, argue that while syllables in some languages such as English are right-branching in structure, the Korean syllable is left-branching. That is, given the body-coda (CV-C) division in Korean, the further division (branching) occurs in the “body” unit (C-V), which is to the left of the syllable, hence left-branching. This is theoretically appealing from a Universal Grammar (UG) point of view, one of whose principles is simplicity. More specifically, the idea that syllable structure of a language is either left-branching or right-branching (except for moraic languages such as Japanese, which is beyond the scope the present discussion) would mean that acquisition of the syllable structure is almost effortless, which is consistent with the speed and ease observed in young children’s acquisition of their mother tongue.

Theoretically appealing as it may seem, the empirical evidence supporting the body-coda hypothesis has been inadequate. After all, given the simple syllables used in prior studies, the same results that have been argued to demonstrate the body-coda

division preference can be equally argued to show a core-syllable preference. The inadequacy in evidence supporting the body-coda division hypothesis, moreover, occurs to coda as well as body as a unit intermediate between the syllable and the individual phonemes. In Korean, the language on which the majority of studies arguing for a body-coda preference have been based, a coda cluster is possible only in morphophonemic forms but never in actual speech. Morphophonemic rules such as re-syllabification or stray erasure, as discussed in Chapter IV, that are responsible for the conversion are thus likely to make the syllable-final consonants unstable as a unit. In contrast, the onset as hypothesized in the onset-rime hierarchy for languages like English has been well received as a unit by itself on both theoretical and empirical grounds (e.g., Bertelson et al., 1997; Browman & Goldstein, 1988; Johnston et al., 1996; Martinot & Gombert, 1996).

The tendency for CV-C division in simple-syllable languages such as Korean would appear to receive the least empirical as well as theoretical resistance, on the other hand, if interpreted under the developmental framework, i.e., as evidence to core-syllable awareness. As the first syllable type acquired developmentally (Levelt et al., 1999; Levelt, Schiller, & Levelt, 2000) and the universally preferred syllable type (Benson, 1988), the core syllable should remain an intact unit for some time developmentally, i.e., at least in the initial stages of language development (Geudens & Sandra, 2003). Applied in the case of Korean, a developmental account would simply treat the observed CV-C division preference as a continuation of the universal preference for core syllable without having to postulate any additional changing mechanism. In contrast, a body-coda account would require a strong theoretical argument for any factors that should motivate the transition from the simple core syllable to the more complex body unit, which would

include not only CV but also more complex subsyllabic structures such as CCV, CCCV, etc., which, however, is obviously lacking in native Korean speaking environment. If, as Levelt et al. (2000) has argued, language acquisition stops when the learner no longer perceives differences between his/her own language and that found in the environment, Korean speakers should have reached the final stage of acquisition after they have acquired all the syllable types found in Korean (which are simple), i.e., CV, CVC, V, VC. In order for a body-coda account to be viable as an explanation of the found CV-C results, therefore, stronger theoretical argument appears mandatory. In sharp contrast to the heavy burden (but lack) of proof imposed on body-coda account, a core-syllable interpretation appears to fall much more naturally as a coherent part of the picture drawn from the available data from earlier studies.

As reviewed in Chapter IV, moreover, Chinese and Korean share a very similar syllable structure, except for the sonority and number of the coda members. That is, they both have the maximal syllable structure of CVC and differ only in the nature and possible members of the final consonant. This similarity in the CV part between the two simple-syllable languages plus the lack of evidence as to why the core syllable in Korean should evolve to become a hierarchical subsyllabic unit, i.e., the hypothesized body, would make a stronger argument that the same core syllable preference found for Chinese, as demonstrated here, should also apply to Korean. With the advance represented by this dissertation study, therefore, the picture concerning the preferred subsyllabic division units appears clearer. Future studies with Korean speakers, however, are still needed preferably with a research design similar to the present study. Such studies should come

up with more conclusive evidence concerning the nature of the preferred subsyllabic division.

Methodological Implications

Another implication of the L1 core syllable transfer findings concerns the validity and reliability of items measuring phonological awareness in a Chinese ESL context. The findings that core-syllable items alone, but not non-core-syllable items, contribute to English nonword reading imply that the inclusion of both core and non-core syllable items would introduce a large amount of measurement errors and make the tests much less reliable. That is precisely what has been found in the present study (see Table 4, repeated here as Table 5). This unreliability of the non-core syllable items would carry its effect in blurring the prediction of decoding skills: while the core-syllable items alone did contribute to English nonword reading, the contribution became non-significant when both core-syllable and non-core-syllable items were entered as one predictor.

The same problem may also help to explain some of the puzzling results obtained in earlier studies. On the English phoneme task in Huang and Hanley's study (1995), for example, while the English subjects were superior to their Hong Kong peers in performance on both English reading (34.84 vs. 16.98) and receptive vocabulary (17.64 vs. 5.02), their performance on phoneme deletion is strangely close (34.09 vs. 32.40).

Table 5

Reliabilities by item types; item types in boldface and italicized indicates core-syllable segmentation items (item number = 4, which is also the 4 levels of sonority)

Task	Phoneme Deletion				Phoneme Isolation			
Item Types	(C)CVC	CVC(C)	<i>C(C)VC</i>	CV(C)C	<i>(C)VC</i>	CV(C)	(C)CVC	CVC(C)
Alpha	.60	.45	<i>.72</i>	.47	<i>.74</i>	.44	.38	.58

Given the well-established contribution of phonological awareness to English reading, these results were surprising. As discussed earlier, this is apparently due to the fact that while the British children outperformed their Hong Kong counterparts on items whose target sounds were the onset itself (i.e., (C)VCC items), they were also inferior in performance on items whose target initial sound was part of the onset (i.e., (C)CVC). The differences, however, also corresponded perfectly to the contrast in terms of the core-syllable (i.e., (C)VCC items) vs. non-core-syllable (i.e., (C)CVC items) distinction. If indeed core-syllable transfer underlies the performance differences, we should be able to further hypothesize that the predictions of English word reading from performance on such items should be significant for the English subjects but not for their Hong Kong counterparts. An examination of Huang and Hanley's results suggests that this may be the case: while English phoneme deletion was highly correlated with English reading for the English subject ($r = .59$; $p < .001$), it was non-significant for the Hong Kong subjects ($r = .29$; *NS*). Thus it is recommended that, in future studies with Chinese ESL learners or other ESL learners whose mother tongue was simple in syllable structure, items created/employed for tapping explicit phonological awareness (segmentation, isolation, or deletion) should be controlled for their target sound's position relative to the core syllable, either by limiting the items to the core-syllable segmentation type only or imposing further classification based on the core- and non-core syllable distinction.

Educational Implications

The contribution to English reading being limited to core syllable items also has practical implications for Chinese ESL education in terms of both assessment and instruction. While it is possible to diagnose learners' sensitivity to syllable structures by

their performance on core-syllable items alone, the performance discrepancy between core- and non-core syllables and its direction may also serve as a probe to ESL learners' stage of proficiency. This is possible given the different performance patterns as demonstrated by native speakers and L2 learners in, for example, Huang and Hanley's research (1995). That is, with development towards better L2 proficiency, the discrepancy in performance between certain specific core- and non-core-syllable items (e.g., (C)VC and (C)CVC items) may decrease. This is because, as Huang and Hanley (1995) have observed, the relative difficulty between such pairs is opposite in direction between English speakers and Chinese ESL students. With increased English proficiency, the ESL learners may pick up the characteristics of English in its subdivision of the syllable into onsets and rimes, hence making onsets of singleton consonant (e.g., words of CVC structure) easier to segment and, as a result, decreasing the performance differences between core- and non-core-syllable items.

On the other hand, since Chinese ESL children appear to find core-syllable items the most difficult to segment but segmentation abilities on such items are critical to beginning literacy, the implication of the core-syllable transfer for instruction is that instruction that sensitizes learners to the divisibility of core syllables may be desirable. Moreover, if this is tenable, it would be also interesting to investigate if Zhuyin or Pinyin instruction is also predictive of English nonword reading. As Zhuyin is necessarily instructed by dividing the core syllable into onsets and rimes, its instruction could potentially contribute to Chinese ESL learners' phonological awareness. Users of Pinyin (taught mainly in mainland China) have also been reported to be instructed in a similar

fashion of division to that of Zhuyin instruction and thus should also find Pinyin instruction conducive to English decoding skills.

In summary, the contribution of the present study is manifold. In addition to demonstrating L1 transfer of phonological processing unit, this study also has implications for psycholinguistic theory, methodology, and education. Theoretically, a developmental account of the preferred subsyllabic division units appears to be better supported by the results of the present study. Methodologically, it is suggested that phonological awareness of Chinese ESL children could be more reliably measured if only core-syllable segmentation items are involved. Educationally, while distinction between the core- and non-core-syllable types could serve as diagnostic tools for Chinese ESL learners' sensitivity to English syllables, the potential impact of Zhuyin or Pinyin instruction on English phonological awareness also seem to deserve further investigation. A series of further studies is therefore expected to ensue from this dissertation study.

APPENDIX A: PHONEME DELETION ITEMS

Marginality	Marginal		Non-marginal	
Position	Prevocalic	Postvocalic	Prevocalic	Postvocalic
Sonority (practice items)	*(C)CVC P1 /s)pæf/	CVC(C) P2 /fəs(p)/	C(C)VC P3 /p(l)ʊk/	CV(C)C P4 /kɪ(f)t/
Liquid	1. /f)rɛp /; 2. /s)læz/	3. /pɛr(f)/; 4. /zæl(f)/	5. /f(r)ɪp/; 6. /s(l)ʌz/	7. /tɪ(r)f/; 8. /zʌ(l)f/
L1 Nasal	9. /s)mɪk/; 10. /s)nɛf/	11. /hɪ(k)/; 12. /fɛŋ(k)/	13. /s(m)ʊk/; 14. /s(n)ɪv/	15. /pʊ(ŋ)k/; 16. /vɪ(ŋ)k/
Non-L1 Nasal	17. /s)mɪf/; 18. /s)nɛp/	19. /fɪm(p)/; 20. /sɛm(p)/	21. /s(m)æf/; 22. /s(n)ɪp/	23. /fæ(m)p/; 24. /tɪ(m)p/
Obstruent	25. /s)kɛ/; 26. /s)pʊk/	27. /fɛs(k)/; 28. /kʊs(p)/	29. /s(k)ʊt /; 30. /s(p)ɛt/	31. /kʊ(s)k/; 32. /tɛ(s)p/

*Parentheses indicate the target sound to be deleted.

Pronunciation key:

Vowels

/ɪ/ as the “i” in “his”

/æ/ as the “a” in “mask”

/ɛ/ as the “e” in “set”

/ʊ/ as the “oo” in “book”

/ɔ/ as the “aw” in “law”

/ʌ/ as the “u” in “hut”

Consonants

/dʒ/ as the “j” in “john”

/tʃ/ as the “ch” in “chair”

/ʒ/ as the “s” in “vision”

/ʃ/ as the “sh” in “shoe”

/ð/ as the “th” in “that”

/θ/ as the “th” in “thank”

/ŋ/ as the “ng” in “thank”

APPENDIX B: PHONEME ISOLATION ITEMS

(Linguistic) Complexity		Simple		Complex	
Position		Prevocalic	Postvocalic	Prevocalic	Postvocalic
Sonority		*(C)VC	CV(C)	(C)CVC	CVC(C)
(practice items)		P1 (/ (v) Δ p /)	P2 (/ p Δ (v) /)	P3 (/ (k) l Δ f /)	P4 (/ f Δ l (p) /)
Liquid		1. / (r) Γ θ /;	3. / θ Γ (r) /;	5. / (f) Γ Γ s /;	7. / s Γ Γ (k) /;
		2. / (l) Δ p /	4. / p Δ (l) /	6. / (s) l Δ k /	8. / k Δ l (f) /
L1 Nasal		9. / (n) u p /;	11. / p u (n) /;	13. / (s) n u t /;	15. / t u η (k) /;
		10. / (m) ϵ k /	12. / k ϵ (η) /	14. / (s) m ϵ t /	16. / t ϵ η (k) /
Non-L1 Nasal		17. / (m) Γ f /;	19. / f Γ (m) /;	21. / (s) m Γ k /;	23. / k Γ m (p) /;
		18. / (n) Δ v /	20. / v Δ (n) /	22. / (s) n Δ v /	24. / v Δ m (p) /
Obstruent		25. / (t) ϵ θ /;	27. / θ ϵ (t) /;	29. / (s) k u f /;	31. / f u s (k) /;
		26. / (f) u p /	28. / p u (f) /	30. / (s) p ϵ t /	32. / t ϵ s (p) /

*Parentheses indicate the target sound to be isolated.

Pronunciation key:

Vowels

/ɪ/ as the “i” in “his”

/æ/ as the “a” in “mask”

/ɛ/ as the “e” in “set”

/ʊ/ as the “oo” in “book”

/ɔ/ as the “aw” in “law”

/ʌ/ as the “u” in “hut”

Consonants

/dʒ/ as the “j” in “john”

/tʃ/ as the “ch” in “chair”

/ʒ/ as the “s” in “vision”

/ʃ/ as the “sh” in “shoe”

/ð/ as the “th” in “that”

/θ/ as the “th” in “thank”

/ŋ/ as the “ng” in “thank”

APPENDIX C: SIMILARITY JUDGMENT ITEMS (OPTIONAL STUDY ONLY)

Practice Items	C ₁ V: /kɪb/: /kɪg/, /pʌd/	VC ₂ : /bɪk/: /gɪk/, /dʌp/	V: /kɪb/: /tɪg/, /pʌd/	C ₁ _C ₂ : /kɪb/: /kɔb/, /pʌd/
Sonority	Shared Units			
	C ₁ V (body)	VC ₂ (rime)	V (nucleus vowel)	C ₁ _C ₂ (random consonants)
Liquid	/lʌt/: /lʌp/, /mɛk/ /rʊv/: /rʊʒ/, /nɪð /	/tʌl/: /pʌl/, /kɛm/ /vʊr/: /zʊr/, /ðɪm/	/lʌt/: /nʌp/, /mɛk/ /vʊr/: /zʊm/, /ðɪm/	/lʌt/: /lʊt/, /mɛk/ /vʊr/: /vɛr/, /ðɪl/
L1 Nasal	/nɛd/: /nɛf/, /mʊθ/ /nʊb/: /nʊtʃ/, /mɛdʒ/	/dɛŋ/: /fɛŋ/, /θʊn/ /bʊn/: /tʃʊn/, /dʒɛŋ/	/nɛd/: /mɛf/, /mʊʃ/ /bʊn/: /tʃʊl/, /dʒɛr/	/dɛŋ/: /dʊŋ/, /gɛn/ /bʊn/: /bɔn/, /dɛŋ/
Non-L1 Nasal	/mɪp/: /mɪð/, /nʊʒ/ /mɛk/: /mɛv/, /nɪð/	/pɪm/: /ðɪm/, /zʊn/ /kɛm/: /vɛm/, /ðɪm/	/mɪp/: /lɪʃ/, /ræʒ/ /sɛm/: /vɛn/, /ðɪm/	/mɪp/: /mʊp/, /næʃ/ /kɛm/: /kʊm/, /tɔn/
Obstruent	/pɛʃ/: /pɛð/, /tʌv/ /ʃæʃ/: /ʃæv/, /vɛð/	/ʃɛp/: /ðɛp/, /vʌt/ /tʃæʃ/: /væʃ/, /ðɛv/	/pɛʃ/: /kɛð/, /tʌv/ /tʃæʃ/: /væθ/, /ðɛf/	/pɛʃ/: /pʌʃ/, /tʊf/ /tʃæʃ/: /tʃɪʃ/, /dʒɛθ/

Pronunciation key:

Vowels

/ɪ/ as the “i” in “his”

/æ/ as the “a” in “mask”

/ɛ/ as the “e” in “set”

/ʊ/ as the “oo” in “book”

/ɔ/ as the “aw” in “law”

/ʌ/ as the “u” in “hut”

Consonants

/dʒ/ as the “j” in “john”

/tʃ/ as the “ch” in “chair”

/ʒ/ as the “s” in “vision”

/ʃ/ as the “sh” in “shoe”

/ð/ as the “th” in “that”

/θ/ as the “th” in “thank”

/ŋ/ as the “ng” in “thank”

APPENDIX D: PREDICTION TABLES

Marks used in the comparison tables (on the following pages):

- * Brackets indicate hypothetical units; parentheses indicate target (to-be-deleted) sound.
- ** Comparison based on relative easiness (accuracy rate)
- *** Indicating the decision bases of the predictions, in accordance with the assumptions and criteria discussed in the *Predictions* section of *Overview of the Present Study* in Chapter II.
- **** “x” indicates uncertain predictions, due to effects opposite in directions.

Phoneme Deletion Task (Obstruent items)

Table A-1: Flat Hypothesis

	Pair	Hypothetic Units			Factor
Position	Marginal	*(C)CVC	**≈	CVC(C)	Level 2 = Level 2; Both marginal
Prevocalic vs. Postvocalic	Non- marginal	C(C)VC	≈	CV(C)C	Level 2 = Level 2; Both nonmarginal
Marginality	Prevocalic	(C)CVC	>	C(C)VC	Level 2 = Level 2; *** Marginal > non-M
Marginal vs. Non-marginal	Postvocalic	CVC(C)	>	CV(C)C	Level 2 = Level 2; Marginal > non-M

Table A-2: Onset-rime Hypothesis

	Pair	Hypothetic Units			Factor
Position	Marginal	*[(C)C][VC]	>	[C][VC(C)]	Level 3 > Level 4
Prevocalic vs. Postvocalic	Non- marginal	[C(C)][VC]	>	[C][V(C)C]	Level 3 > Level 4
Marginality	Prevocalic	[(C)C][VC]	≈	[C(C)][VC]	Level 3 = Level 3
Marginal vs. Non-marginal	(onset) Postvocalic (rime)	[C][VC(C)]	≈	[C][V(C)C]	Level 4 = Level 4

Table A-3: Body-coda Hypothesis

	Pair	Hypothetic Units			Factor
Position	Marginal	[(C)CV][C]	<	[CV][C(C)]	Level 4 < Level 3
Prevocalic vs. Postvocalic	Non- marginal	[C(C)V][C]	<	[CV][(C)C]	Level 4 < Level 3
Marginality	Prevocalic	[(C)CV][C]	≈	[C(C)V][C]	Level 4 = Level 4
Marginal vs. Non-marginal	(body) Postvocalic (coda)	[CV][C(C)]	≈	[CV][(C)C]	Level 3 = Level 3

Table A-4: Core-syllable Hypothesis

	Pair	Hypothetic Units			Factor
Position	Marginal	(C)[CV]C	≈	[CV]C(C)	Level 2 = Level 2; Both marginal
Prevocalic vs. Postvocalic	Non- marginal	C[(C)V]C	<	[CV](C)C	Level 3 < Level 2
Marginality	Prevocalic	(C)[CV]C	>	C[(C)V]C	Level 2 > Level 3; Marginal > non-M
Marginal vs. Non-marginal	Postvocalic	[CV]C(C)	>	[CV](C)C	Level 2 = Level 2; Marginal > non-M

Phoneme Isolation Task (Obstruent items)

Table B-1: Flat Hypothesis

	Pair	Hypothetic Units			Factor
Position: Prevocalic vs. Postvocalic	Simple	*(C)VC	**≈	CV(C)	Level 2 = Level 2
	Complex	(C)CVC	≈	CVC(C)	Level 2 = Level 2
Complexity: Simple vs. Complex	Prevocalic	(C)VC	≈	(C)CVC	Level 2 = Level 2
	Postvocalic	CV(C)	≈	CVC(C)	Level 2 = Level 2

Table B-2: Onset-rime Hypothesis

	Pair	Hypothetic Units			Factor
Position: Prevocalic vs. Postvocalic	Simple	*[(C)][VC]	>	[C][V(C)]	Level 2 > Level 3
	Complex	[(C)C][VC]	>	[C][VC(C)]	Level 3 > Level 4
Complexity: Simple vs. Complex	Prevocalic	[(C)][VC]	>	[(C)C][VC]	Level 2 > Level 3
	Postvocalic	[C][V(C)]	>	[C][VC(C)]	Level 3 > Level 4

Table B-3: Body-coda Hypothesis

	Pair	Hypothetic Units			Factor
Position: Prevocalic vs. Postvocalic	Simple	[(C)V][C]	<	[CV][(C)]	Level 3 < Level 2
	Complex	[(C)CV][C]	<	[CV][C(C)]	Level 4 < Level 3
Complexity: Simple vs. Complex	Prevocalic	[(C)V][C]	>	[(C)CV][C]	Level 3 > Level 4
	Postvocalic	[CV][(C)]	>	[CV][C(C)]	Level 2 > Level 3

Table B-4: Core-syllable Hypothesis

	Pair	Hypothetic Units			Factor
Position: Prevocalic vs. Postvocalic	Simple	[(C)V]C	<	[CV](C)	Level 3 < Level 2
	Complex	(C)[CV]C	≈	[CV]C(C)	Level 2 = Level 2
Complexity: Simple vs. Complex	Prevocalic	[(C)V]C	<	(C)[CV]C	Level 3 < Level 2
	Postvocalic	[CV](C)	≈	[CV]C(C)	Level 2 = Level 2

Phoneme Isolation Task (L1-nasal Items)

Table C-1: Flat Hypothesis

	Pair	Hypothetic Units			Factor
Position: Prevocalic vs. Postvocalic	Simple	*(C)VC	**≈	CV(C)	Level 2 = Level 2;
	Complex	(C)CVC	≈	CVC(C)	Level 2 = Level 2
Complexity: Simple vs. Complex	Prevocalic	(C)VC	≈	(C)CVC	Level 2 = Level 2
	Postvocalic	CV(C)	≤	CVC(C)	Level 2 = Level 2; (Sonority: S < NS)

Table C-2: Onset-rime Hypothesis

	Pair	Hypothetic Units			Factor
Position: Prevocalic vs. Postvocalic	Simple	*[(C)][VC]	>	[C][V(C)]	Level 2 > Level 3; (Sonority: NS > S)
	Complex	[(C)C][VC]	>	[C][VC(C)]	Level 3 > Level 4
Complexity: Simple vs. Complex	Prevocalic	[(C)][VC]	>	[(C)C][VC]	Level 2 > Level 3
	Postvocalic	[C][V(C)]	**** x	[C][VC(C)]	Level 3 > Level 4; (Sonority: S < NS)

Table C-3: Body-coda Hypothesis

	Pair	Hypothetic Units			Factor
Position: Prevocalic vs. Postvocalic	Simple	[(C)V][C]	<	[CV][(C)]	Level 3 < Level 2;
	Complex	[(C)CV][C]	<	[CV][C(C)]	Level 4 < Level 3
Complexity: Simple vs. Complex	Prevocalic	[(C)V][C]	>	[(C)CV][C]	Level 3 > Level 4
	Postvocalic	[CV][(C)]	x	[CV][C(C)]	Level 2 > Level 3; (Sonority: S < NS)

Table C-4: Core-syllable Hypothesis

	Pair	Hypothetic Units			Factor
Position: Prevocalic vs. Postvocalic	Simple	[(C)VC]	≈	[CV(C)]	Level 3 = Level 3;
	Complex	(C)[CVC]	≈	[CVC](C)	Level 2 = Level 2
Complexity: Simple vs. Complex	Prevocalic	[(C)VC]	<	(C)[CVC]	Level 3 < Level 2
	Postvocalic	[CV(C)]	<	[CVC](C)	Level 3 < Level 2; (Sonority: S < NS)

APPENDIX E: GLOSSARY OF LINGUISTIC TERMS

Body: the part of a syllable that consists of the nucleus vowel and the consonant or consonant cluster that precedes it;

Coda: in a syllable, the consonant or consonant cluster after the nucleus vowel;

Fricative: a sound of audible friction produced with the airflow forced through a narrow constriction formed in the vocal tract;

Liquid: a sonorant of less sonority than vowels but more than obstruents, as the English sounds /l/ and /r/.

Nasal: a speech sound produced with airflow going through both vocal and nasal tracts;

Obstruent: a speech sound produced with a certain degree of constriction in the vocal tract;

Onset: in a syllable, the consonant or consonant cluster before the nucleus vowel;

Phoneme: the smallest abstract unit in the sound system of a language that distinguishes meaning, usually indicated by enclosure between slashes, e.g., /p/;

Rime: the part of a syllable that consists of the nucleus vowel and the consonant or consonant cluster that follows it;

Sonorant: a speech sound characterized by relatively free airflow in its production, including vowels, liquids (e.g., the initial sound of *like*), and nasals;

Sonority: the degree of vowel-likeness;

Stop: an obstruent produced with the airflow fully stopped before its release;

APPENDIX F: AN OPTIONAL STUDY

If the results of the study had required that a possible Zhuyin instruction effect be ruled out, an optional study had been planned with kindergartener subjects who had not been taught any Zhuyin. However, as some of the tasks given to the 4th graders could be too difficult, and others simply impossible, for the kindergarteners, task adjustments would have been necessary. The *phoneme deletion* task, to begin with, would have been replaced by a *similarity judgment* task, where a subject hears a target word followed by two choice words, and must judge which of the choice words is more similar to the target. As the kindergarteners are not expected to read, moreover, nonword reading would not have been administered.

In order to address the research questions, however, the *phoneme isolation* task for the 4th graders would have been retained for kindergarteners. Fortunately, while the majority of phoneme awareness tasks, such as phoneme deletion, segmentation, and blending, have been found to be too cognitively demanding to administer to pre-readers, phoneme isolation tasks have appeared to be the only exception. Stahl and Murray (1994), for example, administered four phoneme awareness tasks to 113 kindergarteners and first graders. While the average scores for the other three tasks were all well below 3 (out of a maximal 5; 2.02 for phoneme segmentation; 2.44 for phoneme deletion; 2.69 for phoneme blending), phoneme isolation alone was found to be by far the easiest, with a mean score as high as 4.02. Similar results were obtained in other studies as well (e.g., Yopp, 1988). Besides its relative easiness for pre-readers, the phoneme isolation task is selected also for its high predictive power of decoding skills. Yopp (1988), for example, found phoneme isolation to have the highest predictive correlation with rate of learning

new words among the 10 phonological awareness tasks used. With phoneme deletion replaced by a similarity judgment task on the one hand, and nonword reading removed on the other, the kindergarteners of the optional study would thus have been left with only two tasks: similarity judgment and phoneme isolation.

With different and limited measures, however, the strength of evidence would also have been more limited. *With* the phoneme deletion task in *the study*, for example, the L1 transfer of phonological processing units can be so measured with the number of phonemes held constant across all item types, i.e., either CVCC or CCVC. *Without* the phoneme deletion task in the *optional study*, on the other hand, the burden of evidence falls on phoneme isolation task, which does not control for the number of phonemes. Moreover, the results of the phoneme isolation task in the optional study are burdened with the explanation of both L1 transfer effect (research question one) and nature of preferred subsyllabic division units (research question three). The similarity judgment task, in contrast, serves only to cross-validate the interpretation of preferred units along with phoneme isolation task and to strengthen the evidence for L1 transfer effect. In other words, the focus of the optional study is placed squarely on the phoneme isolation task. Moreover, as decoding skills are not measured, hence no criterion variable, the hypothesis about the nature of the core-syllable awareness (research question two) is solely answered by factor analysis but not multiple regression analyses. As the argument for the hypotheses of the three research questions remain relatively the same with the phoneme isolation task, I will discuss, in what follows, only hypotheses about the results of the similarity judgment task.

Similarity Judgment Task

Pre-reading kindergarteners, if necessary, would have been recruited to examine their subsyllabic division unit preferences as well as phonological processing unit transfer. As one of the only two tasks administered to the pre-reading subjects, the similarity judgment task asks the participants to judge which of the two choice nonwords is more similar to the target nonword. All the nonword items are of the syllable type CVC. One control factor and one experimental factor would have been examined. The control factor is again sonority, whereas the experimental factor is the shared units between the target and the key choice. Four levels of shared units would have been distinguished: body (CV), rime (VC), nucleus (V), and random consonants (C_C). Like all other experimental tasks, vowel length and global similarities would also have been controlled in addition to sonority. It would have been expected that the children should make more correct judgment on items sharing the body units than on items sharing others. The results, however, would have offered two competing interpretations as discussed in Chapter I—indicating either a preference for the body units or simply a preference for the core syllables—and hence would have to be interpreted along with the results from the phoneme isolation task.

Procedure

Other tasks being the same, this subsection introduces the procedure specific to the similarity judgment task. In this task, the participant first hears the target nonword, which is followed by two choice nonwords. The participant is required to judge which of the two choice nonwords sounds more like the target. Thirty-two formal items and four practice items (see Appendix C for all items), all of the syllable pattern of CVC, were

created to vary along two dimensions: *sonority* and *shared units*, each with four levels.

Shared units refer to the phonemes shared by the target and the key, and its four levels are: CV, VC, V, and C_C. While the focus is on the first two types, i.e., CV and VC, the other two types, i.e., V and C_C, serve mainly as reference.

REFERENCES

- Adams, M. J. (1994). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.
- Akamatsu, N. (1999). The effects of first language orthographic features on word recognition processing in English as a second language. Reading and Writing, 11, 381-403.
- Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. Language Learning, 53, 207-231.
- Anderson, J. I. (1987). The markedness differential hypothesis and syllable structure difficulty. In G. Ioup and S. Weinberger (Eds.), Interlanguage phonology: The acquisition of a second language sound system (pp. 279-291). Cambridge, MA: Newbury House Publishers.
- Benson, B. (1988). Universal preference for the open syllable as an independent process in interlanguage phonology. Language Learning, 38, 221-242.
- Bertelson, P., de Gelder, B., & van Zon, M. (1997). Explicit speech segmentation and syllabic onset structure: Developmental trends. Psychological Research/Psychologische Forschung, 60, 183 - 191.
- Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. Scientific Studies of Reading, 9, 43-61.
- Blevins, J. (1996). The syllable in phonological theory. In J. A. Goldsmith (Ed.), The handbook of phonological theory (pp. 206-244). Cambridge, MA: Blackwell.

- Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read - A causal connection. Nature, 301, 419-421.
- Broselow, E. (1983). Non-obvious transfer: On predicting epenthesis errors. In S. M. Gass and L. Selinker (Eds.), Language transfer in language learning (pp. 269-280). Rowley, MA: Newbury House.
- Broselow, E. (1984). An investigation of transfer in 2nd language phonology. IRAL-International Review of Applied Linguistics in Language Teaching, 22, 253-269.
- Broselow, E., Chen, S.-I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. Studies in Second Language Acquisition, 20, 261-280.
- Browman, C. P. C., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. Haskins Laboratories Status Report on Speech Research, 93-94, 85-102.
- Browman, C. P. C., & Goldstein, L. (1989). Articulatory gestures as phonological units. Phonology, 6, 201.
- Bus, A. G. (1985). Preparatory reading instruction in kindergarten. Some comparative research into methods of auditory and auditory-visual training of phonemic analysis and blending. Pedagogische Studien, 62, 269-279.
- Caravolas, M., & Bruck, M. (1993). The effect of oral and written language input on children's phonological awareness - A cross-linguistic study. Journal of Experimental Child Psychology, 55, 1-30.
- Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? Cognition, 91, 77-111.

- Chen, A. S. W., Wang, M., & Cheng, C. X. (2005). Is there an L1 orthographic transfer effect? Unpublished manuscript.
- Chen, J. Y. (2000). Syllable errors from naturalistic slips of the tongue in Mandarin Chinese. Psychologia, 43, 15-26.
- Chen, J. Y., Chen, T. M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. Journal of Memory and Language, 46, 751-781.
- Chen, J. Y., Lin, W. C., & Ferrand, L. (2003). Masked priming of the syllable in Mandarin Chinese speech production. Chinese Journal of Psychology, 45, 107-120.
- Chen, P. (1999). Modern Chinese: History and sociolinguistics. Cambridge: Cambridge University Press.
- Chen, T.-M., Dell, G. S., & Chen, J. Y. (2004, August). A cross-linguistic study of phonological units: Syllables emerge from the statistics of Mandarin Chinese, but not from the statistics of English. Paper presented at the meeting of the 26th Annual Meeting of the Cognitive Science Society, Chicago.
- Cheung, H., & Chen, H. C. (2004). Early orthographic experience modifies both phonological awareness and on-line speech processing. Language and Cognitive Processes, 19, 1-28.
- Cheung, H., Chen, H. C., Lai, C. Y., Wong, O. C., & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. Cognition, 81, 227-241.

- Cisero, C. A., & Royer, J. M. (1995). The development and cross-language transfer of phonological awareness. Contemporary Educational Psychology, 20, 275-303.
- Comeau, L., Cormier, P., Grandmaison, E., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. Journal of Educational Psychology, 91, 29-43.
- De Cara, B., & Goswami, U. (2002). Similarity relations among spoken words: The special status of rimes in English. Behavior Research Methods Instruments & Computers, 34, 416-423.
- De Cara, B., & Goswami, U. (2003). Phonological neighbourhood density: Effects in a rhyme awareness task in five-year-old children. Journal of Child Language, 30, 695-710.
- Derwing, B. L., Yoon, Y. B., & Cho, S. W. (1993). The organization of the Korean syllable: Experimental evidence. In P. M. Clancy (Ed.), Japanese and Korean linguistics, volume 2 (pp. 223-238). Stanford, CA: Center for the study of language and information.
- Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test-Third Edition. Circle Pines, MN, American Guidance Service.
- Durgunoglu, A. Y., Nagy, W. E., & Hancinbhatt, B. J. (1993). Cross-language transfer of phonological awareness. Journal of Educational Psychology, 85, 453-465.
- Eckman, F. R. (2004). From phonemic differences to constraint rankings: Research on second language phonology. Studies in Second Language Acquisition, 26, 513-549.

- Flege, J. E. (1987). The production of new and similar phones in a foreign-language: Evidence for the effect of equivalence classification. Journal of Phonetics, 15, 47-65.
- Fujimura, O. (1975). Syllable as a unit of speech recognition. IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-23, 82-87.
- Fujimura, O. (1979). An analysis of English syllables as cores and affixes. Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung, 32, 471-476.
- Fujimura, O. (1988). Some remarks on consonant clusters. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, 22, 57-64.
- Fujimura, O., & Lovins, J. B. (1978). Syllables as concatenative phonetic units. In A. Bell and J. B. Hooper (Eds.), Syllables and segments (pp. 107-120). New York: North-Holland.
- Geudens, A., & Sandra, D. (2003). Beyond implicit phonological knowledge: No support for an onset-rime structure in children's explicit phonological awareness. Journal of Memory and Language, 49, 157-182.
- Girden, E. R. (1992). ANOVA: repeated measures. Newbury Park, Calif., Sage Publications.
- Goldsmith, J. A. (1990). Autosegmental and metrical phonology. Oxford, England: Blackwell.
- Gombert, J. E. (1996). What do children do when they fail to count phonemes? International Journal of Behavioral Development, 19, 757-772.
- Gottardo, A., Yan, B., Siegel, L. S., & Wade-Woolley, L. (2001). Factors related to English reading performance in children with Chinese as a first language: More

- evidence of cross-language transfer of phonological processing. Journal of Educational Psychology, 93, 530-542.
- Hansen, J. G. (2001). Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. Applied Linguistics, 22, 338-365.
- Hodne, B. (1985). Yet another look at interlanguage phonology - The modification of English syllable structure by native speakers of Polish. Language Learning, 35, 405-422.
- Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. Cognition, 59, 119-147.
- Hoosain, R. (1991). Psycholinguistic implications for linguistic relativity: A case study of Chinese. Hillsdale, NJ: L. Erlbaum Associates.
- Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. Cognition, 54, 73-98.
- Ioup, G. (1984). Is there a structural foreign accent: A comparison of syntactic and phonological errors in 2nd language-acquisition. Language Learning, 34, 1-17.
- Jakobson, R., & Halle, M. (1956). Fundamentals of language. 's-Gravenhage, Mouton.
- Jia, Z., & Zhang, H. (2001, September 26). Zuixin hanzi tongji jiexiao (or Latest statistics on Chinese characters published). Beijing Wanbao (or Beijing Evening News), p. 2.
- Johnston, R. S., Anderson, M., & Holligan, C. (1996). Knowledge of the alphabet and explicit awareness of phonemes in pre-readers: The nature of the relationship. Reading and Writing, 8, 217-234.

- Kabak, B. (2003). The perceptual processing of second language consonant clusters.
Unpublished doctoral dissertation, University of Delaware.
- Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. Journal of Memory and Language, 37, 295-311.
- Kirtley, C., Bryant, P., Maclean, M., & Bradley, L. (1989). Rhyme, rime, and the onset of reading. Journal of Experimental Child Psychology, 48, 224-245.
- Koda, K. (1988). Cognitive process in second language reading: Transfer of L1 reading skills and strategies. Second Language Research, 4, 133-156.
- Koda, K. (1989). Effects of L1-orthographic representation on L2-phonological coding strategies. Journal of Psycholinguistic Research, 18, 201-222.
- Koda, K. (1999). Development of L2 intraword orthographic sensitivity and decoding skills. Modern Language Journal, 83, 51-64.
- Kubozono, H. (1996). Speech segmentation and phonological structure. In T. C. Otake and A. Berlin (Eds.), Phonological structure and language processing (pp. 77-94). New York: Mouton de Gruyter.
- Lado, R. (1957). Linguistics across cultures: Applied linguistics for language teachers. Ann Arbor, University of Michigan Press.
- Levelt, C. C., Schiller, N. O., & Levelt, W. J. (2000). The acquisition of syllable types. Language Acquisition, 8, 237-264.
- Levelt, C. C., Schiller, N. O., & Levelt, W. J. M. (1999). A developmental grammar for syllable structure in the production of child language. Brain and Language, 68, 291-299.

- Li, W.-C. (1999). A diachronically-motivated segmental phonology of Mandarin Chinese. New York: Peter Lang.
- Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. Journal of Educational Psychology, 95, 482-494.
- Lombardi, L. (2003). Second language data and constraints on manner: Explaining substitutions for the English interdental. Second Language Research, 19, 225-250.
- Major, R. C. (1987). A model for interlanguage phonology. In G. Ioup and S. H. Weinberger (Eds.), Interlanguage phonology: The acquisition of a second language sound system (pp. 101-124). Cambridge, MA: Newbury House.
- Major, R. C. (2001). Foreign accent: The ontogeny and phylogeny of second language phonology. Mahwah, NJ: L. Erlbaum.
- Martinot, C., & Gombert, J. E. (1996). Development and control of phonological knowledge in preschoolers. Revue De Neuropsychologie, 6, 251-270.
- McGrew, K. S., & Woodcock, R. W. (2001). Technical manual: Woodcock-Johnson III. Itasca, IL: Riverside Publishing.
- Muljani, D., Koda, K., & Moates, D. R. (1998). The development of word recognition in a second language. Applied Psycholinguistics, 19, 99-113.
- Naglieri, J. A. (1985). Matrix analogies test-expanded form: Examiner's manual. San Antonio, TX: Psychological Corporation.
- Osburne, A. G. (1996). Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. Applied Linguistics, 17, 164-181.

- Prince, A., & Smolensky, P. (2004). Optimality theory: constraint interaction in generative grammar. Malden, MA: Blackwell Pub.
- Read, C., Zhang, Y. F., Nie, H. Y., & Ding, B. Q. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. Cognition, 24, 31-44.
- Sato, C. J. (1984). Phonological processes in 2nd language-acquisition: Another look at interlanguage syllable structure. Language Learning, 34, 43-57.
- Schreuder, R., & van Bon, W. H. J. (1989). Phonemic analysis: Effects of word properties. Journal of Research in Reading, 12, 59-78.
- Sohn, H.-m. (1999). The Korean language. New York, Cambridge University Press.
- Spencer, A. (1996). Phonology: Theory and description. Oxford, England: Blackwell.
- Stahl, S. A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. Journal of Educational Psychology, 86, 221-234.
- Tarone, E. (1987). Some influences on the syllable structure of interlanguage phonology. In G. Ioup and S. H. Weinberger (Eds.), Interlanguage phonology: The acquisition of a second language sound system (pp. 232-247). Cambridge, MA: Newbury House.
- Treiman, R. (1983). Onsets and primes as units of syllables: Evidence from children. Bulletin of the Psychonomic Society, 21, 364-364.
- Treiman, R. (1984). On the status of final consonant clusters in English syllables. Journal of Verbal Learning and Verbal Behavior, 23, 343-356.
- Treiman, R. (1985). Onsets and rimes as units of spoken syllables - Evidence from children. Journal of Experimental Child Psychology, 39, 161-181.

- Treiman, R. (1986). The division between onsets and rimes in English syllables. Journal of Memory and Language, 25, 476-491.
- Treiman, R. (1991). Children's spelling errors on syllable-initial consonant clusters. Journal of Educational Psychology, 83, 346-360.
- Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell. In P. B. Gough and R. Treiman (Eds.), Reading acquisition (pp. 65-106). Hillside, NJ: Lawrence Erlbaum.
- Treiman, R. (1995). Errors in short-term-memory for speech: A developmental study. Journal of Experimental Psychology-Learning Memory and Cognition, 21, 1197-1208.
- Treiman, R. (2000). The foundations of literacy. Current Directions in Psychological Science, 9, 89-92.
- Treiman, R., Bowey, J. A., & Bourassa, D. (2002). Segmentation of spoken words into syllables by English-speaking children as compared to adults. Journal of Experimental Child Psychology, 83, 213-238.
- Treiman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? Developmental Psychology, 33, 771-780.
- Treiman, R., & Danis, C. (1988). Short-term-memory errors for spoken syllables are affected by the linguistic structure of the syllables. Journal of Experimental Psychology-Learning Memory and Cognition, 14, 145-152.
- Treiman, R., & Kessler, B. (1995). In defense of an onset-rime syllable structure for English. Language and Speech, 38, 127-142.

- Treiman, R., & Weatherston, S. (1992). Effects of linguistic structure on children's ability to isolate initial consonants. Journal of Educational Psychology, 84, 174-181.
- Treiman, R., Zukowski, A., & Richmondwelty, E. D. (1995). What happened to the N of sink: Children's spellings of final consonant clusters. Cognition, 55, 1-38.
- Uhry, J. K., & Ehri, L. C. (1999). Ease of segmenting two- and three-phoneme words in kindergarten: Rime cohesion or vowel salience? Journal of Educational Psychology, 91, 594-603.
- Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. Developmental Review, 13, 286-350.
- Wang, M., & Geva, E. (2003). Spelling performance of Chinese children using English as a second language: Lexical and visual-orthographic processes. Applied Psycholinguistics, 24, 1-25.
- Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. Cognition, 87, 129-149.
- Wang, M., Koda, K., & Perfetti, C. A. (2004). Language and writing systems are both important in learning to read: A reply to Yamada. Cognition, 93, 133-137.
- Weinberger, S. (1987). The influence of linguistic context on syllable simplification. In G. Ioup and S. Weinberger (Eds.), Interlanguage phonology: The acquisition of a second language sound system (pp. 401-417). Cambridge, MA: Newbury House.
- Wode, H. (1978). Beginnings of non-school room L2 Phonological Acquisition: Survey of problems and issues based on data from English as L2 with German as L1.

- IRAL-International Review of Applied Linguistics in Language Teaching, 16, 109-125.
- Yamada, J. (2004). An L1-script-transfer-effect fallacy: A rejoinder to Wang et al. (2003). Cognition, 93, 127-132.
- Yavas, M. S., & Core, C. W. (2001). Phonemic awareness of coda consonants and sonority in bilingual children. Clinical Linguistics & Phonetics, 15, 35-39.
- Yavas, M. S., & Gogate, L. J. (1999). Phoneme awareness in children: A function of sonority. Journal of Psycholinguistic Research, 28, 245-260.
- Yoon, H. K., Bolger, D. J., Kwon, O. S., & Perfetti, C. A. (2002). Subsyllabic units in reading. In L. Verhoeven, C. Elbro and P. Reitsma (Eds.), Precursors of functional literacy (pp. 139-164). Philadelphia: John Benjamins Publishing Company.
- Yoon, Y. B., & Derwing, B. L. (2001). A language without a rhyme: Syllable structure experiments in Korean. Canadian Journal of Linguistics-Revue Canadienne De Linguistique, 46, 187-237.
- Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. Reading Research Quarterly, 23, 159-177.
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. Psychological Bulletin, 131, 3-29.