University of Alberta

Experimental Studies of the Syllable and the Segment in Korean

by

Yeo Bom Yoon ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Psycholinguistics

Department of Linguistics

Edmonton, Alberta

Fall 1995

ISBN 0-612-06313-5

University of Alberta

Library Release Form

Name of Author: Yeo Bom Yoon

Title of Thesis: Experimental Studies of the Syllable and the Segment in Korean

Degree: Doctor of Philosophy

Year this Degree Granted: 1995

Yeo Bom Yoon

Yeo Bom Yoon
3, Keumam-Dong, Kunsan-Shi
Jeonlabuk-Do, KOREA (SOUTH)

Date: 15 August 1995

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *Experimental Studies of the Syllable and the Segment in Korean* submitted by **Yeo Bom Yoon** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Psycholinguistics**.

Dr. Bruce L. Derwing (Supervisor)

Dr. Eung-Do Cook

Dr. Terrance M. Nearey

Dr. Bernard L. Rochet

Dr. Gary H. Libben

Dr. Rebecca Treiman (External Examiner)

Date: 4 August 1995

부모님께

To my parents

Abstract

Syllable structure and the basic level of phonological representation in Korean were explored using sound similarity judgments and concept formation techniques. For syllable structure, experiments were designed to test whether a CVC syllable in Korean is divided into CV and C or into C and VC. For the basic level of phonological representation, the experimental question was which is a more basic phonological unit in Korean: the segment or the syllable?

In the sound similarity judgment experiments, subjects listened to pairs of words and judged on a scale how similar each pair was in sound. In Experiments 1 and 2, CVC-CVC pairs with the same CV were judged more similar in sound than those with the same VC, which was in contrast with the previous research on English where the opposite results were found. These results suggested that CV is a more salient subcomponent of a CVC syllable than VC, thus favoring the left-branching (CV-C) over the right-branching (C-VC) analysis for Korean syllable structure. Using disyllabic CV.CVC-CV.CVC pairs, Experiment 3 compared predictions of sound similarity judgments based on a phoneme count vs. a syllable count. It was found that the syllable was a more important predictor than the phoneme, contrasting again with the previous research on English.

Concept formation experiments were designed to test the ease of learning target word sets defined in terms of a common CV vs. VC (Experiment 4) and a common C vs. CVC (Experiment 5). The results of Experiment 4 suggested that the CV-based concept was much easier to learn than VC-based concept, conforming to the results of Experiments 1 and 2. In Experiment 5, subjects learning the syllable concept found the task much easier than those learning the phoneme concept.

Taken together, the results of the present experiments clearly support a left-branching syllable structure and the primacy of the syllable in Korean, and thereby shed

doubt on the hypothesis of a universal right-branching syllable structure and on the role of the segment as a universal basic phonological unit.

Acknowledgments

My deepest appreciation goes to Professor Bruce Derwing, my thesis supervisor. When we first met in Seoul in 1990, he opened the door to experimental phonology for me. Since then he has guided me with grace, caution and insight. Whenever I flip through the numerous drafts of my thesis filled with his insightful comments, I realize that I have had the best possible supervision. He has also constantly encouraged me to participate in many conferences and to publish my work. Although I have now finished my PhD, it is still hard to imagine myself being an independent scholar without his help.

My thanks also go to those who served as my committee members: Professor Terry Nearey, Professor Gary Libben and Professor Bernard Rochet of the University of Alberta, Professor Eung-Do Cook of the University of Calgary, and Professor Rebecca Treiman, my external examiner, from Wayne State University. Their comments and questions were very helpful to me in broadening my views. In particular, Professor Treiman suggested many intriguing new ways to analyze my data. Special acknowledgment is given to Professor John Ohala for his visiting lectures, from which I obtained many interesting ideas on experimental phonology.

I have been lucky to have great colleagues in the Department: Greg Skrypiczajko, Yuhuan Wang, Grace Wiebe and many others. Having a chat with them was always amusing, and I think I have benefited from them more than they did from me. I also thank Mr. Hwakyu Kim, Mr. Chiwoo Park, and their families for making me less lonely.

In Korea, there are also many to whom I owe special thanks. First of all, I gratefully acknowledge Professor Tae-ok Kim, who has been a constant source of encouragement for me throughout my life at Sogang University as well as in Edmonton. I owe a lot to Professor Sook Whan Cho, who encouraged me to continue my academic career at the University of Alberta. Thanks also go to Professor Kook Chung for his phonology classes and supervision of my MA thesis and to professors at Sogang University for their interesting lectures. I also thank all students who served as willing subjects for my experiments.

Throughout my program, I have received financial assistance and travel grants. Thanks are due to the University of Alberta FGSR and to the organizers of ICSLP '92 .

Neither last nor least, my heartfelt thanks go to my wife, Eun Sook Oh. She had to stay at home most of the time looking after Hyoung Sik (Joshua), while I was working at school. I am sure she was happier than I was, when the thesis was finally over.

Finally, this thesis is dedicated to my parents. Without their financial and emotional support, I could not have endured living far away from home, not to speak of studying.

TABLE OF CONTENTS

KOREAN PHONOLOGICAL UNITS AND EXPERIMENTATION

SOUND SIMILARITY JUDGMENTS

CONCEPT FORMATION

SUMMARY AND CONCLUSION

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

KOREAN PHONOLOGICAL UNITS AND EXPERIMENTATION

1.1. Overview

There can be little doubt that the phonemic segment and the syllable both play major roles in the phonology of Korean, as in other languages. This is reflected in the unique writing system of Korean, in which not only individual phoneme-sized letters are written but these letters are combined to form syllable-like orthographic "packages" as well. In fact, a great deal of Korean phonological theory assumes implicitly or explicitly the existence of both phonemes and syllables. Furthermore, subsyllabic units intermediate in size between the phoneme and the syllable have also been hypothesized, such as the rime (VC unit) and the body (CV unit).

Since the 1980's, a variety of experimental techniques have been applied to test the psychological reality of these units in English (the phoneme and the rime, in particular). The increase in experimentation has been based on the recognition that "if linguistic theories are to be considered psychologically real, then even the most basic or self-evident of their claims needs to be supported with empirical evidence" (Jaeger, 1980b: 233). In the same vein, Derwing and Nearey (1986: 188) argued that "the task of psycholinguistics is not to take the findings of purely descriptive linguistics for granted... but rather to ascertain whether the proposed entities have any psychological validity in the first place."[1] In the Korean phonological literature, however, there have been very few *experimental* studies on the psychological reality of these *basic* units.[2]

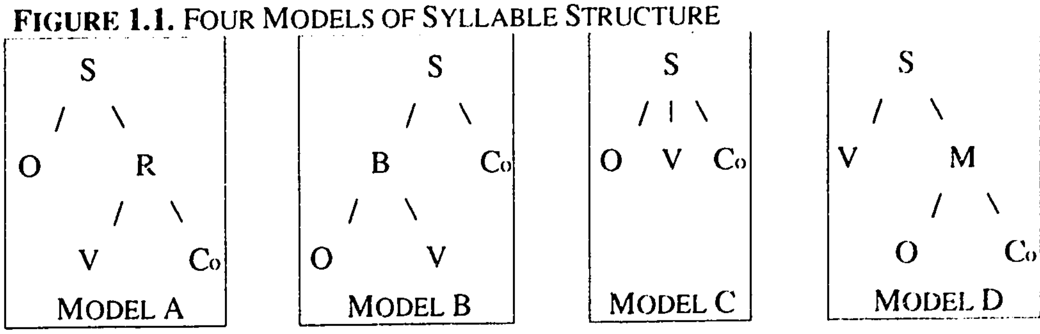
The purpose of the present work is to assess the psychological status of these units in Korean by applying two widely disparate experimental techniques: global sound similarity judgments (SSJ) and concept formation (CF). Both the extension to Korean and the use of diverse techniques were purposeful: Cross-methodological validation is important to ascertain that experimental results do not vary as a function of a particular experimental design, and cross-linguistic investigation is always necessary whenever issues of universality are at stake. In this latter connection, therefore, the present experiments were designed in such a way that the Korean and English results could be compared as objectively and directly as possible.

Since there can be numerous ways to go about testing the psychological status of such units, the present study focuses on the following two issues: syllable structure and the basic level of phonological units in Korean. First, does the Korean syllable have an internal structure that is left-branching [[CV]C] or right-branching [C[VC]]? Second, what is the most basic representational unit in Korean (i.e., the *size* or level of unit that Korean speakers most readily manipulate in a variety of experimental tasks)? The phoneme hypothesis suggests that a CVC syllable can be best analyzed as C-V-C, and that each phoneme plays a key role of a basic discrete unit. On the other hand, the syllable hypothesis states that the syllable as a whole plays a more salient role than do the individual segments that together comprise it.

1.2. Syllable Structure

A variety of theoretical models have been proposed in the literature on syllable structure. Figure 1.1 below represents the four of these hypothesized models.

FIGURE 1.1. FOUR MODELS OF SYLLABLE STRUCTURE



S = syllable, O = onset, V = vowel, Co = coda, R = rime, M = margins, B = body[3]

In English, it has long been assumed that Model A is correct; in fact, claims of universality have ever been made. One notable proponent of this "universal rime hypothesis" (URH) is Fudge (1969, 1987), who simply dismisses Model B, by stating that "No scholar to [his] knowledge has ever espoused [it]." (1987: 360). Instead, arguing primarily against Model C (e.g., Clements and Keyser, 1983), Fudge provided several pieces of naturalistic evidence in support of Model A, including distributional constraints (e.g., Selkirk, 1982), the weight of the rime, speech errors (e.g., Fromkin, 1971), and word games. Briefly, Fudge's (1987) point was that all these spontaneous effects could only be satisfactorily explained by positing a rime unit that is intermediate in size between the individual phonemes and the whole syllable.

In fact, as far as English is concerned, considerable evidence has also been garnered in favor of Model A, using a variety of different experimental techniques, among them: (i) experimental word games (most notably, word-blending, as in Treiman, 1983, 1984, 1986); (ii) unit counting (Dow, 1987); (iii) substitution-by-analogy (Derwing, Nearey, and Dow, 1986); (iv) short-term memory recall (Treiman and Danis, 1988); and others. To illustrate with a word-blending example, for instance, Treiman (1983) asked subjects to take a pair of meaningless but well-formed English monosyllables (e.g., /krɪnt/ and /glʌpɵ/) and to combine them into a new monosyllabic word that contained parts of both. Her subjects produced /krʌpɵ/ (onset of the first syllable plus rime of the other) far more frequently than any other possible combination. All of the other experimental studies cited also showed the prominence of the rime within the English syllable, a result that was compatible with Model A above (i.e., the view that the English syllable is not only hierarchical in structure but also right-branching).

As for Korean, however, this issue is more controversial. Model A has been most widely accepted in the theoretical literature on the subject (e.g., Kang, 1991; Kim, 1986; Kim-Renaud, 1978; Lee, 1993; Sohn, 1987). There has also been some advocacy of

Model C on the grounds of lack of strong evidence for either Model A or Model B (e.g., Kim, 1984a). Arguments for Model A in Korean in general boil down either to a rather non-critical acceptance of the URH or to denying the appropriateness or adequacy of the behavioral evidence used by the proponents of Model B (e.g., Ahn, 1988; Cheon, 1980; Gim, 1987; Kim, 1990). We will now briefly discuss some of this evidence.

First, on the basis of some speech error data in Korean, Cheon (1980) argued that there was much evidence to suggest that the Korean syllable has a CV-C structure, unlike the C-VC structure that Hockett's (1967) data indicated for English. Some of Cheon's examples were as follows ("." indicates a syllable boundary):

cil.sə.ka.mun.lan... → mul.sə.ka.cin.lan... 'not in order' (p. 23)

ca.sip.sə → cap.sə 'reference book' (p.20)

yal.mip.saŋ... → yap.saŋ 'somewhat hateful' (p. 20)

As the underlined segments above indicate, the first example above shows that CV sequences of CVC syllables were transposed, and the last two examples illustrate the omission of CV sequences (together with a coda in the last case). Cheon (1980) maintained that it was in general quite natural in Korean for a syllable-initial CV to be transposed or omitted as a unit.

There is other behavioral evidence for the body-coda structure in Korean, as well. Gim (1987), for example, cited language games, in which "nosa" or "pV" (where V is a copy of the vowel from the preceding syllable) is inserted between the CV and the C of a CVC syllable:

c^həl.su 'person's name' → c^hə.no.sal.su (p.51)

kwaŋ.ho 'person's name' → kwa.paŋ.ho (p.53)

Compare these examples with English word game "Pig Latin," in which *pig*, for example, changes to *igpay*. While the Korean word games require the syllable /c^həl/ to be divided into /c^hə/ and /l/ (i.e., CV and C), the English case shows the division of the syllable /pɪg/ into /p/ and /ɪg/ (i.e., C and VC). This contrast seems to support the idea that the two languages have different syllable structures. However, Lee (1993) argued that these Korean word games had nothing to do with syllable structure; rather, since the to-be-inserted syllable started with a consonant, it could only be inserted after the vowel, since insertion after the initial C would result in unpronounceable sequences with CC clusters, such as *c^hno.sa.əl.su (p. 40).

As another piece of evidence for Model B, Gim (1987) pointed out that children learn a syllable such as "손" <son>[4] as a combination of "소" <so> + "ㄴ" <n> but not by "ㅅ" <s> + "온" <on> (p. 47).[5] However, Kang (1991) argued that, irrespective of syllable structure, the combination of <so> + <n> is easier to learn than the combination of <s> + <on>, because the former has easy-to-recognize /so/ as an initial sequence but the latter has /s/, which is unpronounceable in isolation. (That the same argument could be applied to a final consonant seems to be beside the point.) In sum, while word games and orthography are used as evidence for a left-branching syllable structure by the proponents of Model B, they are considered as "nonarguments" for syllable structure in Korean by the proponents of Model A.

There are also some word formation processes in Korean that are used as evidence for both models:

Model A: /a.ki/ ≈ /ca.ki/ 'charming' (Kang, 1991: 45)

/al.t'il/ ≈ /sal.t'il/ 'frugally' (Kang, 1991: 46)

Model B: /a.sak/ → /a.sa.sak/ 'munching' (Ahn, 1988: 350)

/solJ/ + /namu/ → /so.na.mu/ 'pine tree' (Ahn, 1988: 351)

In the examples for Model A, a null onset alternates with another onset /c/ or /s/, leaving the rime intact. In the examples for Model B, on the other hand, either a CV sequence is reduplicated or the coda is deleted, leaving the body intact. As we discussed in the language game and orthographic evidence above, however, the same kinds of dismissal arguments could also be used for such cases of language-internal evidence.[6]

The first major experimental approach to this issue on the Korean syllable was taken by Derwing, Yoon, and Cho (1993), who used a forced-choice version of Treiman's word-blending task (see also Wiebe and Derwing, 1994). Subjects were asked to choose the better or more natural blending form out of two input syllables. Sample test items were as follows, showing how both 'left-to-right' and 'top-to-bottom' ordering were controlled in that study:

1. (a) səm + tʰoŋ → soŋ
   (b) səm + tʰoŋ → səŋ

2. (a) tʰoŋ + səm → tʰom
   (b) tʰoŋ + səm → tʰəm

Overall, subjects preferred body+coda blends (i.e., 1b and 2a) over onset+rime blends (i.e., 1a and 2b), regardless of the order in which the two input words were presented. Later, Yoon (1994a) confirmed that there was again significant preference for body+coda blends. These results supported Model B, which, as already noted, was also supported by such behavioral evidence as speech errors (Cheon, 1980) and word games (Gim, 1987). Taken together, experimental and naturalistic evidence sheds serious doubt on English-like structure for the Korean syllable. Derwing, Yoon, and Cho (1993: 227) reached the following tentative conclusion:

Korean vowels are indeed more closely linked to preceding consonants than to succeeding ones, a result that is consistent with a left-branching model of the syllable. This is precisely the opposite result to that found for an onset-rime language like English and constitutes a direct challenge to the notion of a universal right-branching approach to syllable structure.

In the present work, two further experimental techniques, sound similarity judgments (SSJ) and concept formation (CF), are applied to the investigation of syllable structure in Korean. Briefly, the SSJ task will test whether syllable structure is reflected in subjects' judgments of overall sound similarity of pairs of Korean words, and the CF task will compare the ease with which subjects learn a concept based on a common CV *vis-a-vis* a concept based on a common VC. The main purpose of this application was to firm up our previous findings (Derwing, Yoon, and Cho, 1993; Yoon, 1994a; Yoon and Derwing, 1994) and to provide a strong cross-methodological validation for them, using less direct kinds of experimental tasks.

1.3. Basic Level of Phonological Representation

The issue of a universal basic representational unit for language has not been completely resolved, as both the segment (phoneme) and the syllable continue to compete for the ultimate prize. Although anecdotal, the classical work of Sapir (1925) provided some early quasi-experimental evidence for the psychological reality of phonemes, as his informants seemed to ignore allophonic differences and to make phoneme-like judgments. The development of alphabetic writing system has also been taken as support for the phoneme, although there is an intractable problem of circularity involved (i.e., whether the knowledge of alphabets precedes phonemic awareness or vice versa; see Dow, 1987); moreover, not all languages have adopted the alphabetic model. There is also considerable evidence in favor of the syllable. In language acquisition, for example, young children seem to manipulate whole syllables much better than smaller units, including the phoneme, irrespective of their native language background. Adults unfamiliar with alphabetic writing systems seem to do much the same (e.g., Morais, Bertelson, Cary, and Alegria, 1986; Morais, Cary, Alegria, and Bertelson, 1979; Read, Zhang, Nie, and Ding, 1986). There is also evidence to suggest that speech is primarily encoded in terms of the syllable (e.g., Fromkin, 1971, argues that phonemes are transposed on the basis of syllable structure).

Psycholinguistic experiments have also provided evidence for both the phoneme and the syllable. Most notably, the detection paradigm in the 1970's provided mixed results for English. The phoneme, which was considered "nonperceptual" in Savin and Bever (1970), was emphasized as psychologically more basic than the syllable by Foss and Swinney (1973). Later, McNeill and Lindig (1973) and Healy and Cutting (1976) suggested that both the phoneme and the syllable are psychologically real phonological units in English, and that their relative prominence can vary as a function of experimental design.

Instead of comparing the two units, Vitz and Winkler (1973) focused on testing the role of phonemes in predicting the judged sound similarity of English word pairs. In spite of its rather simple algorithm based only on a phoneme count, their model was able to predict a substantial portion of the variance in similarity judgments (see section 2.2.2 for details), providing strong evidence for the phoneme as a basic perceptual unit in English. Other experimental studies followed in the 1980's, supporting each of the three fundamental characteristics of the phoneme that were summarized by Derwing, Nearey, and Dow (1986): (i) as a class of allophones; (ii) as a smallest nondiscriminable perceptual unit; and (iii) as a discrete segment. The first characteristic was confirmed by Jaeger (1980ab, 1986a). Using the concept formation technique, she showed that English subjects spontaneously categorized various allophones of /k/ into one phoneme category. Derwing and Nearey (1981) also found that allophones of /t/ are perceived as equivalent to one another, supporting the second characteristic. Finally, using the unit counting task, Dow (1987) showed that English subjects had no trouble segmenting the words into units that were largely consistent with a traditional phonemic analysis. In sum, the experimental results in the 1980's, taken together with Vitz and Winkler's (1973) results, all seemed to confirm that the phoneme was a basic unit in English.

However, as often discussed in the experimental phonological literature (e.g., Derwing, 1992b; Derwing and Dow, 1987; Mann, 1986b; Otake, Hatano, Cutler, and Mehler, 1993; Read, Zhang, Nie, Ding, 1986), most of these results have been strongly colored by orthography. For example, after comparing the phonemic categorization behavior of English and Japanese speakers, Jaeger (1980a) noted a relationship between the basic units recognized and the type of orthography employed:

> For English speakers, the phoneme-sized unit is the basic level at which they [speakers] interact with the sounds of their language; syllables, words, etc. are superordinate levels and features are a subordinate level. For Japanese speakers, the syllable is a basic level of categorization of the sounds of their language; words, etc. are superordinate, and phonemes, features, etc. are subordinate. (p. 146)

Derwing, Nearey, Beinert, and Bendrien (1992) also pointed out the potential role of orthography in testing the universal basic level of phonological representation:

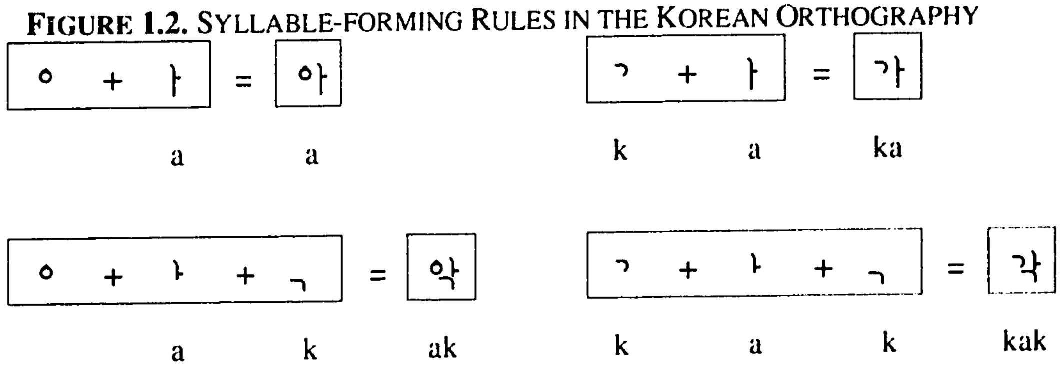
> There is at least some circumstantial evidence that casts suspicion on the idea that the segment is the natural or universal basic representational unit.

> For literate speakers, in fact, there appears to be a *tie between the type of orthographic systems used and the size of the representational units recognized*, and while segment-based alphabetic writing systems are now widespread, large residues persist of historical precursor systems that employ larger units, such as words (logographs), syllables, morae and others. (p. 289) [emphasis added]

In testing the viability of the candidate units, what is necessary then is not only cross-methodological validation but also cross-linguistic comparison, avoiding as much as possible any contaminating orthographic effects.

Korean provides an interesting testing ground to this end, since its orthography employs at once syllable-like units and phoneme-sized letters. The Korean orthographic system *hangul* (pronounced as /han.kɨl/) is at once alphabetic like Indo-European languages in that all phoneme-sized units are individually written, and syllabic in that individual phoneme-sized letters are grouped to form a syllable-like orthographic unit, as well:

FIGURE 1.2. SYLLABLE-FORMING RULES IN THE KOREAN ORTHOGRAPHY



A unique characteristic of the Korean orthography indicated in Figure 1.2 above is that all syllables are written in an equisize square, regardless of the difference in the number of phonemes. Especially notable is the fact that, for syllables that begin with a vowel, the silent letter 'o' is added to maintain the consistent size and the shape of the syllable as a whole.

Few studies in the Korean phonological literature deal with such basic issues as the status of the phoneme *vis-a-vis* the syllable. Based on a traditional phonemic analysis, Martin (1951) posited a phonemic inventory for Korean. Under the influence of generative phonology (e.g., Chomsky and Halle, 1968), most later studies on the Korean phoneme focused on positing underlying forms and rules for deriving surface forms (e.g., Kim-Renaud, 1986; Lee, 1976). As far as our knowledge goes, there has been no study on the issue of basic level of representation in Korean similar to the English studies sketched above. As a result, there has been no experimental evidence either for or against the psychological reality of the phoneme or the syllable in Korean. However, there is some behavioral evidence that the syllable in Korean may play a role that is at least as important as the phoneme is in English.

The first evidence is related to the fact that the number of syllables is a salient factor in versification. The traditional poetry *sico* employs the syllable count as a metric unit. The *sico* consists of three lines, and each line contains four phrases, as presented below.[7]

흥 망 이   유 수 하 니   만 월 대 도   추 초 로 다

hiŋ maŋ i   yu su ha ni   man wəl tæ to c^hu c^ho lo ta

오 백 년   왕 업 이   목 적 에   부 쳤 으 니

o pæk nyən   waŋ əp i   mok cək e   pu c^hyəs ɨ ni

석 양 에   지 나 는 객 이   눈 물 겨 워   하 노 라

sək yaŋ e   ci na nin kæk i   nun mul kyə wə   ha no la

From the number of syllables used in each phrase, at least three basic patterns can be identified: (i) The first phrase in each line always consists of three syllables; (ii) most phrases contain three to four syllables; and (iii) the second phrase in the third line is longer than any other phrase. In sum, the syllable count is an important metric device, which is similar to the way the Japanese mora is used as a counting unit in *haiku*, the Japanese traditional poetic form (Otake, Hatano, Cutler, and Mehler, 1993). Furthermore, as shown in the syllable-forming rule in Figure 1.2 above, the syllable boundary is so clear-cut that the number of syllables is often used as a good cue to recall of the words.

There is also a language game that shows that the syllable is a readily identifiable unit to Korean speakers. In this game, called *maliski* ('connecting words'), players take turns producing a new word on the basis of a part of the previous word; specifically, the first syllable of a new word should be the last syllable of the prior word (e.g., *hak.kyo* → *kyo.sil* → *sil.su* →...). This continues until a player is unable to construct a new word. Thus, a common winning strategy is to produce a new word whose last syllable is relatively rare as a first syllable. Aso, crossword puzzles in Korean are based on the syllable, unlike those in English that are based on individual letters. In sum, the number of syllables and the identity of a particular syllable serve as evidence for the syllable as a highly salient phonological unit in Korean.

In the present thesis, the status of the phoneme and of the syllable are compared by applying the sound similarity judgment (SSJ) and concept formation (CF) tasks. These techniques were chosen for the following two reasons: (i) to permit cross-linguistic comparison of the status of these units and (ii) to provide a direct comparison between the phoneme and the syllable within Korean. First, since both techniques have already been employed in the investigation of phonological units in English, the results of their application to Korean ought to go a long way in assessing the status of the same kinds of units in that language. Second, since both techniques are flexible enough to allow for a direct comparison among a variety of phonological units (features, phonemes, and syllables, as well as intermediate units), if the results of the SSJ studies can be confirmed by means of a CF task, a firm empirical base will have been laid in support of the conclusions to be drawn.

1.4. Organization of the Thesis

Chapter II discusses the application of the SSJ task, and Chapter III the application of the CF task. Both chapters are organized in a parallel fashion. The first section overviews the basics of the task. The second section reviews the previous methodologies and the results of the earlier experiments, primarily on English. Each of the remaining sections then describes one of the present experiments in detail, including the results. The last section of each chapter summarizes the results of all experiments and provides a general assessment of the experimental task. Chapter IV, finally, presents the overall findings from all of the studies and relates them to the two issues of syllable structure and basic phonological units, as well as suggesting directions for future research.

Notes to Chapter I

[1] See Derwing (1973, 1980), Eddington (1993), McCawley (1986), Ohala (1974, 1986), Ohala and Jaeger (1986a), and Prideaux, Derwing, and Baker (1980) for a further discussion of these issues.

[2] See Chung (1994a), however, for some specific experimental phonological questions in Korean. He also explains the circularity of theoretical arguments and the necessity of experimental approaches to Korean phonology.

[3] The term *body* has been taken from Vennemann (1988a) and will be used throughout the present work. Alternative labels that have been used for this unit are *core*, *head*, or simply X.

[4] Throughout the thesis, angled brackets < > are used to denote Romainzed Korean letters.

[5] This orthographic convention is somewhat akin to the Japanese *hiragana* and *katakana* writing systems, in which the same syllable /son/ can be written as /so/ (そ) + /n/ (ん), whereas there is no combination such as /s/ + /on/.

[6] See Ohala (1986).

[7] Kim (1986) translated this poem into English as follows:

The rise and fall of things follow nature's way; autumn grass yellows on
Full-moon Terrace.
Five centuries of royal reign are drowned by the notes of a herdboy's flute;
The traveler at dusk can hardly restrain his tears.

CHAPTER II

SOUND SIMILARITY JUDGMENTS

2.1. Introduction

2.1.1. Overview of the Sound Similarity Judgment Task

Greenberg and Jenkins (1964) introduced the "sound similarity judgment" (SSJ, hereafter) task as a tool for investigating phonological units. In SSJ experiments subjects listen to pairs of words and rate on scale how similar each pair is in sound. Most of the SSJ studies in English focused on monosyllabic CVC-CVC pairs and varied systematically the number and the position of the shared phonemes (e.g., Nelson and Nelson, 1970). These early studies suggested that there was a strong correlation between the number of shared phonemes and similarity scores provided by subjects.

Vitz and Winkler (1973) made explicit the importance of phonemic segments in predicting sound similarity in English. Based on the number of shared phonemes only, their model "predicted phonemic distance" (PPD, hereafter) could account for most of the variation in similarity scores, as the correlations between PPD and similarity scores were about -.9 averaged across four experiments ranging from monosyllabic to trisyllabic words (see section 2.2.2 for details). These results were confirmed by Derwing (1976), Derwing and Nearey (1986), and Bendrien (1992). Although the Vitz and Winkler study, as argued by Derwing, Nearey, and Dow (1986), was not designed to test the phoneme concept *per se*, the results provided strong "indirect" evidence for the phoneme as a basic representational unit in English.[1]

Other factors, however, are ignored in the PPD model, such as the role of sub-phonemic differences (i.e., features). Derwing and Nearey (1986) found that two phonemes sharing all but one feature (e.g., /b-p/) were judged significantly more similar than those differing by two or more features (e.g., /b-s/), suggesting that feature differences should be taken into account in a complete model. Furthermore, the position of shared phonemes was another important factor in predicting similarity scores. For English CVC words, Nelson and Nelson (1970) found that among pairs sharing two phonemes, pairs with the last two phonemes (VC) in common were rated significantly more similar than pairs with the other two phonemes in common (CV or C...C). Similarly, Derwing and Nearey (1986) and Bendrien (1992) found that final consonants were more important than initial consonants in English SSJs. In sum, although (i) the phoneme effect (i.e., simple counting of shared phonemes between two words) was most important in predicting the sound similarity of English words, (ii) the feature effect (i.e., the degree of similarity between phonemes) and (iii) the position effect (i.e., the position of shared phonemes) should also be taken into account to achieve better prediction of SSJs.

On the basis of these earlier findings in English, the SSJ task has been applied to a variety of languages, including Arabic (Beinert and Derwing, 1993), Japanese (Derwing and Wiebe, 1994), Korean (Yoon and Derwing, 1994), and Taiwanese (Wang and Derwing, 1993). The main goal of these cross-linguistic studies was to test basic representational units across language types and the relationship between these representational units and the orthography. The results showed that, while the phoneme was in general the most important variable, other language-specific phonological units also had a significant effect (see Derwing and Nearey, 1994, for details). Most notably, in Japanese, the similarity predictions proved to be based more on the mora than the phoneme. In Korean, too, the body as opposed to the rime played an important role as an additive factor on achieving more precise prediction of similarity scores (cf. Derwing, Yoon, and Cho, 1993; Yoon, 1994a). The C-tier, a hypothesized phonological unit in Arabic, also emerged as a significant variable (but not V-tier). Overall, the SSJ task proved to be useful for investigating the status of a variety of phonological units across languages.

2.1.2. Predicting the Sound Similarity Judgments in Korean

As discussed above, the phoneme proved to be the most basic unit in predicting SSJs in English. In Japanese, however, predictions of SSJs were more accurate on the basis of the mora than on the basis of the phoneme. The results of the concept formation task applied to the two languages also pointed in the same direction (Jaeger 1980a), suggesting that there *is* a difference in the degree to which the phoneme plays a role in the two languages. As often discussed in the literature, however, orthography has also been heavily implicated in all of these studies (Derwing and Dow, 1987; Jaeger 1980a; Mann 1986b; Read, Zhang, Nie, and Ding 1986).

In this regard, Korean provides an interesting testing ground, as already noted in Chapter I, since both phoneme-and syllable-sized orthographic units are used, thus raising the possibility of comparing these units in a way that is independent of orthographic effects. If the phoneme is the most basic unit in predicting SSJs in Korean as in English, predictions based on phoneme counting should be more accurate than those based on syllable counting. If, on the other hand, the syllable is the more salient unit than the phoneme in Korean, an opposite result ought to be found.

The Korean syllable is often viewed as a left-branching [[CV]C] structure, as opposed to the right-branching [C[VC]] structure in English. Although the issue is still a matter of some controversy in Korean, the "left-branching" view is supported by behavioral evidence (e.g., Ahn, 1988; Cheon, 1980; Gim, 1987) and the "right-branching" view mostly by purely theoretical arguments (Kim, 1986; Kang, 1991; Lee, 1993); our prior experimental evidence also favored the former interpretation (Derwing, Yoon, and Cho, 1993; Yoon, 1994a; Yoon and Derwing, 1994). Assuming that syllable structure is invoked in making SSJ judgments, and if the "left-branching" view of the Korean syllable is correct, quite different patterns of results should emerge for Korean and English on the SSJ task. On the other hand, if the two languages have the same "right-branching" syllable structure, the rime ought to play roughly the same role in the two languages.

2.1.3. Organization of the Chapter and Goals of the Experiments

In the next section, 2.2, the literature on SSJ experiments in English is reviewed. This literature review introduces the basic methodology and discusses in some detail the major effects in predicting the SSJs in English. In sections 2.3 to 2.5, three experiments are discussed. Experiment 1 (section 2.3) tests the relative weights of the six subcomponents of a CVC syllable: the onset, vowel, coda, body (CV), rime (VC), and margins (C...C). The goals are to test (i) whether there are position effects (body vs. rime, in particular) on judged similarity, and (ii) to assess the status of the phoneme as a basic phonological unit in Korean.

Using monosyllabic CVC-CVC pairs again, Experiment 2 (section 2.4) focuses on the comparison between body-sharing pairs and rime-sharing pairs. By varying the degree of similarity in mismatched phonemes, i.e., the coda for body-sharing pairs and the onset for the rime-sharing pairs, Experiment 2 measures the relative strength of the position effect and the feature effect, to check whether the former varies as a function of the latter.

Experiment 3 (section 2.5) compares the status of the phoneme and the syllable as the basic phonological unit in Korean. Using disyllabic CV.CVC-CV.CVC pairs (where "." indicates a syllable boundary), predictions of similarity based on the phoneme and the syllable are compared.

Finally, section 2.6 discusses the results of Experiments 1-3 in the light of syllable structure and the basic level of phonological representation in Korean.

2.2. Literature Review

2.2.1. Nelson and Nelson (1970)

Nelson and Nelson (1970) were interested in the effects of the *number* and the *position* of shared "letters" in CVC pairs. It was hypothesized that verbal stimuli were coded by pronunciation, and that there should be some position effects on one-letter and two-letter overlap conditions (e.g., the first letter might be more important than the second and the third letter).

There were eight types of pairs consisting of 12 pairs each. Although stimulus pairs were presented visually (e.g., *pan-fit*), subjects were asked to pronounce the stimuli to themselves and to rate the similarity of sound on a seven-point scale ranging from 1 (completely different) to 7 (exactly the same). Half of the total of 96 subjects rated the pairs in AB order (*pan-fit*), and the other half in BA order. The results were as follows:

TABLE 2.1. RESULTS OF NELSON AND NELSON (1970)

Overlapped Letter	Example (n = 12)	Score (range: 1-7)	Mean by no. of Letters Matched
None	*pan-fit*	1.74	Zero = 1.74
Last	*pan-fin*	2.60	
Middle	*pan-fat*	2.69	One = 2.87
First	*pan-pit*	3.33	
First + Middle	*pan-pat*	3.99	
First + Last	*pan-pin*	4.17	Two = 4.50
Middle + Last	*pan-fan*	5.33	
All	*pan-pan*	6.98	Three = 6.98

An ANOVA showed that the eight overlap conditions were highly significant as a within-subjects variable ($F[7,658] = 743.8$, $p < .001$), but presentation order as a between-subjects variable was not. As can be seen in Table 2.1, the number of shared letters was a highly significant factor on judged similarity; Fisher's least significant difference was .17. However, there were also some position effects which Nelson and Nelson (1970) failed to explore deeply. When the number of shared letters was controlled, for example, pairs sharing the first letter were judged most similar of all one-letter sharing pairs, and pairs sharing the last two letters were judged most similar of all two-letter sharing pairs. In particular, the large discrepancy between pairs sharing the last two letters (mean similarity score of 5.33) and pairs sharing other two letters (3.99 and 4.17) seemed to constitute evidence for a rime effect, which suggested that subjects were not simply counting the number of shared letters.

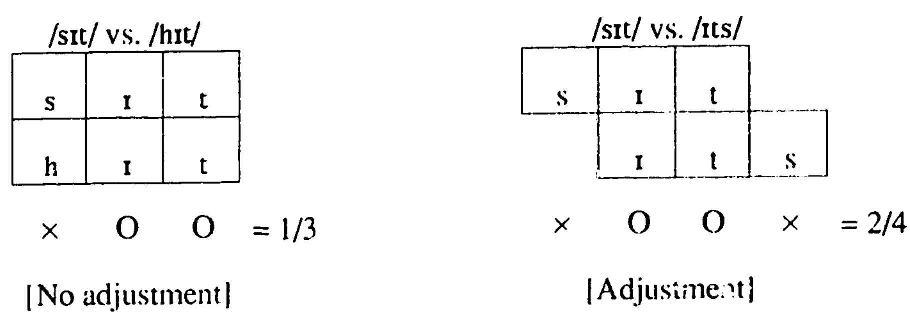
Although stimuli in Nelson and Nelson's (1970) experiments were presented visually, possibly increasing the likelihood of orthographic influence, it was observed that the majority of variations between pairs could be accounted for by articulatory and/or acoustic factors. At least when the similarity attribute was specified as *sound*, subjects seemed to be sensitive to such subsyllabic units as an onset and a rime.

There were, however, some problems involved in generalizing the results. Since the stimulus pairs were derived from the three sets of minimal pairs tightly controlled to allow all possible pairwise comparisons within each set, the kinds of contrasts in each position were substantially limited to a tiny subset of all possible contrasts that are possible with English CVC words. For example, in spite of the large number of pairs ($n = 96$), there were only three kinds of contrasts: *p-f*, *b-l*, *h-r* in the onset, *a-i*, *a-e*, *a-u* in the vowel, and *n-t*, *d-g*, *m-t* in the coda. Furthermore, the degree of similarity between mismatched phonemes (e.g., /p-f/ in the onset vs. /n-t/ in the coda) was not controlled. It is, therefore, difficult to make general claims about sound similarity in English words from Nelson and Nelson (1970).

2.2.2. Vitz and Winkler (1973)

Vitz and Winkler (1973) ran five experiments to demonstrate the usefulness of their Predicted Phonemic Distance (PPD) model, which was built to predict the judged similarity of sound of two English words. The PPD is calculated by a simple procedure. First, two strings of phonemes are aligned to maximize the number of matched phonemes, and PPD is then calculated by the proportion of mismatched phonemes relative to the number of phoneme slots that are needed for the alignment. Thus PPD can range from 0 (for identical pairs) to 1 (pairs sharing no phonemes). For example, the PPD between *sit* and *hit* is 1/3, since only the first consonants are mismatched out of three phonemic slots, while the PPD between *sit* and *its* is 2/4, since two phonemes are mismatched out of four phoneme slots after alignment, as illustrated below:

FIGURE 2.1. ILLUSTRATION OF CALCULATING PPD (O = match, × = mismatch)



Defined in this way, the PPD emphasizes the independent status of an individual phoneme and ignores the position effects discussed in Nelson and Nelson (1970) and the degree of difference between phonemes. One of the most important assumptions behind this model was that all phonemes are *equally* important (or equidistant from one another, using a distance metric).

The first four experiments in Vitz and Winkler (1973) were identical in the following respects: (i) Stimuli were presented aurally and subjects were instructed to focus on similarity in sound, not on meaning or spelling; (ii) there was one standard word which occurred in every pair; (iii) there were 25 comparison words which ranged from no phoneme in common (PPD = 1) to all but one phoneme in common (PPD = close to 0) with the standard word; (iv) there were two trials (the first trial in the standard-first order, and the second with the order reversed); and (v) a five-point scale was used from 0 (zero similarity) to 4 (identity). The experiments differed only with regards to stimuli. In Experiments 1-4, the standard words were *sit, plant, wonder*, and *relation*, respectively, and the comparison words were systematically varied alongside the standard words.

According to Vitz and Winkler, there were two checks on the reliability of their experiments. As for test-retest reliability, the results of the first trial (standard-first order) correlated highly ($r > .80$ in all experiments) with those of the second trial (reversed order). The between-subjects reliability was also high ($r > .75$ in all experiments). The main result was that the correlations between PPD values and subjects' ratings were also very high (from -.81 to -.95 in Experiments 1 to 4).

Experiment 5 applied the PPD to the mean ratings of Nelson and Nelson's (1970) 96 pairs. The correlation was again very high (-.94). This high accuracy of PPD predictions suggested that subjects' judgments were made primarily on the basis of a phoneme code. Derwing and Nearey (1986) also argued that, in spite of its rather simple algorithm, the PPD had a good theoretical basis:

> If, after all, speakers are supposed to be insensitive to the predictable variations manifested in their language, then a phonemic representation,

> which is an attempt to abstract away from precisely this kind of variation, ought to provide a good prediction of the judgments these speakers make about similarity in sound between utterances in their language. Furthermore, if the boundaries between phonemes are more or less categorical ("all or nothing"), then a simple count of the shared phonemes between two forms ought to provide a perfectly adequate basis for this kind of prediction, since all phonemes would presumably be judged as equidistant from each other. (p. 197)

Vitz and Winkler (1973) went further to argue that "relatively little of the variance in the rating of complete words is due to factors existing at a lower or more molecular level than the phoneme" (p. 386). However, this argument ignores other work indicating that all phonemes are not equally dissimilar or confusable (e.g., Cole, Haber, and Sales, 1968; Cole, Sales, and Haber, 1969; Greenberg and Jenkins, 1964; Sales, Cole, and Haber, 1969; Wickelgren, 1965, 1966).[2]

Another major finding in Vitz and Winkler (1973) was the position effect; when the number of shared phonemes was controlled, some pairs were judged more similar than others by virtue of having the shared phonemes in a more salient position. The results of Experiment 1 showed this position effect, as summarized in Table 2.2. The five pairs in Table 2.2 all had the same predicted PPD of .33, since they had one mismatched phoneme out of three, but the first two (sharing a rime) had the highest similarity scores.

In sum, the two main factors in predicting similarity scores in Vitz and Winkler (1973) were the *number* of matched phonemes (cf. matched letters in Nelson and Nelson, 1970) and the *position* of matched phonemes. Although these two studies used different modes of stimulus presentation (i.e., written vs. aural; or letter vs. phoneme), the results were quite similar.

TABLE 2.2. POSITION EFFECT IN VITZ AND WINKLER'S (1973) EXPERIMENT 1

Standard word	Comparison word	Shared unit	Similarity score (range: 0 - 4)
/sɪt/	/hɪt/	rime	3.16
/sɪt/	/pɪt/	rime	3.10
/sɪt/	/sɪk/	body	2.89
/sɪt/	/sɛt/	margins	2.86
/sɪt/	/sæt/	margins	2.34

2.2.3. Derwing and Nearey (1986)

Derwing and Nearey (1986) sought to refine the PPD,[3] suspecting that the effect of sub-phonemic feature differences might be much greater than Vitz and Winkler (1973) had assumed. Here, therefore, all attention was focused on the similarity of consonants, and all the stimulus pairs were CVC-CVC, in which either the initial or the final consonant contrasted. There were six kinds of contrasts in both initial and final positions, as summarized in Table 2.3.

A 10-point rating scale was used, ranging from 0 (totally different) to 9 (identical). According to the PPD, all of these pairs should have the same similarity rating (i.e., .33), since only one phoneme was mismatched out of three. However, there was a significant difference among the pair types ($F[11{,}444] = 20.5$, $p < .0001$). These results suggested that the degree of similarity between phonemes or the potential role of distinctive features could not be dismissed in predicting subjects' judgments of sound similarity.

As can be seen in Table 2.3, pairs containing mismatched initial consonants (μ=.317) were judged more similar than pairs containing mismatched final consonants (μ=.087). Since in CVC pairs, disagreement in initial C is tantamount to agreement in

VC, and disagreement in final C to agreement in CV, this result supported the rime effect found in Nelson and Nelson (1970) and Vitz and Winkler (1973) (i.e., identity of the rime contributes more to the similarity of CVC pairs than that of the body or the margins in English).

TABLE 2.3. RESULTS OF DERWING AND NEAREY (1986)

Position	Contrast	Example	Average Z score	
Initial	Voice	pɪt-bɪt	.520	
	Place in non-voiced	pɔp-tɔp	.310	
	Place in voiced	bɔp-dɔp	.380	
	Manner	pɪt-fɪt	.223	
	Voice & Place	pɛt-gɛt	.251	
	Voice, Place & Manner	bɪt-sɪt	.218	$\mu = .317$
Final	Voice	pɪp-pɪb	.489	
	Place in non-voiced	pɔp-pɔt	.491	
	Place in voiced	pɔb-pɔd	.297	
	Manner	kæp-kæf	-.208	
	Voice & Place	pɪg-pɪp	-.104	
	Voice, Place & Manner	gæg-gæs	-.481	$\mu = .087$

Second, the manner feature was more important than place and voicing in both initial and final consonants. Finally, as expected, there was a strong correlation between the number of feature mismatches and the similarity ratings. All these results suggest that the degree of similarity between phonemes (i.e., sub-phonemic differences) could not be ignored and that some kind of feature analysis was reflected in the SSJ experiments.

2.2.4. Bendrien (1992)

Bendrien (1992) addressed two main issues. First, the effects of features and the relative weights of initial and final consonants were tested by systematically varying consonants in CVC pairs, while keeping the vowel the same. In both positions there were three types of consonant contrasts: (i) match (e.g., /p-p/); (ii) minor mismatch (e.g., /p-b/); and (iii) major mismatch (e.g., /p-z/). The results were largely consistent with Derwing and Nearey (1986): (a) There was a significant difference among the above three types of contrasts and (b) mismatches in the final consonants counted more than those in the initial consonants. However, there were some limitations in generalizing the results found in Bendrien (1992). Since only three consonants /p, b, z/ were used for consonant contrasts, the manner and the place feature could not be explored for the single feature mismatches.

The second issue pointed to the level at which subjects make SSJs. This issue is related to the relative weightings between initial and final consonants discussed in the previous sections. Dow (1987) argued that final consonants could count more in the SSJ experiments because of a recency effect (i.e., more attention is paid to the more recent changes), and that "there is no motivation to claim that the onsets and codas of CVC and CCVCC type syllables have different weights." (p. 164). She found no significant difference between the two positions when this recency effect was controlled for by repeating pairs in two orders and inserting an intrusive element between test items. If the final consonant were more important than the initial one in SSJs just because of the recency effect, as Dow (1987) argued, then this effect should be universal. However, for Arabic CVCVC pairs, Beinert and Derwing (1993) found that the second consonant was most important. Wang and Derwing (1993) found that for Taiwanese CVC pairs the initial consonants were more important than the final ones.[4] These results from cross-linguistic studies suggest that higher-level phonological factors are at work here, not simply a recency effect. In particular, since both super-phonemic (e.g., rime, body, mora) and sub-phonemic units (i.e., features) can play a role, we conclude that the structure of the entire syllable is taken into account when SSJ judgments are made.

2.2.5. Summary

So far, we have observed the following effects on sound similarity in English: (i) the number of matched phonemes; (ii) different weights of subsyllabic units; (iii) the number of mismatched features, i.e., single feature differences vs. multi feature differences; (iv) the kind of features, i.e., manner features being more important than other consonantal features; and (v) the level at which SSJs are made, i.e., phonetic or phonological. The details of these studies are summarized in Table 2.4 below.

TABLE 2.4. SUMMARY DESCRIPTION OF THE LITERATURE REVIEWED

	Stimuli		Scale	Main Finding
Nelson and Nelson	CVC	visual	7-point	phoneme and position effects
Vitz and Winkler	CVC, etc.	aural	5-point	phoneme effect (PPD)
Derwing and Nearey	CVC	aural	10-point	position and feature effects
Bendrien	CVC	aural	percent	level at which SSJs are made

2.3. Experiment 1

2.3.1. Goals

There were two main goals in Experiment 1: to assess (i) the relative importance of various subsyllabic units and (ii) the status of the phoneme as a representational unit in Korean. Among the possible subcomponents of a CVC syllable, the focus of comparison was on the body vs. the rime. The salience of these two subsyllabic units was compared in two ways. One way was to compare the mean similarity scores given to pairs sharing the body with those given to rime-sharing pairs. The other way was to run regression analyses, treating as variables the body and the rime (as well as other units) and to compare the amount of variance in similarity scores explained by the body variable (among other variables) *vis-a-vis* the rime variable (among others).

As discussed in the previous section, we have reason to believe that syllable structure is invoked when the SSJs are made. If the body is a more salient part of the Korean syllable than the rime, the body-sharing pairs (e.g., /pan-pat/) should be judged more similar than the rime-sharing pairs (e.g., /pan-tan/), and the body variable should account for more variance than the rime variable. On the other hand, if the rime is more salient than the body as in the English syllable, opposite results should be found.

A second goal was to test the status of the phoneme as a basic representational unit in Korean. As in the Vitz and Winkler study, a useful measure to test this is a correlation between rated similarity and predicted similarity based on phoneme counting. The previous English SSJ experiments showed that a simple count of matched phonemes between words explained a substantial portion (about 80%) of the total variation in similarity scores (Nelson and Nelson, 1970, for tightly controlled CVC-CVC pairs; and Vitz and Winkler, 1973, for a variety of types of pairs). Focusing on CVC-CVC pairs, the present experiment tested the role of the phoneme in predicting the SSJs of monosyllabic Korean words. If the phoneme in Korean is as basic a representational unit as it is in English, comparable results should be found.

2.3.2. Method

2.3.2.1. Subjects

Thirty undergraduate students at Sogang University participated in the experiment on a voluntary basis.[5] All subjects in this and the following experiments were native speakers of Korean.

2.3.2.2. Stimuli

There were a total of 56 CVC test pairs (seven types of phonemic matches × eight tokens in each type).[6] As presented in Table 2.5 below, the seven types ranged from one 0-phoneme matched, three 1-phoneme matched, to three 2-phoneme matched pairs. In addition to these test pairs, there were eight control pairs of full phonemic identity, e.g., /pan-pan/. Although there were no correct answers indicated for any of the test pairs, these identity pairs should always be rated as nine, i.e., the maximum score on our zero-to-nine 10-point scale, and they were included in part to check if the subjects understood and were following the instructions. A criterion for these eight control pairs was arbitrarily set: if a subject gave a score lower than seven for more than two of the eight control pairs, his or her results were not included in the analysis. (As it turned out, only one subject's results had to be eliminated on this basis.)

TABLE 2.5. EXAMPLES OF STIMULUS PAIRS USED IN EXPERIMENT 1

Number of matched phonemes	Predicted Phonemic Similarity	Subsyllabic units matched	Example
0	0	None	pan-mət
1	.33	Onset	pan-pət
1	.33	Vowel	pan-mat
1	.33	Coda	pan-mən
2	.67	Body	pan-pat
2	.67	Rime	pan-mat
2	.67	Margins	pan-pən

NOTE: For the underlined segments, see below.

The basic stimulus set was derived from the four consonants /p, t, m, n/ and the four vowels /a, ə, o, u/. To control for the degree of difference between phonemes (Greenberg and Jenkins, 1964; Derwing and Nearey, 1986), mismatched consonants in each pair were designed to be always one distinctive feature apart from each other: /p-t/, /m-n/ for place of articulation, and /p-m/, /t-n/ for manner of articulation. Similarly, vowel contrasts were limited to either /a-ə/ or /o-u/, both of which are also only one distinctive feature apart in terms of [height].[7] From these basic sets of consonants and vowels, the following four 'standard words' were chosen: /pan/, /təm/, /mot/, and /nup/, and each standard word was paired with 14 comparison words to make the total of 56 pairs. (See Appendix 1 for a full list of stimuli used in Experiment 1.) Table 2.5 presents seven comparison words paired with the standard word /pan/. On the basis of the totally different comparison word /mət/, the other types of comparison words were constructed by filling out the remaining comparison sets, as illustrated by the underlined segments in the table.

2.3.2.3. Procedure

The instructions were read as follows (in Korean):

> This is an experiment about sound similarity between two Korean words. Each test item is composed of two words, and will be pronounced twice. We ask you to listen to these pairs and to rate how similar the two words *sound to you* on a 10-point scale, as given on the test booklet. If you think two words have a completely identical sound (e.g., *kam-kam*), please choose 9. On the other hand, if you think two words have a completely different sound (e.g., *kam-sil*), please choose 0. However,

most of the items are in between these two extremes (e.g., *kam-kil* and *kam-kal*), in which case you are to rate between 1 and 8.

We understand that this is a very unusual and subjective task. There are no right or wrong answers for most items. So please do not worry about anything else but just try faithfully to rate the sound similarity on the basis of your own intuition. Remember: the more similar the sounds, the bigger the number! Are there any questions?

The stimulus pairs were recorded in a single random order, and played back to subjects roughly as follows:

FIGURE 2.2. THE SAMPLE OF A SINGLE STIMULUS PAIR PRESENTATION

TIME → [1s] [1s] [2s] [1s] [4s]
TAPE: "Number 1" -- "pan" -- "mət" ---- "pan" -- "mət" -------- "Number 2" --
SUBJECT: ________________________________ response __________

Four practice pairs representing the number of shared phonemes from 0 to 3 were presented before the test to get subjects accustomed to the use of a scale, but without providing any suggested answers. Subjects were instructed to focus on the global impression of sound only and to rate the similarity in sound on a 10-point scale ranging from 0 (completely different) to 9 (exactly the same) by filling in the corresponding circle on the answer sheet. No orthographic representations of the stimulus pairs were provided.

2.3.3. Results

Only one subject did not meet the inclusion criterion, making three mistakes out of eight identical pairs. The following analyses were thus based on the remaining 29 subjects.

In the results to be reported in this and the following experiments, mean similarity scores were calculated for each subject and each item, and statistical analyses treated both subjects and items as random factors (Clark, 1973). Thus, in most cases, statistical significance was based on both analyses.

Table 2.6 below presents the mean similarity scores for the seven types of test pairs. In the table, predicted phonemic "distance" (PPD) was converted to the predicted phonemic "similarity" (PPS) for ease of comparison. (PPS was used throughout all Korean SSJ experiments described in this thesis.)

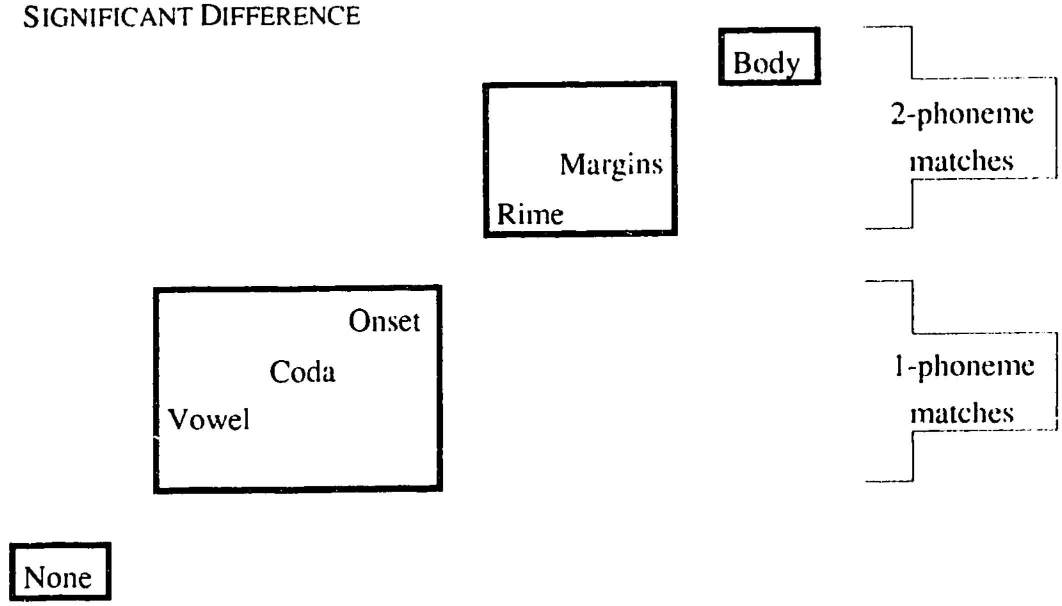
First, mean similarity scores were compared by analyses of variance with the three factors of (i) the seven types of matches, (ii) the four standard words, and (iii) two types of vowels. Only the first factor was significant: the match-types were significantly different both by subjects ($F_1[6,168] = 54.4, p < .001$) and by items ($F_2[6,28] = 40.87, p < .001$). The mean similarity scores did not differ across the four standard words nor the two types of vowels, and no interaction was significant.

TABLE 2.6. RESULTS OF EXPERIMENT 1

Subsyllabic units matched	Example	Mean Score	PPS
None	pan-mət	2.10	0
Vowel	pan-mat	3.07	.33
Coda	pan-mən	3.15	.33
Onset	pan-pət	3.43	.33
Rime	pan-man	4.78	.67
Margins	pan-pən	5.07	.67
Body	pan-pat	5.65	.67

Fisher's least significant differences among the seven types were .49 by subjects and .56 by items, distinguishing in both analyses the four groups in Figure 2.3 below. As can be seen in the figure, while there was no significant difference among one-phoneme matched pairs, in two phoneme-matched pairs, pairs with the body in common were judged significantly more similar than the other two types, which did not differ from one another. Separate *t* tests showed that pairs sharing the body were judged significantly more similar than pairs sharing the rime ($t2[7] = 2.07$, $p < .05$). This difference was clearly indicated by individual subjects' judgments. Of the total of 29 subjects, 25 (86.2%) judged that the body-sharing pairs were more similar in sound than the rime sharing pairs.

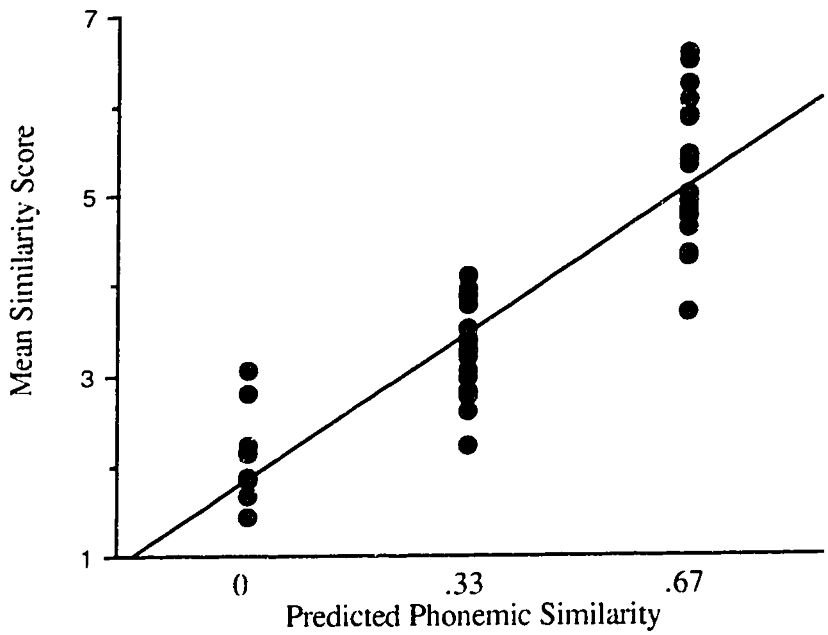
FIGURE 2.3. GROUPINGS OF THE SEVEN TYPES OF PAIRS IN TERMS OF THE LEAST SIGNIFICANT DIFFERENCE



Note: Boxes enclose means that are not significantly different.

Another feature clearly indicated in Figure 2.3 above is the importance of the number of matched segments in predicting the similarity scores. This was also found in the correlation between the PPS and mean similarity scores. Figure 2.4 below shows the correlation based on the item analysis, and each data point represents the mean similarity score of each of the 56 test pairs.

FIGURE 2.4. CORRELATION BETWEEN PREDICTED PHONEMIC SIMILARITY AND SIMILARITY SCORE (R^2 = .77)



The correlation was high enough (r = .88) to account for a significant portion (77%) of the variance ($F_2[1,54]$ =181.6, p < .001). The correlation based on the subject analysis was also significant (r = .59, r^2 = .34; $F_1[1,201]$ = 104.7, p < .001).

Finally, a series of linear regression analyses was run. The main purpose of these analyses was to identify contribution of each subsyllabic variable (e.g., body, rime) and various combinations of these variables (e.g., onset + rime, body + coda) to the prediction of similarity scores. As shown in Table 2.7, six sub-syllabic variables were used, and matches/mismatches for each variable were coded as 1 (when the specific condition was met) or 0 (otherwise):

TABLE 2.7. ILLUSTRATION OF THE LINEAR REGRESSION ANALYSIS

	Subsyllabic Variables (condition)					
Types of Pairs (X = mismatch)	Onset	Vowel	Coda	Body	Rime	Margins
CVC-XXX	0	0	0	0	0	0
CVC-CXX	1	0	0	0	0	0
CVC-XVX	0	1	0	0	0	0
CVC-XXC	0	0	1	0	0	0
CVC-CVX	1	1	0	1	0	0
CVC-XVC	0	1	1	0	1	0
CVC-CXC	1	0	1	0	0	1

Generally, the more variables that are included, the more variance in similarity scores is explained. The combination of all six variables resulted in the best overall coverage (81.1%), but not all variables were significant. The details of this analysis are shown in Table 2.8 below.

TABLE 2.8. RESULTS OF COMBINING ALL SIX VARIABLES

Variable	Coefficient	*t*-ratio	probability
Onset	1.33	4.64	< .001
Vowel	0.97	3.39	< .005
Coda	1.04	3.65	< .001
Body	1.25	3.08	< .005
Rime	0.66	1.64	.11
Margins	0.59	1.46	.15

As can be seen in Table 2.8, all three segment variables were highly significant, while the rime and the margin variables did not play any significant role. When only three segment variables were considered, the percentage of variance explained was 78.7%. Adding the body variable to the three-segment model raised the percentage of variance explained (80.6%), while adding the rime variable to the segment model lowered it (78.3%)! Differential weights between the body and the rime were also found when the following two-term models were compared: body + coda (55.6%) vs. onset + rime (53.5%). As might be expected, the margin + vowel model involving a discontinuous unit resulted in an even lower percentage (44.0%) than the onset + rime model.

2.3.4. Discussion

The two main results of Experiment 1 can be summarized as follows. First, the phoneme emerged as the most important factor in judging sound similarity of Korean CVC-CVC pairs: (i) ANOVAs of mean similarity scores for the seven types of segmental pairs showed significant differences between one-phoneme matched pairs and two-phoneme matched pairs; (ii) the correlation between the PPS and mean similarity scores explained the major portion of variance in similarity scores; and (iii) linear regression analysis found that all three segmental variables made highly significant contributions to predicting similarity scores.

In addition to this phoneme effect, a syllable structure effect also emerged. For pairs with two phonemes in common, pairs sharing the body unit were judged significantly more similar than pairs sharing the rime or the margins. Linear regression analyses also supported this. When combined with the segment model, in particular, the body variable played a secondary additive role in increasing the overall coverage. In sum, these results suggest that the Korean syllable has a left-branching hierarchical structure, which is consistent with our prior research (Derwing, Yoon, and Cho, 1993;

Yoon, 1994a; Yoon and Derwing, 1994). Yoon and Derwing (1994), for example, had English and Korean subjects judge the sound similarity of Korean CVC words. We found directly opposite results from the two groups of subjects for the same set of stimuli: to English subjects the rime was more important than the body in predicting similarity scores, whereas to Korean subjects the body was more important than the rime. Under the PPD model (Vitz and Winkler, 1973) or the "flat" view of syllable structure (e.g., Clements and Keyser, 1983), this differential weighting cannot be explained, since all phonemes ought to play an equal role regardless of their position within the syllable.

2.4. Experiment 2

2.4.1. Goals

One of the main findings of Experiment 1 with CVC-CVC pairs was a syllable structure effect, whereby pairs sharing the CV (body) were judged significantly more similar in sound than those sharing the VC (rime) -- a dramatically different result from English (cf. Nelson and Nelson, 1970). In these CVC-CVC pairs, a match in the first two phonemes is tantamount to a mismatch in the last phoneme, and a match in the last two phonemes is tantamount to a mismatch in the first phoneme. In Experiment 1 we controlled the mismatched phonemes to be one distinctive feature apart, so that the degree of mismatches between words would not be a factor in subjects' judgments about overall similarity of sound. Suppose, for example, that a body-sharing pair /pan-pam/ was judged more similar than a rime-sharing pair /pan-san/. Not only could this result be interpreted as a differential weighting between the body and the rime, (i.e., the syllable structure effect), but it might also reflect a feature effect, since it could be argued that the difference between /n/ and /m/ (involving only a difference in place of articulation) was less than the difference between /p/ and /s/ (involving both manner and place). Experiment 2 was designed to test the relative strength of syllable structure and feature effects, and whether the syllable structure effect found in Experiment 1 would hold up under various types of feature mismatches in the onset and in the coda.

A secondary goal of Experiment 2 was to test the extent to which distinctive features play a role in the perception of similarities among Korean consonants. The degree of feature mismatches in consonants was designed to vary in terms of (i) place of articulation (e.g., /p-t/); (ii) manner of articulation (e.g., /p-m/); and (iii) both (e.g., /p-n/ or /t-m/). If the phoneme is perceived as a set of distinctive features, phonemes differing by only one feature ought to be judged more similar than phonemes differing by two or more features. In English, subjects' judgments are sensitive to sub-phonemic differences of this kind. For instance, a perceptual confusion study by Miller and Nicely (1955) suggested that phonemes agreeing in nasality and voicing are more easily confused than others, and a study of errors in short-term memory by Wickelgren (1966) showed that consonants having the same place of articulation were the most likely to be substituted for each other. Speech errors reported in Fromkin (1971) also indicated that phonemes sharing all but one feature are the most likely to be transposed (e.g., *glear plue sky* ← *clear blue sky*). In a little different vein, Jaeger and Ohala (1984) showed that the concepts based on a manner feature were easier for English speakers to form than those based on a place feature. The implications from these studies are that phonemes tend to group together along various featural dimensions, and that all phonemes are not perceptually equidistant from one another.

Some English SSJ experiments also provided evidence for the effect of features on judged similarity. Using all possible pairwise comparisons of /pa, ta, ka, ba, da, ga/, Greenberg and Jenkins (1964) found that consonants differing by two features (e.g., /p-d/) produced lower similarity scores than consonants differing by only one feature (e.g., /p-b/ or /p-t/). On the other hand, Vitz and Winkler (1973) found that the rated similarity of pairs of words was not influenced by whether consonants differed by one feature or by two or more features, suggesting that words are not perceived and compared at a "molecular" level smaller than the phoneme. Derwing and Nearey (1986) explored this issue further by including fricatives in their CVC stimulus pairs. They found that (i) consonants agreeing in all but one feature (e.g., /p-b/) were rated more similar than consonants with several mismatches of features (e.g., /p-z/); (ii) the difference in manner feature was more detrimental to similarity scores than the difference in place or voicing; and (iii) these differences were larger for the final consonant than for the initial one.

The present experiment was designed to test both syllable structure and feature effects on the SSJs of CVC pairs in Korean, using the nine types of pairs presented in Table 2.9.

TABLE 2.9. NINE TYPES OF PAIRS IN EXPERIMENT 2

Types of pairs (X = mismatch)	Feature Mismatches [place]	[manner]	Both
CVC-CVX	CVX^p	CVX^m	CVX^B
CVC-XVC	X^pVC	X^mVC	X^BVC
CVC-XVX	X^pVX^p	X^mVX^m	X^BVX^B

First, it was hypothesized that, controlling for the kind of mismatched features, CV-sharing pairs should be judged more similar than VC-sharing pairs, if the left-branching view is correct for the Korean syllable. If the right-branching view is correct, an opposite result should be found. Second, controlling for the number of matched phonemes, pairs with the mismatches of the X^p and X^m types should be judged more similar than pairs with X^B mismatches, if subjects are sensitive to sub-phonemic differences. On the other hand, if subjects' judgments operated *only* at the phonemic level (as suggested by Vitz and Winkler, 1973), there should be no significant differences between the CV-sharing pairs and the VC-sharing pairs, nor between pairs with mismatches of any of the three types: X^p, X^m, and X^B.

2.4.2. Method

2.4.2.1. Subjects

Twenty five undergraduate students at Kunsan National University participated in the experiment on a voluntary basis.[8]

2.4.2.2. Stimuli

A total of 72 stimulus pairs were divided into nine types of contrasts as discussed above. Stimulus pairs used in Experiment 2 are given in Table 2.10 below.

TABLE 2.10. EXAMPLES OF STIMULUS PAIRS USED IN EXPERIMENT 2

Contrast type	Example	Contrast type	Example	Contrast type	Example
1. CVX^p	pan-pam	4. X^pVC	pan-tan	7. X^pVX^p	pan-tam
2. CVX^m	pan-pat	5. X^mVC	pan-man	8. X^mVX^m	pan-mat
3. CVX^B	pan-pap	6. X^BVC	pan-nan	9. X^BVX^B	pan-nap

As in Experiment 1, there were eight tokens in each type, and the same four standard words were used. (See Appendix 2 for a full list of stimuli used in Experiment 2)

There are some general characteristics of Korean phonology affecting the present experimental design: (i) Since voicing is not phonemic, only place and manner of articulation features were explored; and (ii) since Korean exhibits a strong tendency to maximize onset contrasts and minimize coda contrasts (i.e., only seven consonants /p, t, k, m, n, ŋ, l / are permissible in the coda), some contrasts frequently found in the onset (most notably, the three series of stops, e.g., /p, p', p^h/) are simply unavailable in the coda. To balance the number and kinds of contrasts in the onset and the coda, the five consonants /p, t, k, m, n/ occurring in both positions were selected. In addition to this basic stimulus set common to both positions, /s/ in the onset and /l/ in the coda were included to have a wider variety of feature contrasts. These two consonants, which were expected to lower similarity scores due to low frequency in the total stimuli, were balanced to occur in four pairs with the onset /s/ and four with the coda /l/. Finally, to check test-retest reliability, 10 pairs were repeated with an interval of at least three pairs. Again, half of them involved the onset contrasts and the other half coda contrasts.

2.4.2.3. Procedure

The procedure was identical to that of Experiment 1.

2.4.3. Results

Of the total of 25 subjects, two gave up about half way through the list of stimuli and one misunderstood the instructions. The data from these three subjects were eliminated. As in Experiment 1, mean similarity scores for the four standard words were not different at all, and the scores were pooled across standard words. Furthermore, subjects were consistent in their ratings, judging from a test-retest measure: the correlation between mean similarity scores of the two sets of 10 repeated pairs was highly significant ($r = .5$; $F[1,217] = 70.89$, $p < .001$).

Table 2.11 below presents the main results. The first analysis involved a one-way analysis of variance on the six types of pairs that differed by on phoneme (the three CVX's and three XVC's). According to a prediction based only on the number of matched phonemes, all these pairs should have the same score, since they all had one mismatched phoneme out of three. However, the results showed highly significant differences among these types ($F_1[5,105] = 15.29$; $F_2[5,42] = 6.27$, $p < .001$ for both).

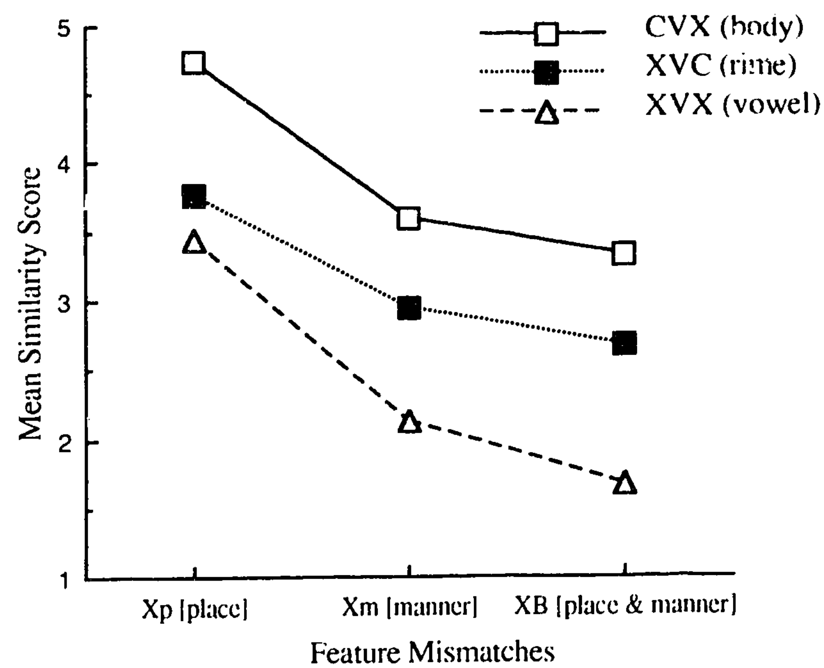
TABLE 2.11. RESULTS OF EXPERIMENT 2

Feature Mismatches	Subsyllabic Matches CVX	XVC	XVX	Mean
X^P [place]	4.73	3.76	3.45	3.98
X^m [manner]	3.60	2.86	2.12	2.86
X^B [place] and [manner]	3.33	2.67	1.67	2.56
Mean	3.89	3.10	2.41	3.13

To test both syllable structure and feature effects as well as an interaction, mean similarity scores for each subject and each pair were compared by ANOVAs conducted on the two effects. First, the three types of subsyllabic matches were significantly different ($F_1[2,42] = 17.59$; $F_2[2,63] = 25.81$, $p < .001$ for both). The three types of feature matches also had a highly significant effect on similarity scores ($F_1[2,42] = 36.96$; $F_2[2,63] = 26$, $p < .001$ for both). However, the interaction was not significant at the .05 level either by subjects or by items, indicating that types of subsyllabic matches did not vary as a function of types of feature mismatches, and vice versa. The results of these analyses are graphed in Figure 2.5 below.

To examine both factors in more detail, Fisher's least significant differences (LSD) were calculated on the three types of subsyllabic matches and the three types of feature mismatches. As for the syllable structure effect, the LSD's were .51 by subjects and .4 by items, showing that CVX pairs were judged significantly more similar than the XVC pairs (the mean difference of .79; $t_1[21] = 3.45$, $p < .005$: $t_2[23] = 3.6$, $p < .001$), and that the XVC pairs were in turn judged significantly more similar than the XVX pairs (the mean difference of .69; $t1[21] = 3.42$, $p < .005$; $t2[23] = 4.08$, $p < .001$). As for the feature effect, results were mixed (the LSD's were .34 by subjects and .38 by items). While the X^p pairs involving differences in [place] only were judged significantly more similar than X^m pairs involving differences in [manner] only (the mean difference of 1.11; $t1[21] = 5.41$; $t2[23] = 5.2$, $p < .001$ for both), there was no significant difference between the X^m and the X^B pairs involving mismatches in both features (the mean difference of .3).

FIGURE 2.5. MEAN SIMILARITY SCORE AS A FUNCTION OF SUBSYLLABIC MATCHES AND FEATURE MISMATCHES



Another way to examine both syllable structure and feature effects is by means of linear regression analyses. In addition to the two subsyllabic variables, [Body] and [Rime], the following four feature variables were used: [Op] = place matched in the onset, [Om] = manner matched in the onset, [Cp] = place matched in the coda, [Cm] =

manner matched in the coda. As illustrated for Experiment 1 in Table 2.7, matches/ mismatches for each variable were coded as 1 when the specific condition was met or as 0 otherwise. For the [Op] variable, for example, while XPVC pairs (i.e., place mismatched in the onset) and $X^{B}VC$ pairs (i.e., place and manner mismatched in the onset) were coded as 0, $x^{m}VC$ pairs (i.e., manner mismatched in the onset) were coded as 1, since place was matched in this latter case. The main purpose of these regression analyses was to compare the variance in similarity scores explained by various combinations of [Body], [Cp], [Cm] *vis-a-vis* [Op], [Om], [Rime] variables. Table 2.12 presents the summarized results of these analyses.

TABLE 2.12. PERCENTAGE OF VARIANCE EXPLAINED BY SUBSYLLABIC AND FEATURE VARIABLES

Variables	% of varinace explained	Variables	% of varinace explained
Body + Cp	37.2%	Op + Rime	7.6%
Body + Cm	87.3%	Om + Rime	73.4%
Body + Cp + Cm	87.4%	Op + Om + Rime	81.0%

Overall, the combinations of Body plus Coda variables explained higher percentage of variance than did the combinations of [Onset] plus [Rime] variables. In particular, the [Body] variable made a significant contribution to the percentage shown in the last two rows of Table 2.12 ($p < .005$ for both), while the [Rime] variable did not make any significant contribution. Also reflected in these results was the significant contribution of manner feature variables both in the onset and in the coda. In all cases, the contribution of both [Om] ($p < .01$) and [Cm] variables ($p < .005$) was significant, while that of [Op] and [Cp] did not reach significance at the .05 level. In sum, [Body] and [Cm] were the two most important variables in predicting the similarity scores, the details of which are presented in Table 2.13 below.

TABLE 2.13. THE BEST COVERAGE ACHIEVED BY THE COMBINATION OF A SUBSYLLABIC AND A FEATURE VARIABLE

Variable	Coefficient	*t*-ratio	probability
Constant	1.90	8.11	< .001
Body	1.55	5.51	< .005
C^m	1.30	4.87	< .005

Finally, to examine individual variation in the similarity ratings, the difference between the CVX pairs and the XVC pairs in all three feature mismatched conditions was analyzed for each subject. First, for place feature mismatched pairs, CVX^p pairs were rated higher than X^pVC pairs by 18 subjects (82%), the same by two, and lower by two. For manner feature mismatched pairs, CVX^m pairs were rated higher than X^mVC pairs by 16 subjects (73%), the same by one, and lower by five. For double-feature mismatched pairs, CVX^B were rated higher than X^BVC pairs by 17 subjects (77%), the same by one, and lower by four. In sum, in 51 instances (77%) out of the total of 66 (i.e., 3 feature mismatch conditions × 22 subjects), the body-sharing pairs were judged more similar than the rime-sharing pairs; in 4 instances (6%) both types were even; and in 11 instances (17%) the rime-sharing pairs were judged more similar than the body-sharing pairs. Interestingly, two subjects rated the body-sharing pairs lower than the rime-sharing pairs in all three feature conditions, and were thus responsible for 6 out of the 11 instances where the rime-sharing pairs predominated. Most of the other subjects, therefore, consistently judged that the body-sharing pairs sounded more similar than the rime-sharing pairs.

2.4.4. Discussion

The results reported above showed that both syllable structure and feature effects were significant. First, matches in the body were generally more important than matches in the rime in making pairs sound more similar, replicating the results of Experiment 1. ANOVAs showed that the body-sharing pairs were judged significantly more similar than the rime-sharing pairs, and these results did not vary as a function of mismatched features. Linear regression analyses also showed that the body variable made a significant contribution to the prediction of similarity scores, while the rime variable did not. Furthermore, these results were maintained consistently for most of the subjects. All these results su[illegible] view (left-branching structure) of the Korean syllable. The difference betw[illegible]aring pairs and the rime-sharing pairs is impressive in that it is as large [illegible]e between the rime-sharing pairs and the vowel-sharing pairs. This is difficult to explain on the basis of the flat view of the Korean syllable, and still more difficult on the basis of the rime view.

The other major finding is that manner of articulation is more important than place of articulation in the SSJs of Korean CVC pairs. Mismatches in manner were more detrimental to similarity ratings than mismatches in place. In a similar vein, linear regression analyses showed that the manner feature variable both in the onset and the coda made a significant contribution to predicting similarity scores, while the place feature variable did not.

Surprisingly, controlling for the number of mismatched phonemes, there was no significant difference between pairs with the manner feature mismatched and pairs with both features mismatched. This result is at odds with the English results reported in Greenberg and Jenkins (1964) and Derwing and Nearey (1986). However, there was a major difference in stimuli between the present experiment and these previous English experiments. Since voicing is not phonemic in Korean consonants, a direct comparison might be impossible with the above English experiments, which included voicing in addition to the features tested in the present Korean experiment. It also appears to provide additional evidence for the importance of manner of articulation relative to place of articulation in the feature dimensions of Korean consonants.

2.5. Experiment 3 [9]

2.5.1. Goals

In Experiment 1 we found the phoneme effect in the SSJs of Korean monosyllabic words, replicating the results reported in Yoon and Derwing (1994): the simple counting of matched phonemes between two words explained a substantial portion of the total variance in similarity scores. These results from Korean add to the body of evidence for the phoneme as a basic representational unit found in previous SSJ experiments with English (Vitz and Winkler, 1973) and other languages (Derwing and Nearey, 1994).

However, since only monosyllabic pairs were used in Experiment 1, the salience of the phoneme could only be measured in terms of the variance explained by the phonemic code itself, but not against the syllable, another potentially salient unit in Korean. When examining the role of the mora in speech segmentation of Japanese, Otake, Hatano, Cutler, and Mehler (1993) expressed an opinion on the role of orthography, as follows:

> Concepts such as phoneme, mora, syllable, and stress unit are not properties of particular languages; they are phonological properties in terms of any language can be described. However, just as there are differences across language communities in the way phonology may be reflected in the orthography, so are there differences across languages in

> the degree to which various phonological constructs play a role in phonological processes within the language. (p. 260)

A recent SSJ study in Japanese (Derwing and Wiebe, 1994) also implicated the orthography. The predictions of SSJs based on the mora were more accurate than those based on the phoneme, which played a secondary role comparable to the rime in English. Since the Japanese *hiragana* and *katakana* alphabets are both mora- rather than segment-based, however, one cannot exclude the possibility of orthographic influence in this latter case.

The primary goal of Experiment 3 was to examine the basic level of representation in Korean by pitting the phoneme against the syllable. All test pairs were CV.CVC structures (where "." indicates a syllable boundary)[10] and were systematically varied from pairs that had all but one phoneme in common (e.g., CV.CVC-XV.CVC, where "X" indicates a mismatched phoneme) to pairs that had no common phonemes. Two hypotheses were tested which, depending on the number of phonemes and syllables mismatched, predicted different patterns of results. The method and predictions were based on the correlation between mean similarity scores and (i) predicted "syllabic" similarity (PSS) vs. (ii) predicted "phonemic" similarity (PPS), as illustrated in Table 2.14. The syllabic hypothesis predicted that, controlling for the number of mismatched phonemes, pairs with mismatches *across* the syllable boundary (e.g., CX.XVC; /ku.saŋ-ko.caŋ/) should be judged less similar than pairs with mismatches *within* a syllable (e.g., CV.CXX; /ku.saŋ-ku.sək/), since mismatches in the former involve both syllables, while those in the latter involve one syllable only.[11] On the other hand, the phonemic hypothesis predicted that there should be no significant difference between the two types of pairs, since both have the same number of mismatched phonemes.

TABLE 2.14. COMPARISONS OF PREDICTED SIMILARITIES ON THE BASIS OF COUNTING MATCHED SYLLABLES AND PHONEMES

Mismatch types	PSS S_1	S_2			PPS P_1	P_2	P_3	P_4	P_5		
XV.CVC	0 +	1	/2	= 0.5	0 +	1 +	1 +	1 +	1	/5	= 0.8
XX.CVC	0 +	1	/2	= 0.5	0 +	0 +	1 +	1 +	1	/5	= 0.6
CX.XVC	0 +	0	/2	= 0	1 +	0 +	0 +	1 +	1	/5	= 0.6
XX.XVC	0 +	0	/2	= 0	0 +	0 +	0 +	1 +	1	/5	= 0.4
CV.XXX	1 +	0	/2	= 0.5	1 +	1 +	0 +	0 +	0	/5	= 0.4

2.5.2. Method

2.5.2.1. Subjects

A total of 117 subjects participated in the experiment on a voluntary basis.[12] There were three groups: (i) 43 middle school students with aural stimuli only (MA); (ii) 44 middle school students with both aural and visual stimuli (MB); and (iii) 30 university students with both aural and visual stimuli (UB). These groupings were designed to test the effects of presentation mode (MA vs. MB) and age/education (MB vs. UB) on SSJs.

2.5.2.2. Stimuli

Twelve types of CV.CVC pairs with four tokens each were selected. As presented in Table 2.15 below, the 12 types ranged from identical pairs (i.e., 0-phoneme or 0-syllable mismatched) to totally different pairs (i.e., 5-phoneme or 2-syllable mismatched). (See Appendix 3 for the full list of stimuli used in Experiment 3.)

TABLE 2.15. TWELVE TYPES OF PHONEMIC MISMATCHES USED IN EXPERIMENT 3

Type		Example	Type		Example
0	CV.CVC	ku.saŋ - ku.saŋ	2C	CV.XXC	co.cəŋ - co.saŋ
1A	XV.CVC	ku.saŋ - pu.saŋ	2D	CV.CXX	ku.saŋ - ku.sək
1B	CV.XVC	co.saŋ - co.caŋ	3A	XX.XVC	su.saŋ - co.caŋ
1C	CV.CVX	co.səŋ - co.sək	3B	CX.XXC	ku.saŋ - ko.cəŋ
2A	XX.CVC	ku.saŋ - po.saŋ	3C	CV.XXX	po.cən - po.saŋ
2B	CX.XVC	ku.saŋ - ko.caŋ	5	XX.XXX	ku.saŋ - po.cən

NOTE: X and numerals indicate location and number of mismatches, respectively.

The focus of our attention was on the 2- and 3-phoneme mismatched pairs, as they yield differential predictions under the Phoneme Hypothesis and the Syllable Hypothesis, depending on whether the mismatched phonemes are within a single syllable or spread over both syllables.

The following controls were built into the stimulus pairs. First, all stimuli were real words, largely to prevent the mis-hearing of stimuli that is common with nonsense words. Second, syllable boundaries (in the orthography) always occurred before the second C. Third, all mismatched phonemes between words differed by one distinctive feature, to control the degree of difference between phonemes. Finally, four identity pairs were included as control items to see if the subjects understood and were following the instructions. (Subjects who rated two or more identity pairs lower than 7 on our 0-to-9 scale were dropped from the experiment.)

2.5.2.3. Procedure

For the subjects in the MA group, the procedure was the same as that used in Experiments 1 and 2. For the subjects in the MB and UB groups, the procedure was the same as for the MA group except that they were provided with orthographic representations of each test pair.

2.5.3. Results

Of the total of 117 subjects, nine subjects did not meet the inclusion criterion (four in the MA and five in the MB group). Results reported below were thus based on the remaining 108 subjects (39 MA, 39 MB, and 30 UB subjects).

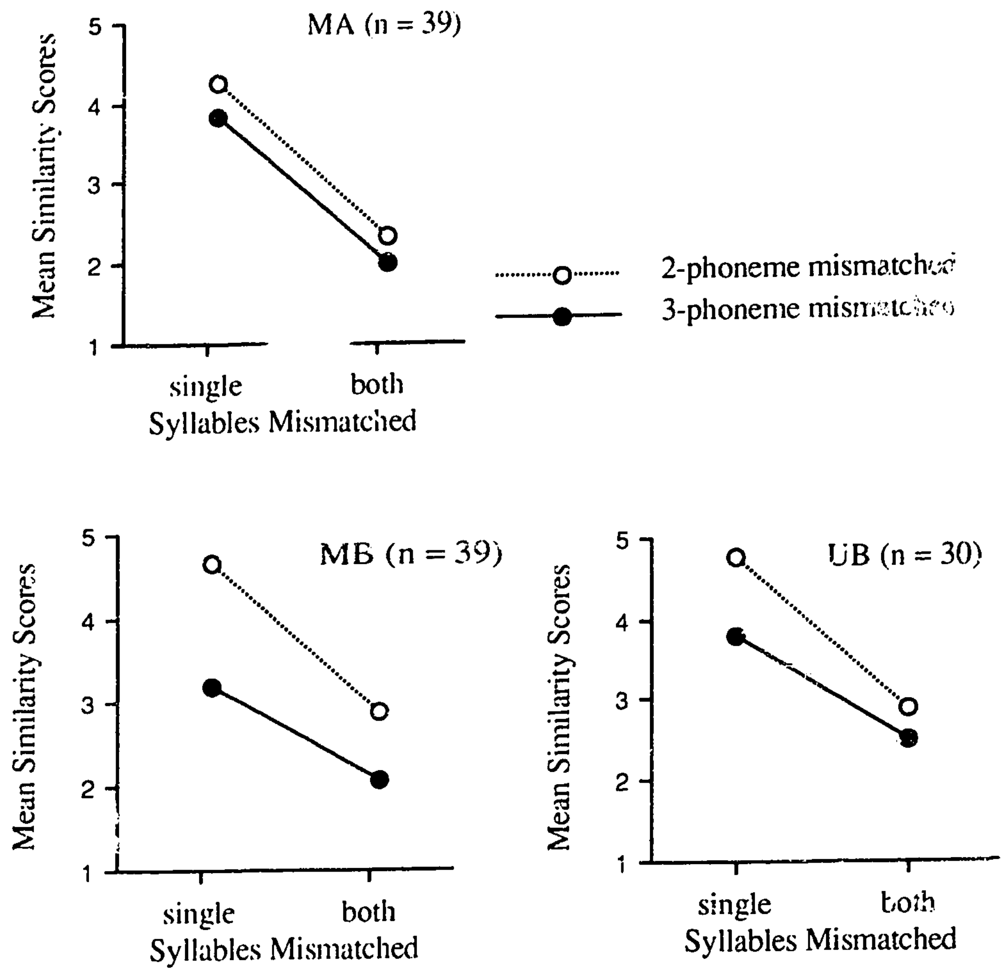
To compare the Syllable and the Phoneme hypotheses, two statistics were used. First, a series of ANOVAs was run on mean similarity scores of four 2-phoneme mismatched pairs and three 3-phoneme mismatched pairs, treating both subjects and items as random factors. In these ANOVAs, the number of mismatched syllables (the Syllable effect) and the number of mismatched phonemes (the Phoneme effect) were within-subject variables. Later, the Group effect was examined as a between-subject variable to test the additional effects of presentation mode (i.e., MA vs. MB) and age/education (i.e., MB vs. UB). In these group comparisons, *Z*-scores for each item were calculated to calibrate the similarity scores of each group, since subjects in the three groups varied somewhat in their use of the scale.

Another way to assess the two hypotheses was to compare the variance in similarity scores explained by the Predicted Syllabic Similarity (PSS) vs. the Predicted Phonemic Similarity (PPS). In these correlation analyses, mean similarity scores of all 44 test pairs were used (excluding the four control pairs).

2.5.3.1. ANOVAs

Overall, the syllable was a more important effect than the phoneme, and the interaction between the two effects was not significant by either subjects or items in all three groups. The results of these analyses are graphed in Figure 2.6.

FIGURE 2.6. MEAN SIMILARITY SCORES AS A FUNCTION OF THE NUMBER OF SYLLABLES AND PHONEMES MISMATCHED



First, in the MA group, the syllable effect was highly significant both by subjects and by items ($F_1[1,38] = 55.4$; $F_2[1,24] = 51.5$, $p < .001$ for both). However, the phoneme effect did not reach significance at the .05 level either by subjects or items. The interaction was also not significant, as displayed in Figure 2.6. As also shown in the figure, the CV.XXX pair ("•" in upper left of the graph) which involved three mismatched phonemes was rated much more similar than the CX.XVC pair ("o" in lower right of the graph), which involved only two mismatched phonemes.

Second, the results for the MB group, where subjects were provided with orthographic representations of stimulus pairs, were somewhat different from those of the MA group, where only aural stimuli were provided. For these subjects, not only did the syllable effect again emerge as a highly significant variable ($F_1[1,38] = 42.7$; $F_2[1,24] = 30.2$), $p < .001$ for both), but the phoneme effect was also significant ($F_1[1,38] = 25.9$; $F_2[1,24] = 17.0$, $p < .001$ for both). However, the interaction was again not significant.

Third, the results of the UB group were similar to those of the MB group, suggesting that the age/education effect might not be important. The syllable effect was again highly significant ($F_1[1,29] = 51.9$; $F_2(1,24) = 33.7$, $p < .001$ for both) and the phoneme effect was also significant, but with a lower level of significance than for the MB group ($F_1[1,29] = 9.03$, $p < .01$; $F_2[1,24] = 6.04$, $p < .05$). As in the other two groups, there was no interaction between the two effects.

Finally, to compare the seven types of pairs in detail, Fisher's least significant differences (LSD) were calculated for the three groups: MA = .81, MB = .96, UB = 1.01. Table 2.16 presents mean similarity scores for each type and *LSD* groupings. (Solid vertical lines include the means that are not significantly different.)

TABLE 2.16. RESULTS OF THE THREE GROUPS IN TERMS OF *LSD* GROUPINGS

	MA	MB	UB
	CV.CXX (4.9)	CV.CXX (5.0)	CV.CXX (5.1)
	CV.XXC (3.9)	CV.XXC (4.6)	XX.CVC (4.8)
	CV.XXX (3.7)	XX.CVC (4.2)	CV.XXC (4.5)
	XX.CVC (3.5)	CV.XXX (3.1)	CV.XXX (3.8)
	CX.XVC (2.2)	CX.XVC (2.8)	CX.XVC (2.9)
	CX.XXC (1.9)	CX.XXC (2.0)	XX.XVC (2.5)
	XX.XVC (1.9)	XX.XVC (1.9)	CX.XXC (2.5)
mean	3.15 (sd=1.18)	3.34 (sd=1.30)	3.72 (sd=1.20)

Again, the syllable effect is obvious in the above groupings. In all three groups, regardless of the number of mismatched phonemes, the four 1-syllable mismatched pairs were rated higher than three 2-syllable mismatched pairs (as indicated by the dotted lines in the Table 2.16 above). This effect was most obvious for the MA group, where there were no overlapping types between 1- and 2-syllable mismatched pairs in terms of the *LSD* groupings.

2.5.3.2. Correlations

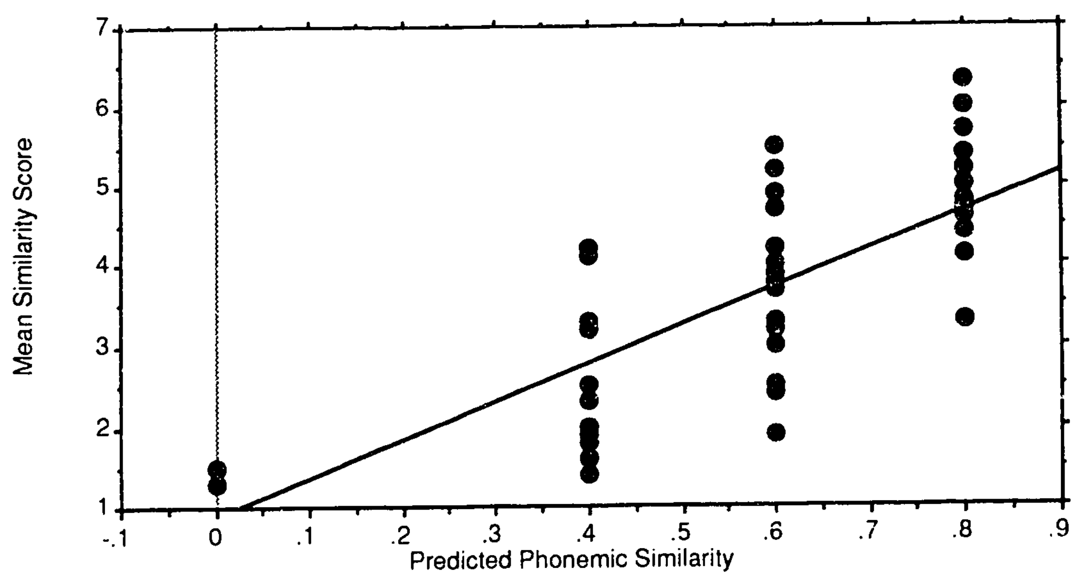
Mean similarity scores of all 44 test pairs were used. The correlations between these mean similarity scores and the Predicted Syllabic Similarity (PSS) vs. the Predicted Phonemic Similarity (PPS) are presented in Table 2.17.

TABLE 2.17. CORRELATIONS BETWEEN SIMILARITY SCORES AND PSS VS. PPS

	PSS		PPS	
	r	(r^2)	r	(r^2)
MA	.84	(.70)	.73	(.53)
MB	.83	(.69)	.84	(.71)
UB	.83	(.69)	.86	(.75)

Except for the PPS in the MA group, all correlations were higher than .8, explaining about 70% of the total variance in similarity scores. Again, the syllable effect was obvious in the MA group, where the percentage of variance explained by the phoneme counting (53%) was much less than by the syllable counting (70%). Figure 2.7 displays all data points and the least squares line for the PPS in the MA group. It can be seen that the data points at the PPS of .6, in particular, are scattered widely along the least squares line. At this PPS, the highest data points were pairs involving mismatches in only one syllable, while the lowest data points were pairs involving mismatches in both syllables. This percentage of 53% was much lower than the approximately 80% that Vitz and Winkler (1973) found for a variety of types of word pairs in English, suggesting that the phoneme in Korean is a less basic representational unit than it is in English.

Figure 2.7. CORRELATION BETWEEN PREDICTED PHONEMIC SIMILARITY AND JUDGED SIMILARITY SCORES (MA group with aural stimuli only, n=39)



2.5.3.3. Group Effects

Since subjects in the three groups varied in their use of the scale (see Table 2.16 above), mean Z-scores for each type were calculated to calibrate the results of the three groups. A syllable effect score (SES) was calculated for each group, as illustrated below:

SES = Mean Z-scores of ((CV.CXX - CX.XVC) + (CV.XXX - XX.XVC))

As indicated in the formula, the SES reflects the extent to which the syllable effect outweighed the phoneme effect. A series of ANOVAs was run on these SES scores with group effects as between-subjects variables. Neither the presentation mode nor the age/education effect was significant.

Another way to test effects of presentation mode and age/education was to conduct two-factor ANOVAs on subject groups and the number of mismatched syllables. Again, neither presentation mode nor age/education effect was significant ($p > .7$ for both), the interactions between the number of mismatched syllables and the group effects were not significant either ($p > .1$ for both).

2.5.4. Discussion

The results reported above suggest that the syllable is the most basic and psychologically the most salient unit in Korean. The results clearly confirmed the Syllable Hypothesis. This hypothesis predicted that, controlling for the number of mismatched phonemes, pairs that had mismatches *within* a syllable should be judged more similar than pairs that had mismatches *across* both syllables. The Phoneme Hypothesis, on the other hand, predicted that there should be no significant difference between the two types of pairs, since both involve the same number of mismatched phonemes. It was found that the syllable effect was highly significant (all at the .001 level). Furthermore, the syllable was a clear winner against the phoneme in all three groups; the phoneme effect was not significant at all in the MA group, where only aural stimuli were provided, and marginally significant in the UB group, where both aural and written stimuli were provided. The difference between the MA group and the other two may suggest that orthographic representations might have led subjects to *count* the number of letters dissimilar between two written words. However, this counting must have been more difficult or impossible for subjects in the MA group, who were not provided with the written word forms.

The correlation used to compare the predictions of the predicted syllabic similarity (PSS) and the predicted phonemic similarity (PPS) produced a similar set of results. In the MA group, the PSS accounted for 70% of the total variance in similarity scores, which was much greater than the 53% explained by the PPS. In the other two groups, both the PSS and the PPS explained about 70% of the total variance. Considering that only two predicted values (i.e., 0 and .5) were used in the PSS as opposed to four predicted values in the PPS, however, the syllable is the more important predictor than the phoneme in the SSJs of Korean words.

The numerical differences (see Table 2.17, in particular) between the MA group and the other two suggests that the presentation mode could be a variable in the SSJ experiments, although the difference did not reach significance. Further research is required on the presentation mode effect on SSJs. However, the effect of age/education was not significant in the ranges tested, judging from quite similar sets of results from both ANOVAs and correlations between the MB and the UB group.

The salience of the Korean syllable, the main outcome of the present SSJ experiment, is comparable to that of Japanese mora, which has been supported by a variety of experimental tasks (e.g., Jaeger, 1980a; Otake, Hatano, Cutler, and Mehler, 1993; Derwing and Wiebe, 1994). In both cases, the basic level of representation is closely related to the orthographic systems, which are based on units larger than the phoneme. In Korean, however, since both the syllable and the phoneme receive orthographic support, the difference between them cannot be attributed solely to orthographic influence.

2.6. Summary of the Main Results

2.6.1. Syllable Structure

The results of Experiments 1 and 2 suggested that, in judging the similarity of pairs of CVC Korean syllables, the CV was more important than the VC. In Experiment 1, CVX pairs (where X indicates a mismatch) were judged significantly more similar than XVC and CXC pairs, while there was no such difference among the pairs that shared only one phoneme. These pattern of results were identical to those found in Yoon and Derwing (1994). Experiment 2 extended the main findings of the above studies in two aspects: (i) by focusing on the CVX and XVC pairs, and (ii) by controlling the types of feature mismatches in the X position. The results of Experiment 2 showed that CVX pairs were always judged more similar than XVC pairs, regardless of the variation of feature mismatches in the X position. Taken together, the above results provide converging evidence that the CV is a more salient than the VC in the SSJs of CVC syllables in Korean.

TABLE 2.18. CROSS-LINGUISTIC COMPARISON OF SYLLABLE STRUCTURE IN THE SSJS OF CVC PAIRS

	Raw Score		% of Varaince Explained	
	CVX	XVC	CV-C	C-VC
English (Nelson and Nelson, 1970)	4.0	(5.3)	53.0	(93.3)
English (Vitz and Winkler, 1973)	2.9	(3.1)	not available	
English (Derwing and Nearey, 1986)	5.6	(6.1)	not available	
Korean (Yoon and Derwing, 1994)	(6.6)	4.9	(79.5)	74.7
Korean (Experiment 1)	(5.7)	4.8	(55.6)	53.5
Korean (Experiment 2)	(3.9)	3.1	(70.6)	54.0

NOTE: In each comparison set, the bigger number is in a parenthesis.

The present results are opposite from those of earlier SSJ studies in English that focused on CVC syllables. Table 2.18 summarizes differential results in the two languages. The most notable feature of the table is that, irrespective of (i) the focus of each study and of (ii) the measurement scales used, the differences between two types were consistently reflected, with a differential prominence depending on the stimulus-language. In both mean similarity score and the amount of variance explained, the results involving the rime in English and those involving the body in Korean were *always* the winner. These consistent results across the two languages and across several experiments strongly suggest the syllable structure effect on the SSJs of monosyllabic pairs and the reliability of the measures used in the SSJ task.

2.6.2. Basic Level

Using all possible subsyllabic comparisons of a CVC syllable, Experiment 1 showed the importance of phonemic segments in predicting the SSJs of Korean. In spite of a differential position effect involving pairs with two-phonemes matched as discussed in the previous section, the correlation between similarity scores and predicted phonemic similarity (PPS) was significant enough to account for most of the variation in similarity scores. In Experiment 3, we tried to extend this finding to disyllabic CV.CVC pairs. As syllable boundaries were involved in these SSJs, however, the role of phoneme was less important than in monosyllabic pairs. Instead, the syllable emerged as a more important predictor of the SSJs.

Table 2.19 below presents a summary of correlations between similarity scores and predicted similarities found in the earlier English studies and the present studies in Korean. In monosyllabic pairs of both languages, the PPS consistently explained about the same high amount of variance in similarity scores, suggesting the importance of phonemic segments in predicting the SSJs of monosyllabic pairs in the two languages. In disyllabic pairs, however, the different pattern of results was found: the PPS in English made as accurate a prediction as in monosyllabic pairs, whereas the PPS in Korean was much less accurate than in monosyllabic pairs. Instead, the syllable-based prediction (PSS) was more accurate than the phoneme-based prediction (PPS), suggesting that the syllable is more basic than the phoneme in disyllabic (and perhaps other polysyllabic) words in Korean. It is interesting to note a parallel results found with the Japanese mora (Derwing and Wiebe, 1994), in which the phoneme played a secondary role to the mora. In sum, therefore, the phoneme's differing role between English *vis a vis* Korean and Japanese cast doubt on the universality of the segment as a basic representational unit; rather, the present studies showed that basic representational units can be language specific. We will discuss this issue in more detail in Chapter IV.

TABLE 2.19. CROSS-LINGUISTIC COMPARISON OF BASIC LEVEL IN THE SSJS

	Monosyllable	Disyllable	
	PPS	PPS	PSS
English (Nelson and Nelson, 1970)	88.4 %		
English (Vitz and Winkler, 1973)	88.4 %	84.6 %	
Korean (Yoon and Derwing, 1994)	86.8 %		
Korean (Experiment 3)		53.3 %	70.6 %

Notes to Chapter II

[1] One of the advantages of the SSJ task pointed out by Derwing and Wiebe (1994: 157) was that it "provides an indirect way to get at phonological units in question, without resorting to the kind of metalinguistic judgments that must be invoked in more direct tasks, such as unit counting."

[2] One obvious counterexample can be found in Vitz and Winkler's own Experiment 2. Although /kræmp/ (PPD = .8) had only one phoneme in common with the standard word /plænt/, it was judged closer to the standard word than was /plɪnt/ (PPD = .2), which shared four phonemes with the standard word. The unexpectedly high similarity rating for the /plænt-kræmp/ pair might be attributable to high similarity of all four of the contrasting consonants, as well as to the importance of the vowel.

[3] Derwing and Nearey (1986) also applied the SSJ technique to test alternative analyses of voiceless stops after /s/ in English, i.e., whether (i) unaspirated [p] in *spill* is identified with /b/ in *bill* in accord with perceptual similarity of the phone [p]; or (ii) as /p/ in *pill* as in the traditional phonemic analyses. It was found that pairs like *spill-pill* were always rated more similar than those like *spill-bill*. Nonsense words used to reduce the orthographic interference also produced a similar result. These results showed that possibly in spite of "phonetic" (physical) reality, subjects' judgments were "phonological" in nature.

[4] More recently, Derwing and Wang (in press) found similar results from the concept formation task, in which the onset concept was easier for Taiwanese speakers to form than the coda concept (see section 3.1.3.3).

[5] Thanks are expressed here to Dr. Sook Whan Cho and to the students in her classes.

[6] Some semi-words, most of which were stems of verbs or adjectives (e.g., *mut, nup*), were included in order to facilitate the tightly controlled comparisons. These semi-words often make it difficult to classify real words and nonwords unambiguously, since semi-words can be considered either as real words in that they carry meaning or as nonwords in that they cannot stand alone (e.g., are not listed in the lexicon as independent words). However, with the similarity attribute specified as *sound* as in the present experiment, the distinction between real and semi-words ought not to be a salient factor in subjects' judgments, since both types are equally plausible in pronunciation (as opposed to some implausible or impossible combinations of consonants and vowels in nonwords.)

[7] These four vowels occur with a frequency of about 60% out of the total of 22 Korean vowels (Lee, 1992; see also Maddieson, 1992). In the Korean orthography, the first two vowels /a/ and /ə/ are written with a vertical stroke and the last two /o/ and /u/ with a horizontal stroke. Thus, the former types of vowels are graphically closer to the onset than to the coda, while the latter types of vowels are located in a position halfway between the onset and the coda (see Figure 4.1). In the present experiment, half of the total of 56 pairs involved the former vowels and the other half involved the latter.

[8] Thanks go to the students in the class and to Dr. Sang Oh Lee.

[9] Earlier versions of this experiment were published as Yoon (1994) and Yoon and Derwing (in press), and were presented at the Alberta Conference on Language, Banff, Alberta, Canada in 1994. The authors thank several participants at the conference for their useful comments.

[10] The location of this break-point has both orthographic and (nonce-word) experimental support (Derwing, 1992a).

[11] Cf. parallel results in Japanese involving the mora unit (Derwing and Wiebe, 1994)

[12] Thanks are expressed here to Ms. Ji Bun Kim, to the students in her classes at N[illegible]san Middle School, to Dr. Sook Whan Cho, and to the students in her class at Sogang University.

CHAPTER III

CONCEPT FORMATION

3.1. Introduction

3.1.1. Definition of the Concept Formation Task

The concept formation (CF) task is a standard experimental technique used in psychological research to test subjects' ability to "manipulate and classify the essential common features embedded in a number of complex stimulus situations" (Vinacke, 1951: 22). That there has been no comprehensive theory of CF is due in part to the difficulty defining the *concept* in the first place. For the purpose of the present discussion, therefore, we adopt the following definition:

> From the point of view of the subject, a concept is a disposition to organize events in a certain way and it implies the expectation that it is *capable of being applied to fresh instances*. In this sense, then, a concept is essentially *predictive*. In so far as a person is open to experiences that bear upon his concepts, these experiences will result in a modification of the concepts [Bolton, 1977: 22; emphasis added].

As its alternative names (e.g., concept attainment, concept learning, concept acquisition) suggest, there are numerous ways to test subjects' ability to form a concept (Bolton, 1977, ch. 2; Deese and Hulse, 1967, ch. 12; Dominowski, 1970; Harris and Harris, 1973). The method I am concerned with is a version of the *identification* method, in which subjects are exposed to one stimulus at a time (Deese and Hulse, 1967). In this method, each stimulus is designated as either a positive or a negative instance of a target concept. Subjects are instructed to respond by saying "yes" or "no" to each instance, and each response is immediately reinforced by the experimenter. Initially, relatively clear positive and negative instances of a target are presented. Later, some unclear instances are presented to see how subjects categorize them. This method is thus better suited to testing the *predictive* status of a concept (as defined in the above quotation) than the method of presenting the entire sample of stimuli at once, for instance. Of primary interest is the ease (e.g., measured by a long succession of correct responses) with which various in[illegible]e categorized to form a concept.

3.1.2. Concept Formation and Linguistic Research

version of the CF technique described above seems well suited to the investigation of a wide range of linguistic concepts (Derwing, 1973, ch. 9; Jaeger, 1980a, ch. 1). Most linguistic theories have assumed implicitly or explicitly the reality of such linguistic units as sentence, word, syllable, phoneme, etc. If real for speakers, such units should find support in speakers' categorical (i.e., sameness or difference) judgments. Lakoff (1981: 142) expresses the point this way:

> Each human language is structured in terms of an enormously complex system of categories of various kinds: phonetic, phonological, morphological, lexical, syntactic, semantic, and pragmatic. Linguistic categories are among the kinds of abstract categories that any adequate theory of the human conceptual system must be able to account for. Linguistics is then an important source of evidence for the nature of cognitive categories. Conversely, *general results concerning the nature of cognitive categorization should apply to categories in linguistics.* [emphasis added]

The CF technique is especially useful in this endeavor, since it can be used to probe subjects' *unconscious* knowledge about their language, as Derwing (1973: 319) pointed out: "One outstanding advantage of the CF paradigm is that subjects are known to be able to derive concepts from common properties even when they are unable to state explicitly what the common properties are." Other advantages of the CF technique are its flexibility and controllability. It is flexible in that it can be used to investigate a wide range of linguistic entities, e.g., phonetic features (Jaeger and Ohala, 1984), the phoneme (Jaeger, 1980a), phonological rules (Wang and Derwing, 1986), sentence types (Baker, Prideaux, and Derwing, 1973), and semantic concepts (Rosch, 1973a). Controllability can also be exercised by varying its procedures (e.g., the degree of explicitness in instructions, depending on the difficulty of concepts subjects are supposed to form) and by the introduction of various types and orders of test and control words, etc.

Although the details of procedures vary greatly in CF experiments depending on the particular concepts being tested, the basic assumption remains the same: the easier it is to categorize on the basis of a concept, the more psychologically salient the concept is. For example, Jaeger and Ohala (1984: 17) assumed that:

> Previously existing or natural categories can be brought to the consciousness of subjects by careful training and feedback techniques... [N]on-existent or unnatural categories should be either extremely difficult or impossible for subjects to learn or manipulate in any consistent way.

Closely related to this assumption are the issues of *basic level* and *prototype* in the theory of natural categorization. A series of experiments by Rosch (1973ab, 1978) found that basic level concepts are easier to learn, faster to recognize, and have shorter names (hence, they are psychologically more salient) than superordinate or subordinate level concepts, and that some (prototypical) members of a category are more representative of the category than other (peripheral) members. Rosch's results provided evidence against the classical theory of categorization, which had been based on fixed relationships and sharp boundaries among categories.[1]

3.1.3. Concept Formation and Phonological Units

In phonology, the two main results of Rosch's experiments, basic level and prototype, were replicated by Jaeger's CF experiments (1980ab, 1986a). Jaeger argued that the CF technique can be useful for investigating all phonological units, since they are all subject to the native speaker's judgments about sameness or difference. Some phonological rules can also be tested using the CF technique on the premise that the forms that undergo some rules (e.g., vowel shift rules) can be differentiated from those that do not (Jaeger, 1980a, ch. 4; Jaeger, 1986b; Wang, 1985; Wang and Derwing, 1986). This section will review briefly the previous CF experiments that have tested phonological units.

3.1.3.1 The Phoneme

Jaeger (1980ab, 1986) and Ohala (1986) applied the CF technique to testing the psychological status of the phoneme in English. Using several versions of the CF technique, Jaeger (1980a) found that the phoneme is a basic level of phonological representation in English. Subjects spontaneously identified various allophones of a phoneme as belonging to one category. Specifically, when the target concept was a word that contained [k^h] (e.g., *kill*), Jeager's subjects extended this concept to [sk]-words (e.g., *skill*) and final [k^o]-words (e.g., *take*), as they formed a category /k/.[2] In her Japanese CF experiments, however, Jaeger (1980a, ch. 3) found that Japanese speakers were much less successful in forming a phoneme concept. Jaeger related this result to orthographic differences in the two languages, which also suggested that a phoneme-sized unit is the basic level of phonological representation for English speakers, whereas it is the subordinate level for Japanese speakers:

	ENGLISH	JAPANESE
SUPERORDINATE LEVEL	word, syllable	word
BASIC LEVEL	phoneme	syllable (or mora)
SUBORDINATE LEVEL	feature	phoneme, feature

3.1.3.2. The Feature

Jaeger (1980a) and Jaeger and Ohala (1984) tested the conceptual saliency of features. In both studies, five consonantal features in English were tested: [+anterior], [-anteror], [+voice], [+sonorant], and [+continuant]. The concepts that subjects learned were "words beginning with a phoneme that has one of the above five features". Overall, the concept of feature was found much more difficult to form than the concept of phoneme. The results also provided evidence that some members of a category are more prototypical than others. In the category [+anterior], for instance, bilabial consonants were more frequently included than alveolar consonants. The status of /w/ was ambiguous, suggesting that the internal structure of the feature [±anterior] is not strictly binary as commonly assumed in generative phonology.[3]

3.1.3.3. Subsyllabic Units

Derwing and Wang (in press) have recently applied the CF technique to the investigation of Taiwanese subsyllabic units. Other experimental techniques applied to Taiwanese such as Word-Blending and Sound Similarity Judgments (Wang and Derwing, 1993) have suggested that the Taiwanese syllable does not have a well-defined internal structure. The question addressed in Derwing and Wang's CF experiments was whether a concept defined in terms of some subsyllabic units (e.g., /p/ or /pa/ in a syllable /pan/) was easier to form than other concepts (e.g., /an/ or /n/ in /pan/). Comparisons among these concepts showed that the onset concept was easier to form than the coda concept by various response measures. This was consistent with Wang and Derwing's (1993) finding from sound similarity judgments in Taiwanese, in which the onset was more salient than the coda.

However, none of these subsyllabic concepts was easy to form, suggesting that the syllable may be the smallest viable phonological unit in Taiwanese for speakers not well-versed in alphabetic orthography (see also Read, Zhang, Nie, and Ding, 1986). In light of Jaeger's (1980a) argument about the basic level, therefore, units smaller than a syllable seem to be subordinate levels of phonological representation in Taiwanese, if they exist as units at all.

3.1.4. Overview of the Chapter and Goals of the Experiments

The organization of the present chapter is as follows. The next section, 3.2, discusses the methodology used in the previous English CF experiments and describes some new modifications used in the present experiments. Section 3.3 discusses a practice task, which illustrates the methodological details of the present CF experiments in Korean. Another goal of the practice task was to check whether the two groups of subjects in the following main experiments had about equal ability to perform a CF-type task.

Sections 3.4 and 3.5, present four CF experiments that were designed to compare the ease with which Korean speakers can learn phonological concepts. Each subject was taught one of the four targets that were defined in terms of a common CV, VC, C or CVC in a CVC syllable. Section 3.4 pits the CV of a CVC syllable (or the body: Experiment

4A) against the VC of a CVC syllable (or the rime: Experiment 4B). Subjects in Experiment 4A had to learn that "words containing /ka/" was the target, while subjects in Experiment 4B had to learn that "words containing /ak/" was the target. The primary goal of Experiment 4, therefore, was to test whether there was a difference in performance between the two experimental groups. If the body is a more salient subsyllabic unit than the rime in Korean, subjects in Experiment 4A should find the task easier than subjects in Experiment 4B.

Section 3.5 (Experiment 5) compares the C unit (the phoneme) vs. the CVC (the whole syllable) in a way similar to Experiment 4. Experiment 5A was designed to parallel the previous /k/ experiments in English (Jaeger, 1980a: ch. 4, *inter alia*), in order to test whether Korean speakers could form a concept defined in terms of a common phoneme /k/. Experiment 5B, in which the target concept was a common syllable /kak/, was designed to be otherwise as similar as possible to Experiment 5A. The results are discussed in light of the notion of basic level of phonological representation in Korean.

Finally, section 3.6 summarizes the results of Experiments 4A-5B and compares the four phonological concepts. The present CF experimental technique will also be assessed.

3.2. Previous Methodology and Modifications for the Present Experiments

3.2.1. Three Types of Exemplar Words

In CF experiments, subjects are provided with positive (target), negative (distractor), and ambiguous (test) words. English examples used to illustrate these three types of words are taken from Jaeger (1986a).

3.2.1.1. Targets

Targets are the examples of the particular concept or category that subjects are trained to form. In Jaeger (1986a), for example, the concept was the phoneme /k/ in English, and targets were words containing the allophone $[k^h]$ (e.g., *kind*). Various orthographic variants of target words were included as well, such as *clear*, *chrome*, *acclaim*, *queen*, etc. Subjects' responses to these spelling variants provided information about orthographic interference in the formation of the concept of /k/.

3.2.1.2. Distractors

Distractors are words that lack the defining property of the target words. There are various types of distractors, which cause different kinds of interference with target words: no interference (e.g., *left*), orthographic interference (*knit*), and phonetic interference (*gift*), where a similar but contrasting category is involved (in this case, the phoneme /g/).

In designing a CF experiment, care should be taken to control the types of targets and distractors, because the relative difficulty of learning a concept can vary as a function of the degree of information that targets and distractors provide for subjects. For example, a larger variety of targets (e.g., $[k^h]$ with various spellings) will render the task more difficult than only one type of target (e.g., $[k^h]$ spelled with <k>). Likewise, distractors are inherently different in their degree of interference with targets; distractors containing no interference are easier to eliminate than those involving some kind of interference. However, non-interfering distractors also provide less information about the nature of the target category. When a series of concepts is being compared, therefore, it is important to control both target types and distractor types.

3.2.1.3. Test Words

Test words are words whose category membership is in question. In Jaeger's /k/ experiment, for example, two types of allophonic variants were tested: unaspirated (*s<u>k</u>in*, *s<u>c</u>are*) and unreleased (*boo<u>k</u>shelf*, *e<u>x</u>cept*). The first of these, in particular, i.e., the nature of unaspirated stops after /s/, has long been a matter of theoretical dispute in the phonetic and phonological literature (e.g., Derwing, 1973; Ladefoged, 1975a; Lisker and Abramson, 1964; Lotz, Abramson, Gerstman, Ingemann, and Nemser, 1960). Since Jaeger's subjects included the above words in their target class, however, her results indicated that subjects' categorization was in accord with a traditional phonemic analysis.

3.2.2. Four Components of the CF Experiments

CF experiments typically involve the following four components, ordered as indicated: instructions →learning session → test session→ post-experimental interview. The learning and test sessions are the major components. The learning session involves a *basic* stimulus set, which measures subjects' ability to classify targets and distractors (i.e., it provides a measure of the general difficulty of the task), while the test session is a *supplementary* set, in which the status of ambiguous and controversial stimuli is evaluated (i.e., whether or not subjects classify these stimuli as targets). Thus, the comparisons of the ease with which various concepts are formed are limited to the results of the learning session.

3.2.2.1. Instructions

The instructions are designed to explain the basic nature of the task, as well as to help direct subjects' attention to the relevant attributes that define the target class. Care must be taken not to provide too much or too little information. Generally, the degree of explicitness depends on the difficulty of the target. For example, if the target is a word beginning with a [+labial] phoneme, the instruction, "the common property has to do with *the first sound of these words*" ought to help subjects focus on relevant attributes of the target. The instruction, "the first sound of the target is *pronounced at the lips*", however, would give too much information, while "the common property has to do with *the way these words sound* to you" would likely provide too little information, allowing subjects to be distracted by any number of salient phonetic properties. On the other hand, if the target is a word containing a [k^h] sound, "the common property has to do with *the way these words sound* to you" would probably be enough for subjects to focus on the relevant attributes, while excluding semantic similarities or other possible distractors.

Care should also be taken to keep the instructions consistent across all of the target categories tested, if the main purpose is to *compare* the psychological status of one category against another. In the present experiments, including the practice sessions, the instructions were thus kept identical throughout.

3.2.2.2. Learning Session

The purpose of learning session is to gradually teach a concept to subjects by reinforcing their responses. Initially, relatively clear examples of targets and distractors (i.e., most typical targets and least interfering distractors) are presented, because subjects will have no idea of how to distinguish targets from distractors. Later, a larger variety of targets and distractors is presented to see what types of targets are representative and what types of interference are significant. Subjects respond to each word by saying "yes" or "no".

In the present experimental design, the learning session is the most important part for two reasons. First, it will provide most of the useful response measures, e.g., the number of subjects reaching criterion used to define the mastery of each concept; and the number of trials each subject needed to reach criterion. The criterion for mastery for this practice task and all of the present experiments was set at 15 consecutive trials with two or fewer errors. This criterion was chosen because it allows some room for accidental errors while remaining extremely difficult to reach by chance ($p < .01$).[4]

The learning session also provides a third response measure, the total number of correct responses for each concept. This third response measure has an advantage over the previous two, the number of subjects reaching criterion and the trials to criterion, which can both vary depending on how the mastery criterion is set. However, the total number of correct responses is independent of any semi-arbitrary choice of a mastery criterion. Therefore, it provides a useful new source of information for between-experiment comparisons.

3.2.2.3. Test Session

The main purpose of the test session is to observe subjects' categorization of various ambiguous or theoretically controversial stimuli (i.e., test words). Unlike in the learning session, no feedback is provided about whether subjects' responses are correct. Separate instructions are thus necessary before the beginning of the test session, and the test words are randomly intermixed with types of targets and distractors used in the learning session.

Since the primary purpose of the present CF experiments in Korean is to compare the relative ease with which subjects form various phonological concepts, the test session is limited to Experiment 5A, where various spelling variants of /k/ and phonemic variants of spelling "ㄱ" are used as test words to see whether subjects' judgments were phoneme-based or spelling-based. In the present series of CF experiments, then, the test session was a supplementary part of the experiment.

3.2.2.4. Post-experimental Interview

In the post-experimental interview, subjects are asked to name the common property that had guided their responses and to explain their decision-making strategy. Although subjects may form a concept without being able to name it (Chafe, 1977), there ought to be a correlation between subjects' naming ability and the ease with which they form a concept (i.e., codability: Deese and Hulse, 1967).

3.2.3. Three Response Measures

A variety of response measures can be used to test the ease with which a concept is formed. The present CF experiments focus on the following three quantitative measures, as outlined above: (i) the number of subjects reaching the mastery criterion; (ii) the number of trials to reach criterion; and (iii) the total number of correct responses for each task. While the first two have been commonly used in the previous CF experiments, the last response measure is relatively new. Since the same number of targets and distractors is used across all of the present experiments, the total number of correct responses ought to provide a reasonable between-experiment comparison about overall difficulty of each task.

In addition to the above three, subjects' ability to name the category (i.e., codability) in the post-experimental interview and responses to test words will also be discussed. Reaction time will not be used, because the present experiments were designed to be group administered. Furthermore, as Jaeger (1980a) pointed out, an interpretation of response times in CF experiments is problematic, due both to great individual variation and to inherent differences involved in positive vs. negative responses.[5]

3.2.3.1. Number of Subjects Reaching Criterion

Data on the number of subjects who can form a concept in the learning session provide information about the difficulty of a concept. For example, Jaeger (1980a, ch. 3) reported that all of the English subjects could form the concept of English /k/, while only half of the Japanese subjects could form the concept of Japanese /t/, indicating that the English task was an easier one.

3.2.3.2. Trials to Criterion

Generally, the more difficult a concept is, the more trials are necessary to form it. For example, Ohala (1986) compared the ease of forming two target concepts in terms of the number of trials required to reach criterion for a list of 75 words: the target for Group I (n = 10) was a word that contained [k^h] or [sk], and the target for Group II (n = 10) was a word that contained [g] or [sk]. The results were consistent with the results from Jaeger's experiments in that subjects in Group I formed the concept much more easily than those in Group II, as summarized below:

TABLE 3.1. RESULTS OF OHALA'S CF EXPERIMENTS (1986: 21)

Response Measure	Group I	Group II
Trials to criterion (range: 15-75)	22.67	45.38
Subjects reaching criterion (n = 10)	9	8

As can be seen in Table 3.1, both response measures showed that the stop in [sk] was easier to be categorized with [k^h] than with [g], especially as measured by number of trials to criterion. Group I required only about 23 trials to reach criterion (15 correct trials in a row with two or fewer errors). This was extremely a small number of trials, given that initially subjects had to resort to random guessing and that later they had to answer correctly to almost all of the words. This was not only because the concept was very easy to form but because the mean number of trials was calculated only on the basis of subjects who reached criterion.

It might also have been informative, however, if the results from the subjects who did not reach criterion had been included in calculating the mean number of trials to criterion. Especially when a target is relatively difficult for subjects to form and only a small percentage of subjects can reach criterion, more information would be provided by calculating a mean number of trials for *all* subjects. One way of doing this is to set the maximum number of trials to 'total number of stimuli + 1' for those subjects who have not formed the concept by the end of the learning session. For the analysis of the present experiments, therefore, the maximum number of trials necessary to reach criterion is set at 61, viz., for subjects who did not reach criterion by the end of the 60-stimulus learning session.

3.2.3.3. Total Number of Correct Responses

The total number of correct responses equals the number of "yes" responses to targets plus the number of "no" responses to distractors. Calculated on the basis of the learning session only, this number provides useful information about category membership for all stimuli except test words. The more correct responses each subclass of stimuli receives, the more prototypical that subclass is of the category. For example, Jaeger (1986a) found that some stimuli involving orthographic interference with the /k/ concept, such as target words spelled with the letter <x> (e.g., *except*) or distractors spelled with "silent k", yielded a much higher error rate than target words spelled with <k> or <c> (e.g., *kind* or *clear*) or distractors which did not involve orthographic interference at all (e.g., *left*).

3.3. Practice Task

3.3.1. Goals

The most important aspect of the present CF technique was to have a maximally similar design across all present experiments (from 4A to 5B). As a sample experiment, the practice task introduces the common methods used in the main experiments. The main goal of the practice task, however, was to provide the basis for later comparisons in Experiment 4, that is, to check whether the performance of subjects in Experiment 4A and 4B were equivalent on the common practice task. The homogeneity of these two groups of subiects is important, in order to ensure that the results of Experiments 4A and 4B did not vary as a function of individual subjects' general ability to perform a CF task.

3.3.2. Subjects

The subjects were 40 university students taking an introductory English composition class at Sogang University, Seoul, Korea. Two different classes were represented.[6] The first class (n = 18) did Experiment 4A and the second (n = 22) did Experiment 4B, both immediately after the practice session. All of the subjects were native speakers of Korean and participated in the experiment on a voluntary basis.

3.3.3. Stimuli

The general methods to be introduced in this section are essentially the same as those used in all present experiments from 4A to 5B. (Differences were in some details of the stimuli specific to each task, of course.) There was a total of 60 real word stimuli, of which half were targets and the other half distractors. Examples of each stimulus type are presented in Table 3.2:

TABLE 3.2. EXAMPLES OF STIMULI USED IN THE PRACTICE SESSION

Targets		Distractors	
Types	Examples	Types	Examples
a-a	paŋ.hak	a-u	san.pul
Ca-Ca	pan.pal	u-a	cʰul.pal
aC-aC	maŋ.saŋ	u-u	cuŋ.sun

In this practice session, all 60 stimuli were two-syllable (CVC.CVC) real words, and the distinguishing property between targets and distractors was vowel-based. The targets were defined as words containing the vowel /a/ in both syllables, while the distractors contained the vowel /u/ in one or both syllables. Both targets and distractors were divided into three sub-types consisting of 10 tokens each, as shown in Table 3.2. (See Appendix 4 for the full list of stimuli used in the practice session.)

Another important aspect of controlling the stimulus presentation was to maintain the same increasing order of difficulty across all experiments. This was done by preparing three "basic sub-lists" of stimuli for each experiment. The first sub-list (from words 1 to 12) was composed of the smallest number of stimulus types, and the last list (from words 37 to 60) had the greatest variation in stimulus types. In the practice task, the stimuli were presented in the following order:

	TARGETS		DISTRACTORS
Words 1-12:	3 Ca-Ca, 3 aC-aC	mixed with	6 u-u
Words 13-36:	4 Ca-Ca, 4 aC-aC, 4 a-a	mixed with	6 a-u, 6 u-a
Words 37-60:	3 Ca-Ca, 3 aC-aC, 6 a-a	mixed with	4 u-u, 4 a-u, 4 u-a

Still another control was built into the reinforcement schedule, viz., the order of individual targets and distractors. In the practice task as well as the following main experiments, a common reinforcement schedule was maintained across all groups (e.g., "T T F T F F T F T T F F, " etc., where (T = target; F = distractor).

3.3.4. Procedures

3.3.4.1. Instructions

The instructions used in the present practice task and the following main experiments were as follows (presented in Korean):

> You are going to play a word game that will help us better understand how our native language Korean works. This game is very similar to the game called ***simukokæ*** that most of you already know. Both games are similar in that feedback to "yes" or "no" responses will guide you to the correct answer.
>
> In this game you will be listening to a tape that contains a series of spoken Korean words. Some of these words have some property in common, while the other words do not. *The property has to do with a certain sound that these words have in common.* Your job is to figure out what this common property is. On your answer sheet there are "yes" and "no" typewritten alongside each item number. I ask you to respond to each word by circling "yes" if you think the word has the common property, or "no" if it does not. After about five seconds, the voice on the tape will say "yes" if a word has the common property, or "no" if it does not. You can determine whether or not your response is correct by checking your response against the response from the tape: the match

between the two indicates that your response was correct, while the mismatch indicates that your response was incorrect.

After the game, I will ask you to try to name the common property you had in mind and to explain your decision-making strategy. Are there any questions?

3.3.4.2. Learning Session

Targets and distractors were semi-randomly ordered to keep the increasing degree of difficulty toward the end of the list, as discussed in section 3.3.3. The stimuli were played back to subjects roughly as follows:

FIGURE 3.1. THE SAMPLE OF A SINGLE STIMULUS WORD PRESENTATION

TIME → [1 sec.] [5 sec.] [2 sec.]

TAPE: "Number 1" ------ stimulus ------------- "yes" or "no" ------ "Number 2"

SUBJECT: ____________________"yes" or "no"__________________________

The time taken for each stimulus word was about 10 seconds and the whole practice task including the instructions took about 12 minutes. "Number ..." was pronounced before each stimulus word to get subjects ready for the next test item, as well as to help subjects keep track of item numbers on their answer sheets.

3.3.4.3. Post-experimental Comments

After the practice task, subjects were asked to write on the answer sheet the common property that had guided their responses and to explain their decision-making strategy.

3.3.5. Results

Correct responses were converted to 1 and incorrect or no responses to 0.[7] Since, unlike the case of oral responses (e.g., Derwing and Wang, in press), there was virtually no instance of "no response," the data from all 40 subjects were analyzed.

Overall, 23 subjects (57.5%) reached the mastery criterion (i.e., provided 15 consecutive answers with two or fewer errors). For these subjects, trials to criterion were set at the end point of the first qualifying string of 15 consecutive stimuli. For the remaining 17 subjects, an arbitrary 61 (the total number of stimuli + 1) was assigned. Table 3.3 shows the results:

TABLE 3.3. RESULTS OF THE PRACTICE SESSION (n = 40)

Response Measure	Group A (n = 18)	Group B (n = 22)
Subjects reaching criterion	11	12
	(61.8%)	(54.5%)
Number of correct responses	43.6	40.4
	(72.7%)	(67.3%)
Trials to criterion (15-61)	38.9	43.0

All three response measures showed that there was no significant difference in the performance on the practice task between the two subject groups. First, about equal proportions of the subjects reached criterion in the two groups. Second, the total number was also not significantly different between the two groups. Trials to criterion were not significantly different either.

Furthermore, the overall 70% correct response rate suggested that the concept of "two-syllable words containing the vowel /a/ in both syllables" was not difficult to learn; the total number of correct responses was significantly larger than the 30 to be expected by chance ($t[39] = 8.48$, $p < .001$). A quite respectable average trials to criterion of 26.5 ($sd = 10.1$) for subjects who reached criterion also suggested that the concept was relatively easy to form.

The most important result of the practice session was that the average performance between the two groups was about equal. Furthermore, performance was significantly better than chance, indicating that subjects generally understood the task. As noted below, Group A's subjects did main Experiment 4A: /ka/, and Group B's subjects did main Experiment 4B: /ak/, both after the practice session.

3.4. Experiment 4

3.4.1. Goals

We found in Experiment 1 that CVC pairs sharing the first two phonemes (CV) were judged more similar than those sharing the last two phonemes (VC) or the first and the last (see also Yoon, 1994a; Yoon and Derwing, 1994). This result was replicated in Experiment 2 with a variety of feature mismatches in the mismatched phonemes (e.g., the last phoneme in CV-sharing pairs). Opposite results have been found in English SSJ experiments (Nelson and Nelson, 1970; Vitz and Winkler, 1973; Derwing and Nearey, 1986), reflecting different internal structures of the syllable in the two languages. A similar set of results was also found in Word Blending experiments (for English, see Treiman, 1983, 1986; for Korean, see Derwing, Yoon, and Cho, 1993; Yoon, 1994a). Evidence from the above two experimental paradigms lent strong support to the idea that the Korean syllable is left-branching, with the body as a major subsyllabic unit.

The primary goal of Experiment 4 was cross-methodological validation. It was designed to test the conceptual saliency of the body as a major subsyllabic unit in Korean *vis-a-vis* the rime. The basic design was similar to that of the detection paradigm (Cutler, Mehler, Norris, and Segui, 1986; Mehler, Dommergues, Frauenfelder, and Segui, 1981;

Otake, Hatano, Cutler, and Mehler, 1993), in which response times were compared between stimuli that contained a target matched with the hypothesized phonological units and those that did not. In the Otake et al study, for example, the syllable hypothesis was disconfirmed in favor of the mora hypothesis in Japanese.

In Experiment 4, about half of the subjects learned the concept of "two-syllable words containing /ka/", and the other half "two-syllable words containing /ak/." It was hypothesized that the first group of subjects would find the task easier. The rationale behind this hypothesis was that, if in Korean the division of the CVC syllable into CV and C is more natural than into C and VC, the targets defined in terms of common CV's ought to be more readily mastered than those defined in terms of common VC's.

3.4.2. Subjects

A total of 85 subjects were drawn from four introductory English Composition classes at Sogang University.[8] Each class was randomly assigned to one of the four groups, as follows:

TABLE 3.4. ASSIGNMENT OF SUBJECTS IN EXPERIMENT 4 (n = 85)

Experimental group	Subject group[9]	Task(s)
4A: /ka/ (n = 44)	AI (n = 18)	Practice & Experiment 4A
	AII (n = 26)	Experiment 4A
4B: /ak/ (n = 41)	BI (n = 22)	Practice & Experiment 4B
	BII (n = 19)	Experiment 4B

3.4.3. Stimuli

All stimuli were two-syllable real words as in the practice session, and were designed to be maximally similar between Experiments 4A and 4B. In addition to the controls already discussed in the previous sections, care was taken to keep all other possible variables as similar as possible between the two experiments. First, targets were divided into two types, depending on the syllable in which /ka/ (Experiment 4A) or /ak/ (Experiment 4B) was located, as shown in Table 3.5 below.

TABLE 3.5. EXAMPLES OF TARGET WORDS USED IN EXPERIMENT 4

	Initial Syllable		Final Syllable	
Experiment	Type	Example	Type	Example
4A: /ka/	kak.-	kak.pyəl	-.kak	sim.kak
	ka-.-	kaŋ.co	-.ka-	sæŋ.kaŋ
4B: /ak/	kak.-	kak.pyəl	-.kak	sim.kak
	-ak.-	cak.sim	-.-ak	si.cak

NOTE: "." indicates a syllable boundary.

There were a total of 30 targets in both experiments. The 10 targets containing the syllable /kak/ either in the initial or in the final syllable were common in the two experiments. The other 20 target words were specific to each group, and the target-bearing syllable was always CVC. Thus the target did not generally comprise a whole syllable but only a syllable sub-part, so no whole syllable analysis was possible for either /ka/ or /ak/.

As presented in Table 3.6, distractors were also designed in a similar way between the two experiments. There were three types of distractors, depending on the number of interfering phonemes (i.e., /k/ or /a/ in this case). The first type was the least interfering type D0, which contained neither /k/ nor /a/; the second type was D1, containing an onset /k/ or a coda /k/; and the third type was D2, containing the non-target sequence /ak/ or /ka/.

TABLE 3.6. EXAMPLES OF DISTRACTORS USED IN EXPERIMENT 4

	D0	D1 (one-phoneme)		D2 (two-phoneme)	
Experiment	Example	Type	Example	Type	Example
4A: /ka/	toŋ.yo	k-	<u>k</u>oŋ.toŋ	-ak	c<u>ak</u>.sim
4B: /ak/	toŋ.yo	-k	ci<u>k</u>.cəp	ka-	<u>ka</u>ŋ.co

Just as the 10 /kak/ targets were common in both Experiments 4A and 4B, so were 10 D0 distractors. These distractors were expected to interfere least with the concepts that subjects were supposed to learn, since they did not contain either /k/ or /a/. (See Appendices 5 and 6 for a full list of the stimuli used in Experiments 4A and 4B.)

The order of stimulus presentation was designed to gradually teach the concept "words containing the body /ka/" for Experiment 4A and "words containing the rime /ak/" for Experiment 4B, and to maintain the same increasing degree of difficulty between the two experiments. The same reinforcement schedule as in the practice task was also used. The three basic sub-lists used were parallel to those in the practice session. (There were no test sessions in Experiments 4A and 4B.)

	TARGETS		DISTRACTORS
Words 1-12:	3 TI, 3 TF words	mixed with	6 D0
Words 13-36:	6 TI, 6 TF words	mixed with	6 D1, 6 D2
Words 37-60:	6 TI, 6 TF words	mixed with	4 D0, 4 D1, 4 D2

3.4.4. Results

3.4.4.1. Experiment 4A

3.4.4.1.1. Overall results

A total of 44 subjects were divided into two groups. Group I subjects performed both the practice and the main experiment, while Group II subjects performed the main experiment only.[9] Table 3.7 presents the results of the main experiment in both groups:

TABLE 3.7. RESULTS OF EXPERIMENT 4A (n = 44)

Response Measure	Group AI (n = 18)	Group AII (n = 26)
Subjects reaching criterion	11	15
	(61.8%)	(57.7%)
Number of correct responses	46.3	43.2
	(77.2%)	(72.0%)
Trials to criterion (15-61)	36.4	39.6

In the above results, a practice effect was not indicated. Although the numerical differences slightly favored Group AI, where subjects performed the practice task, there were no significant differences in subjects' performance between the two groups. (Neither the total number of correct answers nor trials to criterion reached significance at the .05 level.)

3.4.4.1.2. Responses to each stimulus type

The responses to each type of targets and distractors are given in Table 3.8 below. These figures include only responses made at and after (but not before) reaching criterion.

TABLE 3.8. NUMBER AND PERCENT CORRECT RESPONSES TO TYPES OF TARGETS AND DISTRACTORS IN EXPERIMENT 4A (Based on 26 subjects who reached criterion)

	Type	# Correct	# Total	% Correct
Targets	kak	228	239	95.4
	ka-.-	211	227	93.0
	-.ka-	211	232	90.9
	Sum	650	698	93.1
Distractors	∅ (D0)	195	205	95.1
	k- (D1)	202	237	85.2
	-ak (D2)	228	244	93.4
	Sum	625	686	91.1

The overall percent of correct responses for the targets (93.1%) and distractors (91.1%) were not significantly different. There were also no significant differences among the three target types. In the distractor types, however, significant differences were found ($F_1[2,50] = 7.27$; $F_2[2,27] = 6.76$, $p < .005$ for both). *T* tests showed that the D1 type yielded significantly more errors than either D0 ($t_1[25] = 2.79$, $p < .01$; $t_2[9] = 2.87$, $p < .05$) or D2 ($t_1[25] = 2.96$, $p < .01$; $t_2[9] = 2.58$, $p < .05$). These results from the distractor types seem to suggest that, in forming a concept of "words containing the body /ka/," subjects paid more attention to the onset /k/ than to the vowel /a/, and that the occurrence of /k/ in the coda did not cause as much interference with the target /ka/. This result is reminiscent of the sound similarity judgment experiment in Yoon and Derwing (1994), who showed that two CVC Korean words sharing the same consonant but in different positions (e.g., /pin-mup/) were judged to be just as dissimilar as those sharing no consonants (e.g., /pin-mut/), but much less similar than those sharing a consonant in the onset (e.g., /pin-put/.

3.4.4.1.3. Subjects' comments

Finally, subjects' comments after the experiment were analyzed. Out of 26 subjects who reached criterion, 21 mentioned "ka" in various ways. Of these 21 subjects, eight commented on pronunciation (e.g., "the pronunciation /ka/ is in either syllable of the words"), and four specifically pointed out the spelling (e.g., "the words contain the spelling 가 <ka>"), and the remaining nine did not specify either pronunciation or the spelling but simply included "ka" in their comments. Interestingly, four subjects (three of whom were commenting on pronunciation) pointed out that the syllable in question should contain *patchim* (i.e., a coda) after /ka/. Only one subject commented that the words should have a combination of ㄱ <k> and ㅏ <a>.

Most of the subjects who could not reach criterion made no comments. Two comments from these subjects "the words have the initial /k/," indicating that they were forming a (wrong) concept of onset. But there were no instances of forming a concept of /a/. This is presumably related to the fact that the D1 distractor containing the onset /k/ caused more interference than the D2 distractor containing the rime /ak/. (See section 3.4.4.1.2 above)

3.4.4.2. Experiment 4B

3.4.4.2.1. Overall results

A total of 41 subjects were divided into two groups. As in Experiment 4A, the first group of subjects (n=22) performed both the practice and the main experiment, while the second group of subjects (n=19) performed the main experiment only. Table 3.9 below presents the results of the main experiment for both groups. Again, no practice effect was indicated, as in Experiment 4A. Except for a slightly bigger difference between the number of subjects reaching criterion, both trials to criterion and number of correct responses showed that there was no difference in performance between the two groups ($p > .1$ for all). Although the total number of correct responses (38.1; sd = 7.4) was smaller than that in Experiment 4A (44.5; sd = 11.5), it was still significantly larger than the 30 which was to be expected by chance ($t_1[40] = 7.03$, $p < .001$).

TABLE 3.9. RESULTS OF EXPERIMENT 4B (n = 41)

Response Measure	Group BI (n = 22)	Group BII (n = 19)
Subjects reaching criterion	8	5
	(36.4%)	(26.3%)
Number of correct responses	39.0	37.2
	(65.0%)	(62.0%)
Trials to criterion (15-61)	47.5	55.1

3.4.4.2.2. Responses to each stimulus type

The responses to each type of targets and distractors are given in Table 3.10 below. As in Experiment 4A, only responses made at and after (but not before) reaching criterion were included. In Table 3.10, the total number of responses was smaller than in Experiment 4A, since fewer number of subjects reached criterion.

Even including only the responses made at or after reaching criterion, overall proportion of correct answers (83.4%) was smaller than that of Experiment 4A (92.1%). However, many similar trends were found between Experiments 4A and 4B. First, as in Experiment 4A, correct responses for the targets and distractors were not significantly different. There were also no significant differences among the three target types. However, there were significant differences among the three distractor types ($F_1[2,24] = 6.16$, $p < .005$; $F_2[2,27] = 31.62$, $p < .001$). T tests showed that the D0 type containing neither /a/ nor /k/ had yielded more correct responses than either D1 ($t_1[12] = 2.96$, $p < .01$; $t_2[9] = 4.71$, $p < .001$) or D2 ($t_1[12] = 3.45$, $p < .005$; $t_2[9] = 7.29$, $p < .001$).

TABLE 3.10. NUMBER AND PERCENT CORRECT RESPONSES TO TYPES OF TARGETS AND DISTRACTORS IN EXPERIMENT 4B (Based on 13 subjects who reached criterion)

	Type	# Correct	# Total	% Correct
Targets	kak	80	94	85.1
	-ak.-	87	98	88.8
	-.-ak	90	100	90.0
	Sum	257	292	88.0
Distractors	∅ (D0)	80	82	97.6
	-k (D1)	88	108	81.5
	ka- (D2)	68	109	62.4
	Sum	236	299	78.9

3.4.4.2.3. Subjects' comments

Overall, subjects' comments on the target /ak/ in this Experiment 4B varied more than those on the target /ka/ in Experiment 4A. Out of 13 subjects who reached criterion, seven correctly commented on the defining property of the target words. Three subjects mentioned "the words have the spelling 악" (/ak/ with the phonologically null symbol "ㅇ", which is required to write a vowel-initial syllable), while the other four said that the words had a combination of the vowel /a/ plus the consonant /k/.

Of the 28 subjects who could not reach criterion, three subjects commented on the vowel /a/, two on the coda /k/, and one on the consonant /k/. These comments reflected different interference types with forming the concept /ak/. Compared to Experiment 4A, where little inference seemed to be caused by the vowel /a/ with the target /ka/, these comments revealed that the distractors containing the vowei caused more errors in forming the rime concept in Experiment 4B than in forming the body concept in Experiment 4A.

3.4.4.3. Comparison between Experiments 4A and 4B

As discussed above, no practice effect was found in either Experiments 4A or 4B. A series of two-factor analyses of variance was also conducted on the average trials to criterion and on the total number of correct responses in the main experiments. In these analyses, practice effect (i.e., between the two groups in either experiment) and experimental effect (i.e., between the two target groups) were treated as independent variables. Both in the total number of correct responses and trials to criterion, the experimental effect was significant ($F_1[1,81] = 9.74$, $p < .005$ for the total number of correct answers; and $F_1[1,81] = 10.11$, $p < .005$ for trials to criterion). The practice effect did not reach significance at the .05 level in either experiment by either response measure, nor was there a significant interaction between the two effects.

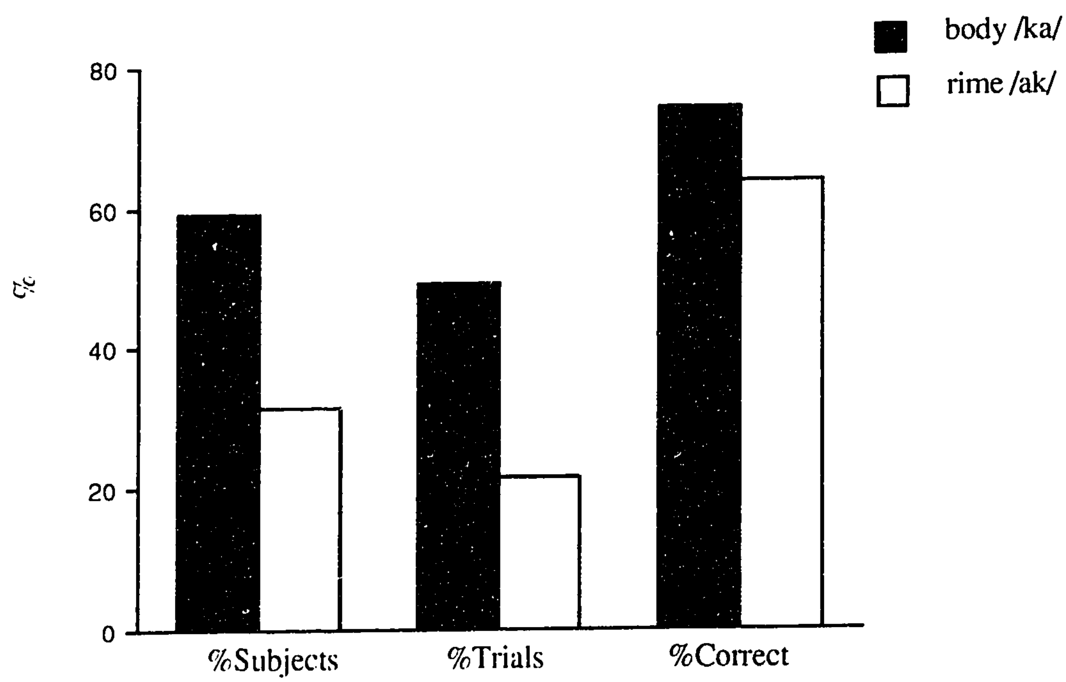
The results were thus pooled across the two groups in each experiment. Table 3.11 shows the pooled results from both experiments.

TABLE 3.11. RESULTS OF EXPERIMENTS 4A AND 4B (n = 85; Two groups within both Experiments combined)

Response Measure	4A: /ka/ (n = 44)	4B: /ak/ (n = 41)
Subjects reaching criterion	26	13
	(59.1%)	(31.7%)
Number of correct responses	44.5	38.1
	(74.2%)	(63.5%)
Trials to criterion (15-61)	38.3	51.0

These results are also graphed in Figure 3.2 below, using a percentage "success rate" measure.

FIGURE 3.2. DIFFERENCE BETWEEN THE BODY AND THE RIME IN TERMS OF "SUCCESS RATE" * OF THE THREE RESPONSE MEASURES



* For the ease of comparison, trials to criterion were converted to percentage by the formula, "(61-trials to criterion) × 100/46", in which the earliest possible trial 15 is converted to 100%, and the 61 (for a subject who did not reach criterion) to 0%.

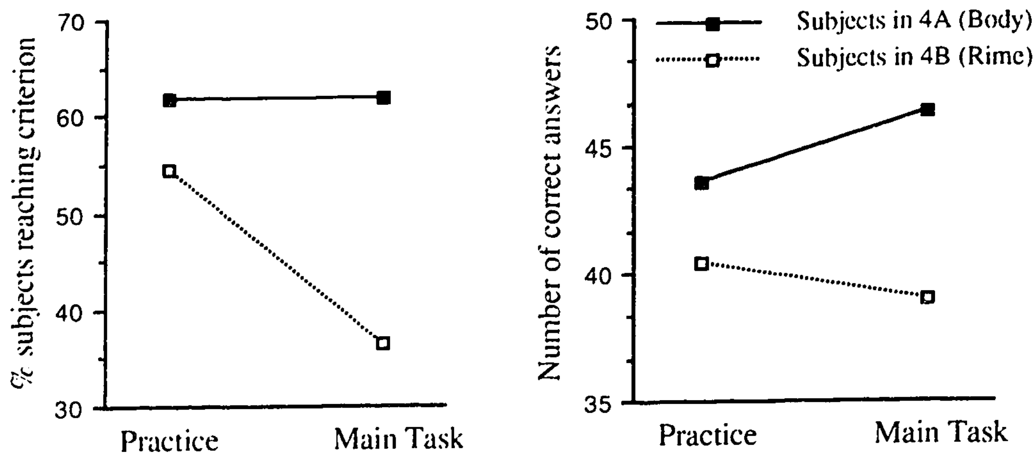
As can be seen in Table 3.11 and Figure 3.2, subjects in Experiment 4A found the task much easier than those in Experiment 4B. Not only did a greater percentage of subjects in Experiment 4A reach criterion (59.1% vs. 31.7%), the other two response measures also pointed in the same direction: significantly fewer trials were required to learn the concept (38.3 vs. 51.0; $F_1[1,81] = 10.11, p < .005$), and the total number of correct responses was significantly higher (44.5 vs. 38.1; $F_1[1,81] = 9.74, p < .005$). Ignoring the subjects who did not reach criterion, the average trials to criterion in Experiment 4A was a quite impressive 22.6 (sd = 11.5), which is comparable to the results of an earlier series of English /k/ CF experiments (see, for example, section 3.2.3.2 for the results reported in Ohala, 1986). Furthermore, the total number of 44.5 correct responses, which amounts 74.2 % correct overall was significantly higher than the 30 correct responses (or 50%) expected by chance ($t_1[43] = 8.36, p < .001$).

Finally, the difference between the two target groups was also found in subjects' post-experimental comments. Many more subjects could point out the defining property for /ka/ targets than /ak/ targets (21 vs. 7 subjects). Interestingly, among the subjects who reached criterion, only one subject in Experiment 4A identified the target /ka/ as a combination of /k/ and /a/, while three subjects in Experiment 4B treated the target /ak/ as a combination of /a/ and /k/.

3.4.4.4. Covariance between the Practice and Main Experiments

The results analyzed between Groups I and II of Experiment 4A and 4B have shown that there was no significant practice effect, and the results were thus pooled across these groups. For all three response measures, these pooled results strongly suggested that the concept of "words containing /ka/" in Experiment 4A was more readily mastered than the concept of "words containing /ak/" in Experiment 4B. However, the performance on the practice task was slightly better for subjects in Experiment 4A than subjects in Experiments 4B, as graphed in Figure 3.3 below. Therefore, it is necessary to check whether the differences between Experiments 4A and 4B (i.e., for only those subjects who did the practice task) would be significant if these subjects' performance on the practice task were taken into account.[10]

FIGURE 3.3. COMPARISON OF SUBJECTS' PERFORMANCE ON THE PRACTICE AND MAIN EXPERIMENTS 4A AND 4B



One possible way to test this is to perform an analysis of covariance, treating the results of the main experiments as variates and the results of the practice task as covariates. Since the main interest is in the variate, the covariate can be manipulated as a classification factor (Winer, Brown, and Michels, 1991). Thus, the total number of correct responses (ranging from 24 to 59) in the practice session was divided into the three classes as shown in Table 3.12 below.[11] As shown in the first column of the table, the subjects (n = 40) were classified into three groups in terms of their performance on the practice task. The figures in the second and third column are the number of correct responses taken from main experiments.

Using an experimental factor (i.e., 4A and 4B) as the variate and a classification factor (i.e., I, II, and III) as the covariate, an ANOVA was conducted both on the number of correct responses and on trials to criterion. In the number of correct responses, the experimental factor was significant ($F_1[1,34] = 5.25$; $p = .028$), while neither a classification factor ($F_1[2,34] = .25$; $p > .7$) nor any interaction was significant ($F_1[2,34] = .05$; $p > .9$). Similarly, in the trials to criterion, the experimental factor was marginally significant ($F_1[1,34] = 3.87$; $p < .06$), while neither classification factor nor any interaction was. The results from these analyses thus indicated that the difference in performance on Experiments 4A and 4B did not vary as a function of subjects' performance on the practice task.

TABLE 3.12. ILLUSTRATION OF ANALYSIS OF COVARIANCE OF TOTAL NUMBER OF CORRECT RESPONSES (Comparison of Experiments 4A and 4B on the basis of subjects' performance on Practice task)

Practice task (Covariate)	Experiments (Variate) 4A (n = 18)	4B (n = 22)
I. 24-35	32 48 53 60	32 34 38 39 42 46 46 50
(n = 12)	(mean = 48.3)	(mean = 40.9)
II. 36-47	30 31 35 37 46 52 58 59 60	30 32 33 34 36 38 49 56
(n = 17)	(mean = 45.3)	(mean = 38.5)
III. 48-59	32 33 50 59 59	27 32 33 34 43 54
(n = 11)	(mean = 46.6)	(mean = 37.2)

3.4.5. Discussion

The results reported in Section 3.4 have shown that the concept defined in terms of the body unit was easier for subjects to form than the concept defined in terms of the rime. Not only did the body concept have (i) more subjects reaching criterion, it also had (ii) fewer trials to reach criterion and (iii) more overall correct responses than the rime category.

These results confirmed that CV/C is a more natural division of a CVC syllable in Korean than C/VC and that CV has a more salient property than VC. In terms of the *basic level* notation discussed in Jaeger (1980a), the subsyllabic unit CV (body) in Korean appears to be more basic than the VC (rime). In other words, CV is basic in that it consists of a single subsyllabic unit (the body), whereas, VC is not in that it consists of two units (i.e., part of the body plus the coda). This difference seems to have been reflected in subjects' comments; most subjects considered /ka/ as one unit, while /ak/ was often described as a combination of /a/ and /k/.

In sum, the results of Experiment 4 were in accord with those of Experiments 1 and 2, which showed that CVC word pairs were judged more similar if they shared a common CV than a common VC.

3.5. Experiment 5

3.5.1. Goals

Experiment 5 was designed to compare the conceptual saliency of phonemes and syllables. Which concept in Korean is easier to learn, one defined in terms of phonemes or one defined in terms of syllables? Two groups of subjects participated. The target for the first group was "a word containing the phoneme /k/", and the target for the second group was "a word containing the syllable /kak/". Closely related to this comparison is the notion of *basic level* discussed in Section 3.3.1. If the phoneme is psychologically more basic than the syllable, as suggested by prior English CF experiments, subjects in the first group should find the task easier than subjects in the second group. That is, that there would be (i) more subjects learning the target; (ii) fewer trials to criterion, on average; and (iii) more overall correct responses. On the other hand, if the syllable is a more salient unit than the phoneme, as suggested by the results of Experiment 3, the opposite results should be found.

In the discussion of her CF experiments, Jaeger (1980a) linked the "basic units" that emerged to orthographic properties. She found, for example, that English subjects could learn easily the concept "words containing /k/", while only half of her Japanese subjects could learn the concept "words containing /t/". Jaeger argued that the basic level is a phoneme for English and a syllable for Japanese, and she attributed this discrepancy in basic levels to orthographic differences between the two languages:

> English speakers talk about the sounds of their language in terms of letters; therefore 'a sound' is equivalent to a letter-sized segment of the speech stream. Japanese speakers talk about the sounds of their language in terms of a syllabary; therefore 'a sound' is equivalent to a syllable. Because the unit of sound captured by the orthography has been learned consciously, it is the unit most easily available to consciousness... [T]he basic level of categorization of the sound of literate speaker's language is the level at which the orthography captures (p. 145-146).

However, there are some flaws in this line of reasoning. First, the conclusion that the phoneme is a basic level in English was drawn from the comparison with the phoneme in Japanese but not with other units (e.g., syllable) in English. Thus, the possibility still remains for English that the syllable is just as psychologically basic as the phoneme is. Similarly, all that was tested in Jaeger's Japanese experiment was the psychological status of the phoneme in that language, not the syllable concept *per se*. Furthermore, more recent evidence shows that the mora, rather than the syllable, is the basic level of speech perception in Japanese (Otake, Hatano, Cutler, and Mehler, 1993; Derwing and Nearey, 1994). In fact, the level that the Japanese orthography captures is not the syllable but the mora.[12]

An important aspect of Experiment 5, therefore, is a direct comparison between the segment and the syllable within a language. In addition, the orthographic interference found in the previous CF experiments can be minimized in Korean, owing to its unique writing system. Most notably, the fact that *both* phoneme-sized *and* syllable-sized orthographic units are represented in Korean permits a comparison between the phoneme and syllable that is as independent as possible of orthographic influences.

3.5.2. Experiment 5A

3.5.2.1. Subjects

The subjects were 23 freshmen taking an introductory English composition class at Sogang University.[13]

3.5.2.2. Stimuli

Experiment 5A was designed to test whether various allophones of a phoneme can be classified into a single category by Korean subjects. The target word was a "word containing /k/." There were several reasons for choosing /k/ as a target phoneme. First, using the same target phoneme as in the previous English experiments provides a maximally similar cross-language comparison. Second, the frequency of /k/ as an onset and as a coda is well balanced. Third, the allophonic variations for the Korean phoneme /k/ are suitably diverse, as summarized below:

(i) Unreleased syllable-finally: /k/ → $[k^{o}]$ / ____ $]_{\sigma}$

(ii) Voiced intervocalically: /k/ → [g] / [+voice] ____ [+voice]

(iii) Plain voiceless elsewhere: /k/ → [k] / $_{\sigma}[$____

Fourth, the orthographic situation for /k/ is also complex. Due to the syllable-final unreleasing that neutralizes all tense consonants and consonant clusters, /k/ has the following four spelling variants in addition to the standard <k> (ㄱ):

Standard spelling: ㄱ in 속: so<k> → /sok/ 'inside'

(i) ㄲ in 묶다 : mu<kk>.ta → /muk.t'a/ 'to tie'

(ii) ㅋ in 부엌 : pu.ə<kh> → /pu.ək/ 'kitchen'

(iii) ㄳ in 몫 ; mo<ks> → /mok/ 'share'

(iv) ㄺ in 흙 ; hɨ<lk> → /hɨk/ 'soil'

Interestingly, there is a directly opposite situation from the above spelling variants of /k/, where a segment is spelled as <k> but pronounced as /k'/: mul.<k>a → /mul.k'a/ 'price'. These orthographic variants of /k/ and the spelling variant of /k'/ will be reserved for the test session.

In the learning session there were a total of 60 words; 30 targets and 30 distractors. Targets were further divided into three types of allophones consisting of 10 words each as shown below:

TABLE 3.13. EXAMPLES OF STIMULI USED IN EXPERIMENT 5A

	Type	Example: One-syllable	Example: Two-syllable
Targets	[k]	kɨl	
	[k°]	pak	
	[g]		ma.kam
Distractors	∅ (D0)	pan	
	/ŋ/ (D1)	paŋ	so.paŋ
	/k'/ or /k^h/ (D2)	k^hal	k^hɨn.cip

Distractors were also divided into three types, depending on the degree of interference with the target: 10 distractors containing no velar consonants, 10 containing /ŋ/, and 10 containing /k'/ or /k^h/. In Table 3.13, numbers in D0, D1, and D2 indicate an increasing degree of interference, e.g., D0 distractors were assumed to be least interfering with the target words, since they did not contain any velar consonants, and D2 the most interfering, since they also involved velar stops.

The stimuli were played to subjects in the following order, to gradually teach the concept "words containing /k/":

	TARGETS		DISTRACTORS
Words 1-12:	3 [k], 3 [k^o] words	mixed with	6 D0
Words 13-36:	4 [k], 4 [k^o], 4 [g] words	mixed with	6 D1, 6 D2
Words 37-60:	3 [k], 3 [k^o], 6 [g] words	mixed with	4 D0, 4 D1, 4 D2

In the first list (words 1-12), two types of targets were intermixed with non-interfering distractors, and all stimuli were one-syllable words. In the second list (words 13-36), [g] words and D1 and D2 words were introduced, and stimuli were either one- or two-syllable words. In the last list (words 37-60), all three types of targets were intermixed with all three types of distractors. (See Appendix 7 for a full list of the stimuli used in Experiment 5A.)

In the test session, subjects heard the following list of words:

1. kɨm	T	2. k^hwæ.c^ha	D	3. pak	T	4. k'ul	D
5. hɨ<lk>	/k/	6. som	D	7. pu.ə<kh>	/k/	8. kyəŋ	T
9. mul.<k>a	/k'/	10. k^hoŋ	D	11. ta<lk>	/k/	12. cuk	T
13. in.<k>i	/k'/	14. po.ki	T	15. p^hum.sa<ks>	/k/	16.k'u.mim	D
17. hyo.<k>wa	/k'/	18. a.ki	T	19. sul	D	20. næs.<k>a	/k'/

The 20 words were divided into six targets, six distractors, and eight test words, and were designed to ascertain whether subjects' judgments were phoneme-based or spelling-based. Since the main interest in the test session was the way in which the test words were categorized by subjects, no reinforcement was made as to the correctness of subjects' responses in this session. T and D words indicate targets and the distractors, respectively; these words were used as controls to reinforce the concept. There were two types of test words: one with spelling variants of /k/ and the other with /k'/ words spelled with the letter <k> (ㄱ). "Yes" responses to /k/-test words indicate that subjects' judgments were phoneme-based, while "yes" responses to /k'/-test words indicate that they were spelling-based.

3.5.2.3. Results

3.5.2.3.1. Overall results

Out of 23 subjects, only seven reached criterion in the learning session (30.4 %). The mean number of trials to criterion was 55.6 (sd = 10.8), which was poorer than that of Experiments 4A and 4B. Even after excluding subjects who did not reach criterion, average trials to criterion was still a very unimpressive 43.1 (sd = 12.9). Compared with the English /k/ CF experiments, in which 90 % of the subjects reached the criterion and the average trials to criterion was only 23 (Ohala, 1986), the present results showed that the task was much more difficult for Korean speakers.

The results from the total number of correct answers also showed that the subjects' performance on the /k/ task was relatively poor. The average total number of correct responses was 36.3 (sd = 6.3; 60.5 % overall correct rate), which was smaller than that of any of the four groups in Experiments 4A an 4B. However, it was still significantly larger than the average of 30 which was expected by chance ($t[22] = 4.78$, $p < .001$).

3.5.2.3.2. Responses to each stimulus type

The responses to each type of target and distractor are given in Table 3.14. These figures include responses made at and after (but not before) reaching criterion. In the three target types, [k]-words had the highest percent correct responses and [k^{o}] -words the lowest. However, an ANOVA showed that the differences in individual subjects' percent correct responses to the three types of targets were not significant. Similarly, the distractor type without any velar sound had the higher percent correct responses than the other two types, but the differences were again not significant.

TABLE 3.14. NUMBER AND PERCENT CORRECT RESPONSES TO TYPES OF TARGETS AND DISTRACTORS IN LEARNING SESSION OF EXPERIMENT 5A (Based on seven subjects who reached criterion)

	Type	Total #	# Correct	% Correct
Targets	[k]	28	24	85.7
	[k^{o}]	31	23	74.2
	[g]	49	41	83.6
Distractors	no velars	31	31	100
	velar nasal	41	38	92.6
	velar stops	41	37	90.2

That there were no differences in percent correct responses among the three target types was similar to the results of Jaeger's English /k/ experiment (1980a: 201). It appears, therefore, that phonetic environments of the phoneme /k/ were not a significant factor on subjects' performance in either the Korean or the English /k/ CF experiments. Instead, orthography seemed to have played a more important role, at least in the English case. In Jaeger's English /k/ experiment above, in fact, there were significant differences when target words were grouped by spelling (e.g., <k>-words, <c>-words, etc.)

3.5.2.3.3. Responses to test words and subjects' comments

The test session was designed to test whether subjects' judgments were phoneme-based or spelling-based. The analyses were again based on the responses made by the seven subjects who reached criterion. In the 20 stimuli in the test session, four were words containing orthographic variants of the phoneme /k/; another four were words containing the phoneme /k'/, which were spelled as <k>; and the remaining 12 were targets or distractors that were learned in the learning session. The responses to this latter set of 12 stimuli had a quite high overall correct percentage (92.9%; 78 correct responses out of 84), indicating that the subjects did learn the concept of "words containing /k/."

Table 3.15 shows the responses to both kinds of test words.

TABLE 3.15. NUMBER AND PERCENTAGE OF "YES" RESPONSES TO TEST WORDS AND COMMENTS ON THE TARGETS WORDS MADE BY INDIVIDUAL SUBJECTS

Subject (n=7)	/k/ test words (n=4)	/k'/ test words (n=4)	Comments on target words
#1	3 (75%)	1 (25%)	If there is the letter <k>, "Yes"; if there are letters <s> <l>, "No".
#5	2 (50%)	3 (75%)	No comments
#7	1 (25%)	4 (100%)	The <k> is in the first sound or in the last sound.
#11	0 (0%)	4 (100%)	Voiced sound vs. unvoiced sound
#14	0 (0%)	4 (100%)	The letter <k>
#22	0 (0%)	4 (100%)	Strong sound vs. soft sound
#23	4 (100%)	0 (0%)	Harsh sound vs. liquid sound
Total (n=28)	10 (35.7%)	20 (71.4%)	

As can be seen in the table, the percent "yes" responses to /k/ test words (i.e., spelling variants of the phoneme /k/) was only 35.7%, indicating that the subjects' judgments were in most cases spelling-based. On the other hand, the /k'/ test words which are spelled as <k> but pronounced as [k'] had 20 "Yes" responses out of 28 (71.4%). In particular, subjects #7, #11, #14, and #22 consistently categorized the /k'/ test words into the target. Interestingly, although subjects #11 and #22 described the defining property as "sound," these subjects' actual judgments were based on the spelling <k>.

In sum, the results of the test session suggest that orthography had a significant effect on the task of classifying various allophones of /k/ into a category. The results of the learning session also showed that phonetic details of the target types were not a significant factor. These two sets of results are in general accord with Jaeger's English /k/ experiments (1980ab, 1986a).

3.5.3. Experiment 5B

3.5.3.1. Subjects

The subjects were 21 freshmen taking an introductory English composition class at Sogang University.[14]

3.5.3.2. Stimuli

The concept taught was "a word containing the syllable /kak/". The main concern in designing this experiment was to keep both the types of stimuli and the procedures maximally similar to those in Experiment 5A. First, the instructions were identical in both experiments. Second, the stimuli in the learning session were again composed of 30 targets and 30 distractors, and the targets and distractors were divided into the following three types, which parallel as closely as possible the types used in Experiment 5A (see Table 3.16). The targets were further divided into 10 TI (Initial), 10 TF (Final), and 10 TM (Middle) words, depending on the location of the syllable /kak/, a distribution that was designed to parallel the three allophonic variations of /k/ in Experiment 5A. Numbers in D0, D1, and D2 indicate the number of critical phonemes (i.e., /k/ and /a/ in this case) contained in the distractors, reflecting the increasing degree of interference with the targets. In order to minimize any potential effects of the number of syllables on subjects' performance, the ratio of two-syllable words to three-syllable words was also maintained the same as the ratio of one-syllable words to two-syllable words in Experiment 5A.

TABLE 3.16. EXAMPLES OF STIMULI USED IN EXPERIMENT 5B

	Type	Example: Two-syllable	Example: Three-syllable
TARGETS	kak.- (TI)	kak.pyəl	
	-.kak (TF)	sim.kak	
	-.kak.- (TM)		ta.kak.to
DISTRACTORS	no /k/ or /a/ (D0)	toŋ.yo	
	/k/ or /a/ (D1)	koŋ.toŋ	ku.səŋ.wən
	/ka/ or /ak/ (D2)	kam.sok	kan.ho.wən

The stimuli were played back to subjects in the following order, to gradually teach the concept "words containing the syllable /kak/"

	TARGETS		DISTRACTORS
Words 1-12:	3 TI, 3 TF words	mixed with	6 D0
Words 13-36:	4 TI, 4 TF, 4 TM words	mixed with	6 D1, 6 D2
Words 37-60:	3 TI, 3 TF, 6 TM words	mixed with	4 D0, 4 D1, 4 D2

The relative ordering between types of targets and distractors and the reinforcement schedule were also the same as in Experiment 5A, in order to maintain the same increasing degree of difficulty toward the end of the list. (See Appendix 8 for a full list of the stimuli used in Experiment 5B.) There was no test session in Experiment 5B.

3.5.3.3. Results

3.5.3.3.1. Overall results

Out of 21 subjects, 18 reached criterion (85.7%). The mean number of trials to criterion was 27.8 (sd = 16.3), which was the best of all the Korean CF experiments discussed so far. Excluding the three subjects who did not reach criterion, the result was a quite respectable 22.2 (sd = 9.3). Most subjects could identify the target very easily and early. These results of Experiment 5B on Korean syllable /kak/ were much more comparable to those of Jaeger' English /k/ experiments than to those of Experiment 5A on the Korean phoneme /k/.

The results from the total number of correct answers also pointed in the same direction. The average total number of correct responses was 51.8 (sd = 8.5; 86.2% overall correct rate), which was again best of all the Korean CF experiments.

3.5.3.3.2. Responses to each stimulus type

Calculated on the basis of responses made at and after reaching criterion, the percentage of correct responses was higher than 94% for all three target types and all three distractor types. As in Experiment 5A, there was a trend in which TI (target-initial syllable in Experiment 5B and target initial phoneme in Experiment 5A) had the highest percentage of correct responses of all target types, but there was again no significant difference either among the three types of targets or among the three types of distractors. Thus the responses to each stimulus type will not be analyzed in detail. Suffice it to say here that neither the location of the target syllable nor the supposed degree of interference mattered in the task of mastering a syllable-based concept in Korean.

3.5.3.3.3. Subjects' comments

Among the 18 subjects who reached criterion, all but one subject commented in one way or the other on the common property of the target words. Fourteen subjects (72.2%) mentioned the letters <kak>, e.g., "The words contained the letters <kak>." Only three subjects (16.7%) attributed the common property to pronunciation, e.g., "The sound /kak/ was pronounced in the words." It seemed thus obvious that orthography played an important role in Experiment 5B, as well as in Experiment 5A.

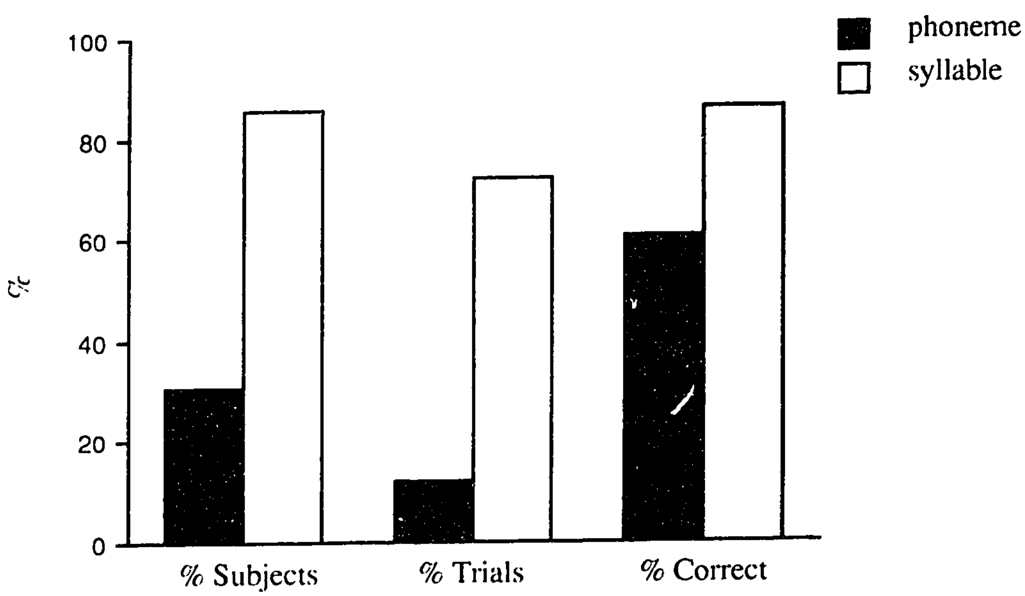
In Experiment 5B, not only were there more subjects who correctly commented on the target, but there was less variation in individual subjects' comments than in Experiment 5A. In addition to the three response measures discussed in sections 3.5.2.3.1 and 3.5.3.3.1, therefore, subjects' naming ability also showed a difference between the phoneme-based concept and the syllable-based concept in Korean.

3.5.4. Discussion

3.5.4.1. Overall Comparisons between Experiments 5A and 5B

As discussed in sections 3.5.2. and 3.5.3., the phoneme-based concept (in Experiment 5A) was much more difficult to form than the syllable-based concept (5B). All three response measures showed that subjects in Experiment 5A found the task more difficult than subjects in Experiment 5B, as shown in Figure 3.4. Table 3.17 below shows the results of *t* tests for the three sets of comparisons that were available from the present experimental design. All of the three *t* scores were highly significant ($\alpha = .0001$), revealing a big difference between the phoneme-based concept and the syllable-based concept in Korean.

FIGURE 3.4. DIFFERENCE BETWEEN THE PHONEME AND THE SYLLABLE IN TERMS OF "SUCCESS RATE" * ON THE THREE RESPONSE MEASURES



* For the ease of comparison, trials to criterion were converted to percentage by the formula explained in Figure 3.2.

TABLE 3.17. COMPARISON OF RESULTS OF EXPERIMENTS 5A AND 5B (n = 44)

	Experiments		Differences	
Response Measures	5A: /k/	5B: /kak/	$t_1(42)$	$t_2(59)$
Subjects reaching criterion	7 / 23	18 / 21		
Trials to criterion	55.6	27.8	6.73	
Total correct answers	36.3	51.7	- 6.93	- 11.45

3.5.4.2. Cross-linguistic Comparison

The present section compares the results of Experiments 5A and 5B with those of the English (Ohala, 1986a) and Taiwanese (Derwing and Wang, in press) phoneme tests. Table 3.18 shows some comparisons among these studies:

TABLE 3.18. CROSS-LINGUISTIC COMPARISON OF THE PHONEME CONCEPT

Response Measures	Korean	English	Taiwanese
Trials to criterion used	13/15 *	13/15	12/14
% Ss reaching criterion	30% (86%)**	90%	23% ***
% correct answers	61% (86%)	not available	53%
Trials to criterion for Ss reaching criterion only	43 (22)	23	not available

* 13/15 indicates 13 correct answers out of 15 trials in a row; ** numbers in parenthesis indicate the results for the Korean syllable /kak/; and *** the Taiwanese results were averaged across the onset and the coda.

Although the experimental designs were slightly different (e.g., trials to criterion used), the figures in Table 3.18 provide a reasonable ground for a cross-linguistic comparison. First, the percent of subjects reaching criterion shows that the phoneme-based concept in Korean was much more difficult to form than in English but about as difficult as that in Taiwanese. Instead, the syllable-based concept in Korean had about the same percent of subjects reaching criterion as the phoneme-based concept in English. The other response measures also clearly show that the phoneme concepts in Korean and Taiwanese were generally more difficult to form than the phoneme concept in English, but the latter was about equal in difficulty to the syllable concept in Korean.

3.5.4.3. Conclusion

The comparisons made in the last two sections have shown that the phoneme-based concept was much more difficult for Korean speakers to form than the syllable-based concept. Although the phoneme in Korean receives orthographic support (unlike the Taiwanese case where individual phonemes are not written), the phoneme in Korean as well as in Taiwanese has turned out to be a less accessible phonological unit than the

English counterpart. Instead, the syllable in Korean has manifested itself as a more viable phonological unit, on the basis of which concepts are more readily formed. In sum, the results of Experiments 5A and 5B strongly suggest that the syllable is the most basic phonological unit in Korean, which plays a role that roughly parallels the phoneme in English.

3.6. General Discussion

3.6.1. Summary of the Four Phonlogical Units Tested

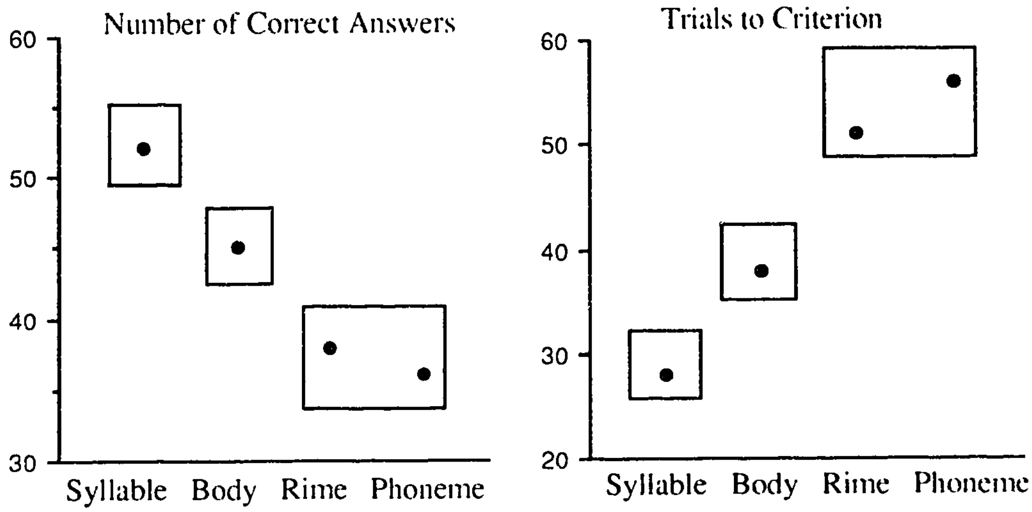
It is clear from the above results that the syllable was by far the easiest concept to form. In raw quantitative terms, the overall difficulty of the four phonological concepts tested in the present experiments was as follows (ordered as indicated, from easiest to most difficult): syllable → body → rime → phoneme. Table 3.19 below gives the summarized results for the four concepts in Korean:

TABLE 3.19. COMPARISON OF THE FOUR PHONOLOGICAL CONCEPTS IN TERMS OF THE THREE RESPONSE MEASURES

Response Measures	Syllable	Body	Rime	Phoneme
% Ss reaching criterion	86	59	32	30
# Correct answers (0-60)	52	45	38	36
Trials to criterion (15-61)	28	38	51	56

All three of the response measures pointed in the same direction for all four concepts tested (e.g., in all three measures, the syllable was best, followed by the body, etc.). Figure 3.5 below shows the statistical significance of the differences among the four concepts (α = .05, two-tailed *t* tests by subjects).

FIGURE 3.5. GROUPINGS OF THE FOUR CONCEPTS IN TERMS OF LEAST SIGNIFICANT DIFFERENCE



NOTE: Boxes enclose means that are not significantly different.

As can be seen in Figure 3.5, by both response measures the only non-significant difference was between the rime and the phoneme. The status of the rime in this regard was similar to that of the segment. The segmental status of the rime in Korean was also observed in subjects' comments (Experiment 4B), in which the target /ak/ was described as a combination of /a/ and /k/ or erroneously as /a/ or /k/. In the results of both response measures, on the other hand, the body was intermediate in difficulty between the syllable and the other two units. It can also be seen in Figure 3.5 that the two response measures had a very strong (negative) correlation, indicating the reliability of both measures.

3.6.2. Assessment of the Present Concept Formation Technique

There were two main modifications from the previous English CF experiments, one in experimental procedures and the other in dependent variables.

3.6.2.1. Experimental Procedures

In the present experiments, subjects were tested as a group and were asked to give written responses. Overall, group administration proved to be effective, as relative difficulties among the four concepts tested were consistently reflected in all response measures used. Since a *comparison* of various phonological concepts was the main purpose in the present CF experiments and it was predicted that there would be great individual variation, more subjects were required than if the focus had been on the categorization of some ambiguous members of a single concept. For this purpose, group administration was an efficient means to obtain the requisite amount of data.

The written response mode, together with the exclusion of response time, may have made the present CF experiments tap a relatively high level of processing only. As Jaeger (1980a) pointed out, however, response time can be dispensed with in CF experiments on the following grounds:

> The concept formation experiments (without measuring response time) tapped a more volitional level of response than the CC [Classical Conditioning] experiment. But since subjects gave essentially the same responses with the two formats, it appears that the responses in the CF experiment were based on the same intuitive preexisting categorizations as those elicited by the CC design. (p. 349)

3.6.2.2. Response Measures Used

Instead of response time, we have used the total number of correct answers as a new response measure in the present experiments. In this section, the viability of the three response measures including this new one will now be discussed.

First, the number of subjects reaching criterion has provided useful initial information on the overall difficulty of the concepts. A potentially serious probelm, however, is that the number of subjects reaching criterion is completely dependent upon how the criterion was set. In other words, trials to criterion can be manipulated by the experimenter, such as 13 correct answers out of 15 consecutive trials (in the present experiments and most previous English CF experiments), 12 out of 14 (Derwing and Wang, in press), and 9 out of 10 (Jaeger, 1980b). All three criteria are extremely difficult to be reached by chance ($p < .01$).

On the other hand, the total number of correct answers cannot be manipulated and proved to be particularly useful in the present experiments. For example, a subject in Experiment 4B, where the target was the rime /ak/, commented that the words had the common <a> *and* were all concrete nouns as opposed to abstract ones. Although this subject happened to reach criterion at the earliest possible trial (i.e., 15th), his overall performance was relatively poor (46 correct answers, or 77% correct rate) because he had rejected all abstract nouns. Interestingly, this subject could not have reached criterion, if it had been set at 10 consecutive correct answers. In that case, however, rare accidental response mistakes would be expected to take their toll.

In general, we would expect the total number of correct answers to be negatively correlated with trials to criterion. The three alternative trials to criterion measures are compared below in terms of their correlations with the total number of correct responses, in order to check the reliability of each measure. The results are given in Table 3.20. It is obvious that the criterion 13/15 had the best overall correlations with the total number of correct responses. Correlations given in Table 3.20 provide additional information on the overall difficulty of the concepts: the stronger the correlation, the easier the concept. Specifically, most correlations were around -.9 in Experiments 4A and 5B, in which the concepts were easier; whereas most correlations were around -.7 in Experiments 4B and

5A, in which the concepts were more difficult. In terms of the criterion 13/15, in particular, the relative order of difficulty of the four concepts as discussed in section 3.6.1 was directly reflected in the degree of correlations: the syllable (r = .95); the body (r = .93); the rime (r = .77); and the phoneme (r = .75).

TABLE 3.20. CORRELATIONS BETWEEN TOTAL NUMBER OF CORRECT RESPONSES AND THREE TRIALS TO CRITERION

	Trials to Criterion					
Experiments	12/14		13/15		10/10	
4A: /ka/	-.91	(3)	-.93	(2)	-.95	(1)
4B: /ak/	-.77	(1)	-.77	(1)	-.74	(3)
5A: /k/	-.73	(2)	-.75	(1)	-.59	(3)
5B: /kak/	-.81	(3)	-.95	(1)	-.93	(2)

NOTE: Numbers in parentheses indicate rank order of correlations within each experiment (row).

Furthermore, the results of the present experiments suggested that trials to criterion (A) including all subjects, could be more informative than trials to criterion (B), including only subjects reaching criterion. Table 3.21 below shows the level of significance among the four phonological concepts tested. As indicated by the results in Table 3.21, trials to criterion (A) and (B) show drastic contrasts: all differences significant under criterion (A) were not significant under criterion (B), and the one significant difference under criterion (B) was not significant under criterion (A). Moreover, it is clear that the results using criterion (A) were more in accord with the number of correct answers than were the results for criterion (B). Furthermore, the percent of subjects reaching criterion was more strongly correlated with criterion (A) than with criterion (B), as graphed in Figure 3.6. In sum, the major modifications proved to be successful in the present experiments. The total number of correct answers, in particular, provided a useful basis against which the reliability of other response measures could be checked.

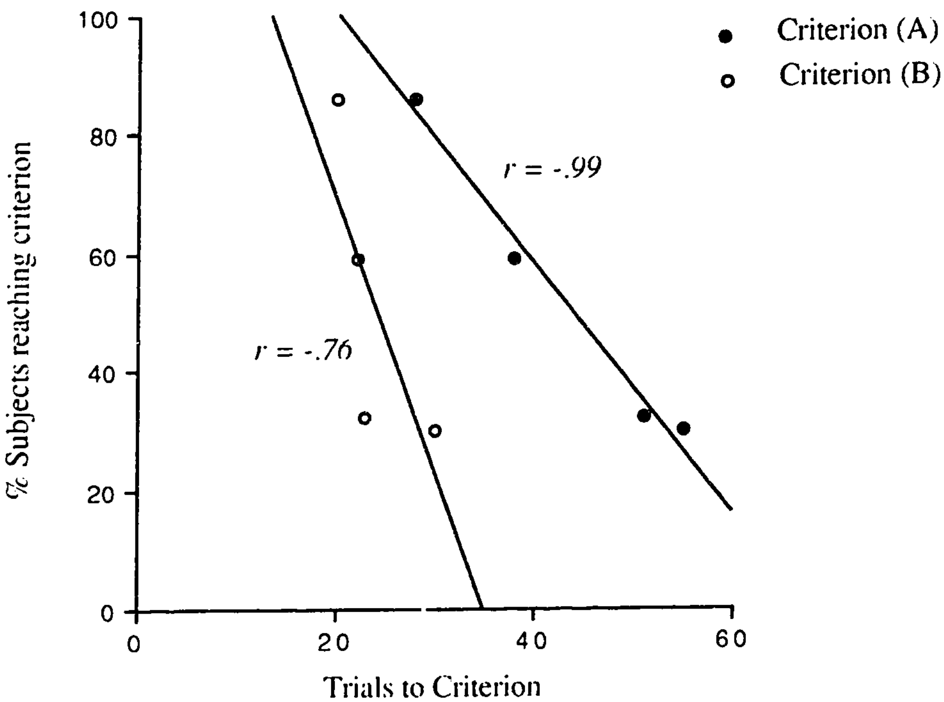
TABLE 3.21. COMPARISON OF FOUR PHONOLOGICAL CATEGORIES IN TERMS OF FOUR RESPONSE MEASURES

Response Measures	Syllable		Body		Rime		Phoneme
% Ss reaching criterion	86		59		32		30
Number of correct answers	52	←*→	45	←**→	38	=	36
Trials to criterion (A)	28	←*→	38	←**→	51	=	55
Trials to criterion (B) for Ss reaching criterion only	22	=	23	=	30	←*→	43
		←-------	=	--------→			

Note: = non-significant; ←*→ significant at the .05 level; ←**→ significant at the .01 level [by unpaired, two-tailed *t* tests by subjects]

Furthermore, trials to criterion when set at 13 correct answers out of 15 trials in a row had the strongest correlation with the total number of correct answers. It was also discussed that, when a relatively small number of subjects reached criterion, trials to criterion calculated on the basis of all subjects turned out to be more informative than those which excluded subjects who did not reach criterion. Therefore, the two new response measures used throughout the present CF experiments, (i) the total number of correct responses and (ii) trials to criterion set at 13 correct answers out of 15 trials in a row and including all subjects, proved to be very useful for the comparison between concepts.

FIGURE 3.6. CORRELATIONS BETWEEN THE PERCENT OF SUBJECTS REACHING CRITERION (A) AND CRITERION (B)



Notes to Chapter III

[1] In linguistics, this notion of sharp boundaries among categories is illustrated in the strict binariness of distinctive features in generative phonology, where phonemes sharing a certain feature are in contrast with other phonemes lacking the feature. Thus, there can be no phoneme whose category membership is ambiguous in the binary classification of features, and phonemes sharing a feature are all *equally* good instances of that feature.

[2] Since subjects included in their target concept such words as *choir*, *quit*, *cash*, etc., the results, as Jaeger argued, could not simply be dismissed as orthographic artifacts. Jaeger's (1980a) results concerning the categorization of [sk]- words were replicated by Ohala (1986). (See section 3.2.3.2.)

[3] Jaeger and Ohala (1984: 24) argued that some phonemes are better instances of the feature [±anterior] than others (e.g., labial consonants are more "anterior like" than alveolar consonants).

[4] Other criteria can also be used such as 10 correct answers in a row or 14 consecutive trials with two or fewer errors (Derwing and Wang, in press). A comparison of these criteria will be discussed in section 3.6.

[5] Jaeger (1980a: 176-177) suggested that "The small number of subjects, small amount of error data, and positional interference have rendered these experiments [CF with a motor response] less than ideal for studying response time phenomena."

[6] Thanks are expressed here to Dr. Chung Ja Kwon and Dr. Kyoung Sun Hong and to the students in their classes.

[7] Thanks are due here to Dr. Terry Taerum for his statistical advice as well as help in tabulating the data.

[8] Thanks are expressed here to Dr. Chung Ja Kwon, Dr. Kyoung Sun Hong, Mr. Jae Hee Cho, Mr. In Key Chung, and to the students in their classes.

[9] It was originally planned to have all subjects do the practice task before the main experiment, the main purpose of the practice task being to check whether the average ability of each subject group was about equal on a CF-type task. However, the time prearranged for the experiment turned out to be insufficient for both tasks, primarily because of many questions the subjects had about the instructions. After running the first two groups, the experimenter decided to focus on the main experiment with the remaining groups. Thus, the subjects in the remaining groups were given more time for the instructions and questions, instead of doing the practice task.

[10] Thanks are expressed here to Dr. Rebecca Treiman for suggesting this analysis.

[11] There can be various criteria for the division of covariates. In general, the frequency of each cell after the division is one of the most important criteria. Since the present data from the total number of correct answers in the practice task ranged from 24 to 59 (with 36 possible scores), it was decided that subjects be divided into three groups (consisting of 12 scores each) in terms of their performance on the practice task.

[12] Jaeger seems to have used the term syllable quite loosely. Although the mora and the syllable often coincide in Japanese, syllables such as /tan/ consisting of two morae (i.e., /ta/ + /n/) can be superordinate in comparison with monomoraic syllables.

[13] Thanks go to Dr. Chung Ja Kwon and to the students in her class.

[14] Thanks go to Dr. Kyoung Sun Hong and to the students in her class.

CHAPTER IV

SUMMARY AND CONCLUSION

Chapter I introduced the two main issues addressed in the present work: syllable structure and the basic representational unit in Korean. The results of the experiments discussed in Chapters II and III suggested that Korean contrasts with English on both counts. This last chapter reviews the main outcomes of the experiments (4.1), discusses the directions for future research (4.2), and makes closing remarks (4.3).

4.1. Review of the Experiments

Using global sound similarity judgments (SSJ) and concept formation (CF), five experiments were carried out to test the issues of *syllable structure* and *basic level of representation* in Korean. Table 4.1 summarizes the main outcomes of these experiments. As for syllable structure, the CV (body) and the VC (rime) of a CVC syllable were compared; and as for the second issue, the C (phoneme) was pitted against the CVC (syllable). In both cases, comparisons were made in terms of the saliency of these units in making the SSJ judgments and in terms of the ease with which these units were learned in a CF task.

TABLE 4.1. SUMMARIZED RESULTS OF ALL PRESENT EXPERIMENTS

Experiment	Task	Stimuli	Issue	Results
1	SSJ	CVC pairs	CV vs. VC	CV
2	SSJ	CVC pairs	CV vs. VC	CV
3	SSJ	CV.CVC pairs	C vs. CVC	CVC
4	CF	CVC.CVC words	CV vs. VC	CV
5	CF	CVC.(CVC.CVC) words	C vs. CVC	CVC

4.1.1. Syllable Structure

In Experiment 1, subjects judged the sound similarity of pairs of CVC words, which varied in both the number and the position of matched phonemes. One of the main results was that the phoneme was a significant variable, much as in the previous English SSJ experiments (e.g., Vitz and Winkler, 1973), as the similarity scores for most pairs could be differentiated in terms of the number of matched phonemes only. The only exception to this generalization was that pairs sharing the CV were judged significantly more similar than pairs sharing the VC. In addition, the CV contributed more to predicting similarity scores than did the VC.

Using CVC pairs again, Experiment 2 focused on the two types of pairs (CVX-CVX and XVC-XVC, where X indicates a mismatched phoneme) and varied the degree of similarity between the mismatched phonemes (Derwing and Nearey, 1986; Bendrien, 1992). Though some feature effects emerged, the main finding of this experiment was again the persistent pattern of higher similarity scores for CV matches over VC matches, regardless of the type of mismatch in the X position. The results of Experiments 1 and 2 thus both indicated that the CV (body) is a more salient subcomponent of a Korean CVC syllable than the VC (rime). Not only did the present SSJ experiments in Korean replicate our earlier findings using the same SSJ technique (Yoon and Derwing, 1994), they were also compatible with the results from prior word-blending studies (Derwing, Yoon, and Cho, 1993; Yoon, 1994a).

Experiment 4 was designed to test whether the salience of CV in the SSJs of Korean could be reflected in a radically different type of experimental task. We used the CF technique to compare the relative ease with which a CV-based concept and a VC-based concept were formed. Specifically, subjects in one group had to identify a target set defined in terms of a common body unit /ka/, while subjects in the other group did the same with a target set defined in terms of the rime /ak/. The former concept was easier to learn than the latter by all three response measures used: (i) more subjects learning the concept; (ii) more correct responses; and (iii) fewer trials to reach the mastery criterion.

The results of Experiments 1, 2, and 4, together with our earlier findings, all clearly indicate that the Korean syllable *is* left-branching, with the body as a major subsyllabic unit. As far as syllable structure is concerned, Korean and English thus do seem to contrast sharply, since most prior research on English has supported the rime as a viable subsyllabic unit in that language (Derwing and Nearey, 1991; Dow, 1987; Treiman, 1989, to name but a few). As noted in Chapter I, what naturalistic behavioral evidence there is available has also supported the body unit in Korean (Ahn, 1988; Cheon, 1980; Gim, 1987) but the rime in English (e.g., Fromkin, 1971; Hockett, 1967; Shattuck-Hufnagel, 1983; Stemberger, 1983). Only in the English case, however, are most theories of syllable structure (Fudge, 1969, 1987; Selkirk, 1982, etc.) supported by empirical evidence, which is not in the case in Korean (e.g., Kang, 1991; Kim, 1986; Kim-Renaud, 1978; Lee, 1993; Sohn, 1987). The results supplied here thus constitute a major challenge to the theoretical approach that has been adopted by most Korean phonologists heretofore, and to the kinds of evidence adduced.

4.1.2. Basic Level of Representation

Using the same two experimental techniques, the SSJ and CF, three experiments were carried out to test the status of the phoneme *vis-a-vis* the syllable as the basic level of phonological representation in Korean. These experiments were designed to be maximally similar to the ones previously carried out in English, in order to allow for objective cross-linguistic comparisons. Experiment 1 tested the role of the phoneme in predicting the SSJs of Korean CVC syllable pairs. As with the SSJs for English CVC syllables (e.g., Derwing and Nearey, 1994), the individual phonemes played a significant role. Since only monosyllabic words were used in Experiment 1, however, the question remained as to whether the findings could be extended to polysyllabic words, in which the syllable could also play some role. If the phoneme in Korean were as basic a representational unit as it is in English, much the same effects should be found in polysyllabic word pairs as in the English case (Vitz and Winkler, 1973), regardless of the intervening syllable boundaries.

Experiment 3 thus compared the predictions of similarity based on the phoneme and on the syllable, using disyllabic CV.CVC real words. Overall, predictions based on the syllable were more accurate than those based on the phoneme. For example, pairs such as /ku.saŋ-po.saŋ/ and /ku.saŋ-ku.sək/ were all judged more similar than pairs such as /ku.saŋ-ko.caŋ/ and /ko.cəŋ-ku.səŋ/. These differences are hard to explain on the basis of a phoneme count, since all of the above pairs had two phonemes mismatched out of five (as indicated by underlining); they are easy to explain on the basis of a syllable count, however, since mismatches in the former types of pairs involved only one syllable while the latter involved both syllables. The results of Experiment 3 thus suggested that the syllable plays a more important role than the phoneme in predicting the SSJs of Korean.

In Experiment 5, we used the CF technique to further compare the phoneme and the syllable in Korean. A series of CF experiments in English (Jaeger, 1980ab; Jaeger, 1986a; Ohala, 1986) provided evidence for the phoneme as the basic level of representation, as subjects could easily learn a phoneme-based concept. CF experiments applied to Japanese (Jaeger, 1980a) and to Taiwanese (Derwing and Wang, in press), however, showed that the task was much more difficult for speakers of these languages, suggesting that the phoneme may not be an ideal unit for them. The results of Experiment 5A also showed that a phoneme-based concept in Korean was about as difficult to learn as in Japanese and Taiwanese. Instead, a concept based on the syllable in Korean (Experiment 5B) was the one whose ease of mastery was comparable to a phoneme-defined concept in English.

In terms of Jaeger's (1980a) arguments about the basic level of phonological representation, the phoneme and the syllable in the languages discussed above would be classified as follows:

TABLE 4.2. CROSS-LINGUISTIC COMPARISON OF BASIC LEVEL OF PHONOLOGICAL REPRESENTATION

	English	Korean	Japanese	Taiwanese
Super-ordinate	syllable			
Basic	phoneme	syllable	syllable (mora)	syllable
Sub-ordinate		phoneme	phoneme	phoneme

Why is there a difference in the *size* of the basic level between English and the other languages, including Korean? One possible interpretation might be that the basic level is closely related to the orthographic systems employed, as Jaeger argued: "The basic level of abstraction for conceptualizing the sounds of one's language is that level which the writing system captures." (1980a: 366). In the case of Korean, however, this argument is not persuasive, because the phoneme receives orthographic support just as clearly it does in English; in fact, even more so, as the individual phonemes are much more systematically represented in the spelling in Korean than in English, where the level of regular grapheme-phoneme correspondence is much less. (As a sidelight, our Experiment 3 also showed that the phoneme effect was greater when orthographic representations were provided than when not.)

A second interpretation of our results may be that the basic level in a language may depend on the complexity of its syllable types. The maximal syllable representation in Korean (as in both Japanese and Taiwanese cases) is CVC,[1] whereas the syllable types in English can range from a simple V to CCCVCCC. Given even roughly comparable phonemic inventories, the number of possible syllables is vastly larger in the English case; in the Korean (and other) cases, however, the relatively small number of syllables allowed may prescribe a more efficient strategy for utilizing syllables as wholes, rather than continually breaking down them into their parts.

Yet another possible interpretation is that the syllable may be a basic unit in Korean because of the relatively clear syllable boundaries involved (Derwing, 1992a). In English (and perhaps other stress-timed languages), especially those that permit a large variety of consonant clusters, ambiguous or fuzzy syllable boundaries may make the syllable less than an ideal unit. In both interpretations, whether (i) simple vs. complex syllable types or (ii) clear vs. unclear syllable boundaries, the Korean syllable seems to have some obvious processing advantages over the English syllable, and the present experiments all provide clear evidence for the prominence of the syllable as the most basic representational unit in Korean.

4.2. Directions for Future Research

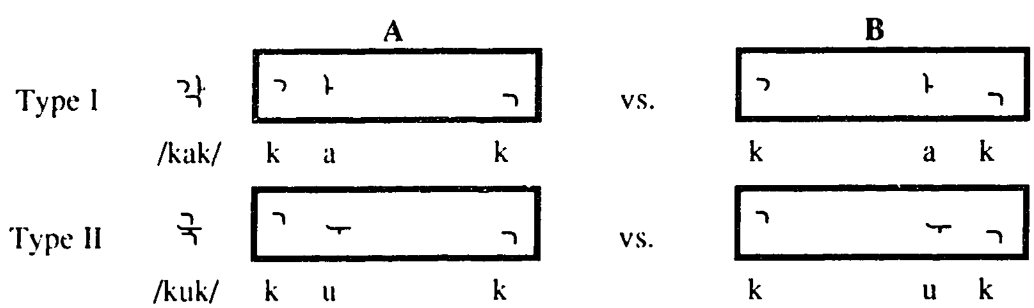
The experimental techniques used in the present research proved to be useful primarily for the following reasons: the results of the SSJ task showed that similarity ratings were consistent both across subjects and across items of the same types. In other words, individual subjects' judgments were stable enough to draw strong conclusions. In the CF task, although there was greater individual variation, the response measures used were still sufficiently reliable in that they all pointed to the same conclusion.

The present study, however, still tapped a relatively high level of conscious analysis of real word stimuli, especially in the CF task, and thus might have incurred some degree of orthographic influence. One of the remaining questions was thus whether the present results can extend to real-time speech processing. In this section, we will briefly discuss directions for future research along this line.

4.2.1. Orthographic Aspects

As already noted, an interesting and unique aspect of the Korean orthography is that individual letters are "packaged" to form syllables. In this syllable-forming rule, there are two types of vowels: one type, involving a large vertical stroke (e.g., ㅏ /a/), is graphically closer to the initial consonant than to the final consonant; the other, with a large horizontal stroke (e.g.,ㅜ /u/), is positioned between the two consonant letters and is thus neutral to this distinction (see Figure 4.1 below). If this orthographic convention were a significant variable, the relative saliency of the CV over the VC might be different depending on the types of vowels involved, i.e., the initial C might be more likely to form a CV unit with the former type of V than with the latter. Our Experiments 1 and 2, however, showed that there was no such difference reflected in sound similarity judgments. Another way to probe this difference might be to measure response time in visual word recognition. In English, there has been some evidence that subsyllabic units such as onsets and rimes (Santa, 1976-7), as well as orthographically defined units (e.g., Taft, 1979), facilitate recognition of printed words. It would be interesting to compare latencies in recognizing the following printed forms:

FIGURE 4.1. COMPARISONS OF PRINTED WORD RECOGNITION IN TERMS OF VOWEL TYPES AND SUBSYLLABIC UNITS



In a Type I syllable, response time for A is expected to be faster than B for orthographic reasons (among others). A more interesting comparison might involve Type II, in which the vowel is neutral in its orthographic influence. In English, Santa (1976-7) showed that words were faster to recognize when they were divided into onset and rime (e.g., C AT) than otherwise (e.g., CA T). Similarly, if response time were still faster for A than B in Type II above, this would constitute evidence that the body unit in Korean is involved in visual word recognition. The results might also shed light on whether or not our present findings on syllable structure can generalize to on-line processing

4.2.2. Statistical Dependency

Selkirk (1982) argued that phonotactic constraints in English provide evidence for the onset/rime division. In the English syllable, virtually any onset can occur with any vowel, but there are some collocational restrictions between the vowel and the coda.[2] In the Korean syllable, on the other hand, there is neither a phonotactic constraint nor a phonological rule that refers to the rime, such as stress assignment rule in English (Selkirk, 1982). In sum, while in English the vowel and the coda are closely related in various phonological aspects, they are independent of each other in Korean (Yoon, 1994a).

Recently, Treiman (personal communication) pointed out a computer-based approach to the phonological patterns in the English syllable. She found that the frequencies of VCs are much more difficult to predict than the those of CVs from knowing the frequencies of individual phonemes that may occur in CVC syllables. In other words, what kind of coda can occur often depends on the preceding vowel and vice versa, whereas the occurrence of an onset and a following vowel are statistically independent of each other. Interestingly, the results of these statistical analyses of CVCs are consistent with the general syllable structure conditions in English that involve much more complicated syllable types. If we could find an analogous statistical dependency between the V and the preceding C in the Korean syllable, it would reveal a great deal about the differences between the two languages. This statistical approach seems particularly interesting in that it tests the entire population, in effect, rather than samples.

4.2.3. Target-monitoring Tasks

One of the most frequently used methods to compare the status of the phoneme and the syllable as a basic representational unit is the target-monitoring task, in which response time is compared by manipulating targets and target-carrying words. The previous research in English has provided some mixed results supporting both units (e.g., McNeill and Lindig, 1973; Foss and Swinney, 1976). In French, on the other hand, the technique provided clear evidence in favor of the syllable (Mehler, Dommergues, Frauenfelder, and Segui, 1981). Follow-up studies showed that speech segmentation strategies do differ between English and French (e.g., Cutler, Mehler, Norris, and Segui, 1986); regardless of types of stimuli (e.g., French, English, or nonwords), the syllable effect was always found for French speakers but it was not for English speakers. These results strongly suggested that the syllable's special function in French is language-specific.

Our present experiments also found evidence that the syllable in Korean is more basic than the phoneme, unlike English. However, the inherent limitations of this experiment were that all stimuli were real words and that the experiments tapped only a relatively high level of processing, leaving us unclear whether the syllable effect could be found with nonword stimuli, which would provide evidence for the syllable as a prelexical representation. Therefore, an interesting and desirable application, in which the present results might be further strengthened, would be to use nonword stimuli with response time as a dependent variable, while keeping a comparison between the phoneme and the syllable as similar as in the present experiments. For example, Experiment 3 showed that, when the number of mismatched phonemes between the two words was controlled, the sounds of two disyllabic words were judged more similar when mismatches were limited to one syllable only (e.g., /ku.saŋ- po.saŋ/) than when mismatches spread over both syllables (e.g., /ku.saŋ-ko.caŋ/). In other words, pairs were judged more similar when one syllable was perfectly matched than when both syllables were partially matched, with the number of matched and unmatched segments held constant. This basic design could also be used in a target-monitoring task with nonword stimuli. For example, when the stimuli are CV.CVC nonwords, response time might be faster for VC and CVC targets than V.C targets or CV.C targets. But there would be no difference expected between V.C and V.CV, if the number of target phonemes were not important, just as the number of mismatched phonemes was less significant than the number of mismatched syllables in Experiment 3 here. If the comparable results were found in a target-monitoring task with nonsense stimuli, this would provide strong evidence for the syllable not only as the basic level in the sense of Jaeger (1980a) but also as a prelexical representation in Korean.

4.3. Closing Remarks

Even without these future additions and improvements, however, we have good reason to believe on the basis of the present studies (and their precursors) that (i) the Korean syllable is left-branching rather than right-branching (i.e., the body is a more viable phonological unit than the rime) and (ii) that the syllable is a better candidate for a basic phonological unit than the phoneme is, both quite unlike the English case. Moreover, that these results have been confirmed using two widely disparate experimental techniques attests not only to their underlying validity but also to the value of the experimental approach to the answering of otherwise quite indeterminate theoretical questions.

Notes to Chapter IV

[1] Allowances must be made in both Korean and Japanese for syllables beginning with clusters of the type Cy (i.e., consonant followed by the glide /y/).

[2] The sequence Vy+sonorant can only be followed by [+coronal] consonant (e.g., /paynt/ or /faynd/ but not */paymp/ or */payŋk/), whereas the sequence V+sonorant does not have such constraint. (Selkirk, 1982: 351)

REFERENCES

AHN, S. C. 1988. A revised theory of syllabic phonology. *Ene* (Language) **13**, 333-362. Korean Linguistics Society, Seoul, Korea.

BAKER, W. J., G. D. PRIDEAUX, and B. L. DERWING. 1973. Grammatical properties of sentences as a basis for concept formation. *Journal of Psycholinguistic Research* **2**, 201-220.

BEINERT, R. A., and B. L. DERWING. 1993. Segment, rime, syllable, tier or root? Evidence from global sound similarity judgments in Arabic. *Proceedings of the Eastern States Conference on Linguistics*, Ithaca: Department of Modern Languages and Linguistics, Cornell University.

BENDRIEN, T. A. 1992. *Sound similarity judgements in English CVC's.* B.A. honor's thesis, University of Alberta.

BOLTON, N. 1977. *Concept formation.* Pergamon Press.

BOURNE, JR., L. E. 1966. *Human conceptual behavior.* Boston: Allyn and Bacon.

BOURNE, JR., L. E., B. R. EKSTRAND, and R. L. DOMINOWSKI. 1971. *The psychology of thinking.* Prentice-Hall.

BOWEY, J. A. 1990. Orthographic onsets and rimes as functional units of reading. *Memory and Cognition* **18** (4), 419-427.

BROWMAN, C., and GOLDSTEIN, L. 1986. Towards an articulatory phonology, *Phonology Yearbook* **3**, 219-252.

BRUNER, J. S., J. J. GOODNOW, and G. A. AUSTIN. 1956. *A study of thinking.* New York: Wiley.

BYRNE, B. 1984. On teaching articulatory phonetics via an orthography. *Memory and Cognition* **12**, 181-189.

CAMPBELL, L. 1986. Testing phonology in the field. In J. J. Ohala and J. J. Jaeger (eds.), *Experimental phonology*, 163-173. Academic Press.

CHAFE, W. L. 1977. Creativity in verbalization and its implications for the nature of stored knowledge. In R. Freedle (ed.), *Discourse production and comprehension*, 41-55. Norwood, NJ: Ablex.

CHEON, S. B. 1980. Lapsus linguaeuy umwunloncek haysek (A phonological interpretation of lapsus linguae). *Ene* (Language) **5** (2), 15-32. Korean Linguistics Society, Seoul, Korea.

CHOMSKY, N., and M. HALLE. 1968. *The sound pattern of English.* New York: Harper and Row.

CHUNG, K. 1980. *Neutralization in Korean: a functional view.* Ph.D. thesis. University of Texas.

CHUNG, K. 1988. Umwunlonuy ceilonkwa woykuke palumkyoyuk (Phonological theories and the teaching of foreign language pronunciation). *Yengeyengmunhak* (English linguistics and English literature) **34**, 363-395.

CHUNG, K. 1994a. Experimental phonology. In S. J. Chang (ed.), *Current trend in modern linguistics*, 483-502. Hanshin Publishing Company, Seoul, Korea.

CHUNG, K. 1994b. *Saengseng umwunlonuy ihae* (The understanding of generative phonology). Hanshin Publishing Company, Seoul, Korea.

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

CLEMENTS, G. N., and S. J. KEYSER. 1983. *CV Phonology.* Cambridge, MA: MIT press.

COLE, R. A., R. N. HABER, and B. D. SALES. 1968. Mechanisms of aural encoding: I. Distinctive features for consonants. *Perception and Psychophysics* **3**, 281-284.

COLE, R. A., B. D. SALES, and R. N. HABER. 1969. Mechanisms of aural encoding: II. The role of distinctive features in articulation and rehearsal. *Perception and Psychophysics* **6**, 343-348.

COOK, E. D. 1987. Sai-sios (ㅅ): A mistaken identity of gemination. *Language Research* **23** (4), 595-605.

CUTLER, A. 1986. Phonological structure in speech recognition. *Phonology Yearbook* **3**, 161-178.

CUTLER, A., J. MEHLER, D. NORRIS, and J. SEGUI. 1986. The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* **25**, 385-400.

DAVIS, S. 1989. On a non-argument for the rhyme. *Journal of Linguistics* **25**, 211-217.

DEESE, J., and S. HULSE. 1967. *The psychology of learning.* New York: McGraw-Hill.

DERWING, B. L. 1973. *Transformational grammar as a theory of language acquisition: A study in the empirical, conceptual and methodological foundations of contemporary linguistics.* London: Cambridge University Press.

DERWING, B. L. 1976. Morpheme recognition and the learning of rules of derivational morphology. *The Canadian Journal of Linguistics* **21**, 38-66.

DERWING, B. L. 1979. Psycholinguistic evidence and linguistic theory. In G. D. Prideaux (ed.), *Perspectives in experimental linguistics*, 113-138. Amsterdam: John Benjamins.

DERWING, B. L. 1980. Against autonomous linguistics. In T. Perry (ed.), *Evidence and argumentation in linguistics*, 113-138. Amsterdam: John Benjamins.

DERWING, B. L. 1983. On drawing the right conclusions in psycholinguistics: Some critical areas of control in experimental and analytical practice. *Proceedings of the Ninth Annual Meeting, Berkeley Linguistics Society*, 64-73. Berkeley Linguistics Society.

DERWING, B. L. 1991. A cross-linguistic experimental investigation of syllable structure. Part I: Background and methodology. In P. Delancy and R. M. Tomlin (eds.), *Proceedings of the Third Annual Meeting of the Pacific Linguistic Conference*, 93-102. Eugene, OR: Department of Linguistics, University of Oregon.

DERWING, B. L. 1992a. A 'pause-break' task for eliciting syllable boundary judgement from literate and illiterate speakers: preliminary results for five diverse languages. *Language and Speech* **35**, 219-235.

DERWING, B. L. 1992b. Orthographic aspects of linguistic competence. In P. Downing, S. Lima, and M. Noonan (eds.), *Linguistic aspects of literacy*, 193-210. Amsterdam: John Benjamins.

DERWING, B. L., and W. J. BAKER. 1978. On the re-integration of linguistics and psychology. In R. N. Campbell and P. T. Smith (eds.), *Recent advances in the psychology of language. Part B: formal and experimental approaches*, 193-218. New York: Plenum Press.

DERWING, B. L., S. W. CHO, and H. S. WANG. 1991. A cross-linguistic experimental investigation of syllable structure: some preliminary results. *Proceedings of the Twelfth International Congress of Phonetic Sciences* **3**, 110-113. Aix-en Provence, France.

DERWING, B. L., S. W. CHO, Y. B. YOON, D. B. PARKINSON, and H. S. WANG. 1992. Syllable structure experiments in Arabic, Korean, and Taiwanese. *Proceedings of the Third International Symposium on Language and Linguistics* **2**, 720-734. Chualalongkorn University, Bangkok, Thailand.

DERWING, B. L., and M. L. DOW. 1987. Orthography as a variable in psycholinguistic experiments. In P. E. Luelsdorff (ed.), *Orthography and phonology*, John Benjamins, 171-186.

DERWING, B. L., M. L. DOW, and T. M. NEAREY. 1987. Experimental evidence for phonological units. *Proceedings of the Eleventh International Congress of Phonetic Sciences* **2**, 249-252. Tallin, Estonia.

DERWING, B. L., M. L. DOW, and T. M. NEAREY. 1988. Experimenting with syllable structure. In J. Powers and K. de Jong (eds.), *Proceedings of the Eastern States Conference on Linguistics*, 83-94. Columbus, OH: The Ohio State University.

DERWING, B. L., and T. M. NEAREY. 1981. On the perceptibility of sub-phonemic differences: The *tough-duck* experiment. *Journal of the Atlantic Provinces Linguistic Association* **3**, 29-40.

DERWING, B. L., and T. M. NEAREY. 1986. Experimental phonology at the University of Alberta. In J. J. Ohala and J. J. Jaeger (eds.), *Experimental phonology*, 187-209. Academic Press.

DERWING, B. L., and T. M. NEAREY. 1991. The 'vowel-stickiness' phenomenon: Three sources of experimental evidence. *Proceedings of the International Conference on Phonetic Sciences* **3**, 210-213.

DERWING, B. L., and T. M. NEAREY. 1994. Sound similarity judgments and segment prominence: a cross-linguistic study, *Proceedings of the International Conference on Spoken Language* Processing, 351-354.

DERWING, B. L., and T. M. NEAREY. In press. On the status of the segment in four typologically diverse languages: results from global sound similarity judgments. *Proceedings of the International Conference on Phonetic Sciences.*

DERWING, B. L., T. M. NEAREY, R. A. BEINERT, and T. A. BENDRIEN. 1992. On the role of the segment in speech processing by human listeners: Evidence from speech perception and from global sound similarity judgments. *Proceedings of the International Conference on Spoken Language Processing*, 289-292.

DERWING, B. L., T. M. NEAREY, and M. L. DOW. 1986. On the phoneme as the unit of the second articulation. *Phonology Yearbook* **3**, 45-69.

DERWING, B. L., D. B. PARKINSON, and R. A. BEINERT. In press. Experimental investigations of Arabic syllable structure. *Perspective on Arabic Linguistics*, **3**

DERWING, B. L., and H. S. WANG. In press. Concept formation as a tool for the investigation of phonological units. *Proceedings of the International Conference on Phonetic Sciences.*

DERWING, B. L., Y. B. YOON, and S. W. CHO. 1993. The organization of the Korean syllable: Experimental evidence. In P. M. Clancy (ed.), *Japanese and Korean Linguistics* **2**, 223-238. Stanford: Center for the Study of Language and Information, Stanford University.

DOMINOWSKI, R. L. 1970. Concept attainment. In M. H. Marx (ed.), *Learning: Interactions*, 152-191. New York: Macmillan.

DOW, M. L. 1987. *On the psychological reality of sub-syllabic units.* Ph.D. thesis, University of Alberta.

DOW, M. L., and B. L. DERWING. 1989. Experimental evidence for syllable-internal structure. In R. Corrigan, F. Eckman, and M. Noonan (eds.), *Linguistic Categorization*, 81-92. New York: John Benjamins.

EHRI, L., and L. WILCE. 1980. The influence of orthography on readers' conception of the phonemic structure of words. *Applied Psycholinguistics* **1**, 371-385.

EDDINGTON, D. S. 1993. *Psychological reality and linguistically significant generalizations: Experimental evidence from Spanish*. Ph.D. thesis, University of Texas.

FALLOWS, D. 1981. Experimental evidence for English syllabification and syllable structure. *Journal of Linguistics* **17**, 309-317.

FOSS, D. J., and D. A. SWINNEY. 1973. On the psychological reality of the phoneme: perception, identification, and consciousness. *Journal of Verbal Learning and Verbal Behavior* **12**, 246-257.

FOWLER, C. A., R. TREIMAN, and J. GROSS. 1993. The structure of English syllables and polysyllables. *Journal of Memory and Language* **32**, 115-140.

FROMKIN, V. 1971. The non-anomalous nature of anomalous utterances. *Language* **47**, 27-52.

FUDGE, E. 1969. Syllables. *Journal of Linguistics* **5**, 253-286.

FUDGE, E. 1987. Branching structure within the syllable. *Journal of Linguistics* **23**, 359-377.

FUDGE, E. 1989. Syllable structure: A reply to Davis. *Journal of Linguistics* **25**, 219-220.

GIM, C. G. 1987. A study on syllable structure and some processes in its nucleus in Korean. *Mal* (Speech) **12**, 25-69. Korean Language Institute, Yonsei University, Seoul, Korea.

GOYVAERTS, D. L. 1981. *Phonology in the 1980's*. E. Story-Scientia, Ghent, Belgium.

GREENBERG, J. H., and J. J. JENKINS. 1964. Studies in the psychological correlates of the sound system of American English I & II. *Word* **20**, 157-177.

GREENBERG, J. H., and J. J. JENKINS. 1966. Studies in the psychological correlates of the sound system of American English III & IV. *Word* **22**, 207-242.

HARRIS, M. L. and C. W. HARRIS. 1973. *A Structure of concept attainment abilities*. Wisconsin Monograph Series. Wisconsin Research and Development Center for Cognitive Learning.

HAN, M. S., and R. S. WEITZMAN. 1970. Acoustic features of Korean /P,T,K/, /p,t,k/ and /p^h,t^h,k^h/. *Phonetica* **22**, 112-128.

HEALY, A. F., and J. E. CUTTING. 1976. Units of speech perception: phoneme and syllable. *Journal of Verbal Learning and Verbal Behavior* **15**, 73-83.

HOCKETT, C. F. 1967. Where the tongue slips, there slip I. In V. Fromkin (ed.), *Speech errors as linguistic evidence*, 120-131. The Hague: Mouton.

HOGG, H., and C. B. MCCULLY. 1987. *Metrical phonology: A coursebook*. Cambridge: Cambridge University Press.

HOMBERT, J. -M. 1986. Word games: some implications for analysis of tone and other phonological constructs. In J. J. Ohala and J. J. Jaeger (eds.), *Experimental phonology*, 175-186. Academic Press.

HOOPER, J. B. 1972. The syllable in phonological theory. *Language* **48**, 525-540.

HOOPER, J. B. 1976. *An introduction to natural generative phonology*. New York: Academic Press.

HUNT, E. B. 1962. *Concept learning*. John Wiley and Sons.

INGRAM, J. C. L. 1980. Perceptual dimensions of phonemic recognition. In G. D. Prideaux, B. L. Derwing, and W. J. Baker (eds.), *Experimental linguistics: Integration of theories and applications*, 273-291. Ghent: E. Story-Scientia.

JAEGER, J. J. 1980a. *Categorization in phonology: An experimental approach*. Ph.D. thesis, University of California, Berkeley.

JAEGER, J. J. 1980b. Testing the psychological reality of phonemes. *Language and Speech* **23**, 233-253.

JAEGER, J. J. 1986a. Concept formation as a tool for linguistic research. In J. J. Ohala and J. J. Jaeger (eds.), *Experimental phonology*, 211-237. Academic Press.

JAEGER, J. J. 1986b. On the acquisition of abstract representations for English vowels. *Phonology Yearbook* **3**, 71-97.

JAEGER, J. J. and J. J. OHALA. 1984. On the structure of phonetic categories. *Proceedings of the Tenth Annual Meeting, Berkeley Linguistics Society*, 15-26. Berkeley Linguistics Society.

KANG, O. M. 1992. *Korean prosodic phonology*. Ph.D. thesis, University of Washington.

KANG, Y. S. 1991. *Phonology of consonant-vowel interaction: with special reference to Korean and dependency phonology*. Ph.D. thesis, University of Illinois.

KIM, C. W. 1965. On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word* **21**, 339-359.

KIM, C. W. 1970. A theory of aspiration. *Phonetica* **21**, 107-116.

KIM, J. H. 1990. *Umcel umwunlon* (Syllabic phonology). Hanshin Publishing Company, Seoul, Korea.

KIM, J. M. 1986. *Phonology and syntax of Korean morphology*. Ph.D. thesis, University of Southern California.

KIM, Y. S. 1984a. *Aspects of Korean morphology*. Ph.D. thesis, University of Texas.

KIM, Y. S. 1984b. On the treatment of consonant cluster reduction. *Language Research*, **20** (4), 345-366.

KIM-RENAUD, Y. K. 1978. The syllable in Korean phonology. In C. W. Kim (ed.), *Papers in Korean linguistics*, 85-98. Columbia, SC: Hornbeam Press.

KIM-RENAUD, Y. K. 1986. *Studies in Korean linguistics*. Hanshin Publishing Company, Seoul, Korea.

KIM-RENAUD, Y. K. 1991. *Korean consonantal phonology*. Hanshin Publishing Company, Seoul, Korea.

KUBOZONO, H. 1989. The mora and syllable structure in Japanese: evidence from speech errors. *Language and Speech* **32**, 249-278.

LADEFOGED, P. 1975a. *A course in phonetics*. Harcourt Brace Jovanovich, New York.

LADEFOGED, P. 1975b. *Three areas of experimental phonetics*. Oxford University Press.

LADEFOGED, P. 1980. What are linguistic sounds made of? *Language* **56**, 485-502.

LAKOFF, G. 1982. Categories: An essay in cognitive linguistics. *Linguistics in the morning calm*, 139-193. Hanshin Publishing Company, Seoul, Korea.

LEE, B. G. 1982. A well-formedness condition on syllable structure. *Linguistics in the morning calm*, 489-506. Hanshin Publishing Company, Seoul, Korea.

LEE, S. O. 1992. Graphical ingenuity in the Korean writing system. *Paper presented at the Department of Anthropology*, University of Alberta, Canada.

LEE, K. S. 1993. *Umcelkucowa Umwunwonli* (Syllable structure and phonological principles). Hanshin Publishing Company, Seoul, Korea.

LEVITT, A., A. F. HEALY, and D. W. FENDRICH. 1991. Syllable-internal structure and the sonority hierarchy: Differential evidence from lexical decision, naming, and reading. *Journal of Psycholinguistic Research* **20**, 337-363.

LISKER, L., and A. ABRAMSON. 1964. A cross-language study of voicing of initial stops: Acoustic measurements. *Word* **20**, 384-422.

LINNEL, P. 1979. *Psychological reality in phonology: A theoretical study*. Cambridge: Cambridge University Press.

LOTZ, J., A. S. ABRAMSON, L. J. GERSTMAN, F. INGEMANN, and W. J. NEMSER. 1960. The perception of English stops by speakers of English, Spanish, Hungarian, and Thai: a tape-cutting experiment. *Language and Speech* **3**, 71-77.

LORCH, JR., R. F., and J. L. MYERS. 1990. Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology* **16**, 149-157.

MADDIESON, I. 1992. The structure of segment sequences. *Proceedings of the International Conference on Spoken Language Processing* (addendum), 1-4.

MALMBERG, B. 1963. *Phonetics*. New York: Dover Publications.

MARTIN, S. E. 1951. Korean phonemics. *Language* **27**, 519-533.

MARTIN, S. E. 1954. *Korean morphophonemics*. Baltimore: Linguistic Society of America.

MARTIN, S. E. 1982. Features, markedness, and order in Korean phonology. *Linguistics in the morning calm*, 601-617. Hanshin Publishing Company, Seoul, Korea.

MANN, V. A. 1986a. Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English "l" and "r". *Cognition* **24**, 169-196.

MANN, V. A. 1986b. Phonological awareness: The role of reading experience. *Cognition* **24**, 65-92.

MCCAWLEY, J. D. 1986. Today the world, tomorrow phonology. *Phonology Yearbook* **3**, 27-43.

MCNEILL, D., and K. LINDIG. 1973. The perceptual reality of phonemes, syllables, words, and sentences. *Journal of Verbal Learning and Verbal Behavior* **12**, 419-430.

MEHLER, J., J. DOMMERGUES, U. FRAUENFELDER and J. SEGUI. 1981. The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior* **20**, 298-305.

MEYER, A. S. 1991. The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language* **30**, 69-89.

MILLER, G. A., and P. NICELY. 1955. An analysis of some perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* **27**, 338-352.

MOON, Y. S. 1988. The syllable in phonological theory. *Enehak* (Linguistics) **9**, 45-73. Korean Linguistics Society, Seoul, Korea.

MORAIS, J., BERTELSON, P., CARY, L., and ALEGRIA, J. 1986. Literacy training and speech segmentation. *Cognition* **24**, 45-64.

MORAIS, J., CARY, L., ALEGRIA, J., and BERTELSON, P. 1979. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* **7**, 323-331.

NEAREY, T. M. 1990. The segment as a unit of speech perception. *Journal of Phonetics* **18**, 347-373.

NELSON, D. L., and L. D. NELSON. 1970. Rated acoustic (articulatory) similarity for word pairs varying in number and ordinal position of common letters. *Psychonomic Science* **19**, 81-82.

OHALA, J. J. 1974. Experimental historical phonology. In J. M. Anderson and C. Jones (eds.), *Historical linguistics* **2**, 353-389. Amsterdam: North-Holland.

OHALA, J. J. 1981. The phonological end justifies any means. In S. Hattori and K. Inoue, (eds.), *Proceedings of the Thirteenth International Congress of Linguistics*, 232-243.

OHALA, J. J. 1986. Consumer's guide to evidence in phonology. *Phonology Yearbook* **3**, 3-26.

OHALA, J. J. 1987. Experimental phonology. *Proceedings of the Annual Meeting, Berkeley Linguistic Society* **13**, 207-222. Berkeley Linguistics Society.

OHALA, J. J., and J. J. JAEGER. 1986a. *Experimental phonology*. Academic Press.

OHALA, J. J., and J. J. JAEGER. 1986b. Introduction. In J. J. Ohala and J. J. Jaeger (eds.), *Experimental Phonology*, 1-12. Academic Press.

OHALA, J. J. and H. KAWASAKI. 1984. Phonetics and prosodic phonology. *Phonology Yearbook* **1**, 113-127.

OHALA, J. J., and M. OHALA. 1986. Testing hypothesis regarding the psychological manifestation of morpheme structure constraints. In J. J. Ohala and J. J. Jaeger (eds.), *Experimental phonology*, 239-252. Academic Press.

OHALA, M. 1994. Experiments on the syllable in Hindi. *Proceedings of the International Conference on Spoken Language Processing* **2**, 487-490.

OTAKE, T., H. HATANO, A. CUTLER, and J. MEHLER. 1993. Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language* **32**, 258-278

PALLIER, C., S-C NURIA, T. FELGUERA, A. CHRISTOPHE, and J. MEHLER. 1993. Attentional allocation within the syllable structure of spoken words. *Journal of Memory and Language* **32**, 373-389.

PETERS, R. W. 1963. Dimensions of perception of consonants. *Journal of the Acoustical Society of America* **35**, 1985-1989.

PRIDEAUX, G. D. 1980. In rejection of autonomous linguistics. *Lingua* **50**, 245-247.

PRIDEAUX, G. D., B. L. DERWING, and W. J. BAKER. 1980. *Experimental linguistics: Integration of theories and applications*. Ghent: E. Story-Scientia.

PRINZMETAL, W., R. TREIMAN, and S. H. RHO. 1986. How to see a reading unit. *Journal of Memory and Language* **25**, 461-475.

READ, C., Y-F. ZHANG, H-Y. NIE, and B-Q. DING. 1986. The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition* **24**, 31-44.

ROSCH, E. 1973a. Natural categories. *Cognitive Psychology* **4**, 328-350.

ROSCH, E. 1973b. On the internal structure of perceptual and semantic categories. In T. E. Moore (ed.), *Cognitive development and the acquisition of language*, 111-144. New York: Academic Press.

ROSCH, E. 1978. Principles of categorization. In E. Rosch and B. B. Lloyd (eds.), *Cognition and categorization*, 27-48. Hillsdale, NJ: Erlbaum.

ROSCH, E., C. B. MERVIS, W. D. GREY, D. M. JOHNSON, and P. BOYES-BRAEM. 1976. Basic objects in natural categories. *Cognitive Psychology* **8**, 382-439.

RUNQUIST, W. N. 1968a. Formal intralist similarity in paired-associate learning. *Journal of Experimental Psychology* **78**, 634-641.

RUNQUIST, W. N. 1968b. Functions relating stimulus similarity to performance in paired associate learning with a variety of materials. *Journal of Verbal Learning and Verbal Behavior* **7**, 549-553.

RUNQUIST, W. N. 1968c. Rated similarity of high m CVC trigrams and low m CVC trigrams. *Journal of Verbal Learning and Verbal Behavior* **7**, 967-968.

RUNQUIST, W. N., and P. A. JOINSON. 1968. Rated similarity of low association value CVC trigrams. *Journal of Verbal Learning and Verbal Behavior* **7**, 317-320.

SAKAMOTO, T. 1980. Reading of Hiragana. In J. F. Kavanagh and R. L. Venezky (eds.), *Orthography, reading, and dyslexia*, 15-24. Baltimore: University Park Press.

SALES, D. B., R. A. COLE, and R. N. HABER. 1969. Mechanisms of aural encoding: V. Environmental effects of consonants on vowel encoding. *Perception and Psychophysics* **6**, 361-365.

SANTA, C. M. 1976-7. Spelling patterns and the development of flexible word recognition strategies. *Reading Research Quarterly* **12**, 125-144

SANTA, J., C. SANTA, and E. SMITH. 1977. Units of word recognition: Evidence for the use of multiple units. *Perception and Psychophysics* **22**, 585-591.

SAPIR, E. 1925. Sound patterns in language. *Language* **1**, 37-51.

SAVIN, H. B., and T. G. BEVER. 1970. The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior* **9**, 293-302.

SELKIRK, E. O. 1982. The syllable. In H. van der Hulst and N. Smith (eds.), *The structure of phonological representations, Part II*, 337-383. Dordrecht: Foris.

SHATTUCK-HUFNAGEL, S. 1983. Sublexical units and suprasegmental structure in speech production planning. In P. MacNeilage (ed.), *The production of speech*. Springer.

SOHN, H. S. 1987. *Underspecification in Korean phonology*. Ph.D. thesis, University of Illinois.

STEMBERGER, J. P. 1983. *Speech errors and theoretical phonology*. Bloomington: Indiana University Linguistic Club.

TAFT, M. 1979. Lexical access via an orthographic code: The bas[illegible]c orthographic syllable structure (BOSS) [illegible] *of Verbal Learning and Verbal Behavior* **18**, 21-39.

TOGERSON, W. S. 1[illegible] [illegible]nensional scaling of similarity. *Psychometrika* **30** (4), 379-393.

TREIMAN, R. 198[illegible] [illegible]e of spoken sylables: evidence from novel word games. *Cognition* [illegible]5, 4[illegible]

TREIMAN, R. 1984. On the st[illegible]us 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. 19[illegible] The internal structure of the syllable. In G. N. Carlson and M. K. Tanenhaus (eds.), *Linguistic structure in language processing*, 27-52.

TREIMAN, R. 1992. The role of intrasyllabic units in learning to read and spell. In P. B. Gough, L. C. Ehri, and R. Treiman (eds.), *Reading acquisition*, 153-198. New Jersey: Lawrence Erlbaum Associates, Inc.

TREIMAN, R., and J. BARON. 1981. Segmental analysis ability: development and relation to reading ability. In G. E. Mackinson and T. G. Waller (eds.), *Reading research: Advances in theory and practice* **3**, 153-198. New York: Academic Press.

TREIMAN, R., and A. M. BREAUX. 1982. Common phoneme and overall similarity relations among spoken syllables: Their use by children and adults. *Journal of Psycholinguistic Research* **2**, 569-598.

TREIMAN, R., and C. DANIS. 1988. Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology* **14**, 145-152.

TREIMAN, R., J. GROSS, and A. Cwikhel-Glavin. 1992. The syllabification of /s/-clusters in English. *Journal of Phonetics* **20**, 383-402.

TREIMAN, R., and A. ZUKOWSKI. 1990. Toward an understanding of English syllabification. *Journal of Memory and Language* **29**, 66-85.

VENNEMANN, T. 1974. Words and syllables in natural generative grammar. In A. Bruck, R. A. Fox, and M. W. La Galy (eds.), *Papers from the Parasession on Natural Generative Phonology*, 346-374. Chicago: Chicago Linguistic Society.

VENNEMANN, T. 1988a. *Preference laws for syllable structure and the explanation of sound change*. Berlin: Mouton de Gruyter.

VENNEMANN, T. 1988b. The rule dependence of syllable structure. In C. Duncan-Rose and T. Vennemann (eds.), *On language: rhetorica, phonologica, syntactica: A festschrift for Robert P. Stockwell from his friends and collegues* 257-283. London: Routledge.

VINACKE, W. E. 1951. The investigation of concept formation. *Psychological Bulletin* **48**, 1-31.

VITZ, P. C., and B. S. WINKLER. 1973. Predicting the judged similarity of sound of English words. *Journal of Verbal Learning and Verbal Behavior* **12**, 373-388.

WANG, H. S. 1995. *Experimental studies in Taiwanese phonology*. The Crane Publishing Company, Taipei, Taiwan.

WANG, H. S., and B. L. DERWING. 1993. Is Taiwanese a 'body' language? *Toronto Working Papers in Linguistics*, 679-694. University of Toronto, Ottawa.

WICKELGREN, W. A. 1965. Distinctive features and errors in short-term memory for English vowels. *Journal of the Acoustical Society of America* **38**, 583-588.

WICKELGREN, W. A. 1966. Distinctive features and errors in short-term memory for English consonants. *Journal of the Acoustical Society of America* **39**, 388-398.

WIEBE, G. E. 1992. *The psychological reality of phonological units for speakers of an unwritten language*. Ph.D. thesis, University of Alberta.

WIEBE, G. E., and B. L. DERWING. 1994. A forced-choice word-blending task for testing intra-syllabic break points in English, Korean and Taiwanese. In M. J. Powell (ed.), *The Twenty-first LACUS Forum*, 142-152.

WINER, B. F., D. R. BROWN, and K. M. MICHELS. 1991. *Statistical principles in experimental design*. New York: McGraw-Hill.

YOON, Y. B. 1987. *Syllable, phonotactic constraints, and syllabification.* M.A. thesis, Sogang University, Seoul, Korea.

YOON, Y. B. 1990. Arguments against the universality of the onset/rime division. *Sogang Working Papers in Linguistics* **4**, 93-104. Sogang University, Seoul, Korea.

YOON, Y. B. 1992. Closure duration and the three series of Korean stops: An experimental study. In S. W. Cho, and H. H. Lee (eds.), *Festshrift for Taeok Kim*, 487-497. Hankuk Publishing Company, Seoul, Korea.

YOON, Y. B. 1994a. CV as a phonological unit in Korean. *Proceedings of the International Conference on Spoken Language Processing* **2**, 483-486.

YOON, Y. B. 1994b. Syllable saliency in Korean: Evidence from the sound similarity judgement task. *Papers in Experimental and Theoretical Linguistics* **2**, 76-87. Department of Linguistics, University of Alberta.

YOON, Y. B., and B. L. DERWING. 1994. Sound similarity judgements for Korean words by Korean and English speakers. *Proceedings of the 1994 Annual Conference of the Canadian Linguistic Association*, 657-666.

YOON, Y. B., and B. L. DERWING. In press. Syllable saliency in the perception of Korean words. *Proceedings of the International Conference on Phonetic Sciences.*

APPENDIX 1. STIMULI IN EXPERIMENT 1

XXX-XXX = NO MATCHES

pan-mət	təm-nap	mot-pun	nup-tom
pan-təm	təm-pan	mot-nup	nup-mot

CXX-CXX = ONSET MATCHES

pan-pət	təm-tap	mot-mun	nup-nom
pan-pəm	təm-tan	mot-mup	nup-not

XVX-XVX = VOWEL MATCHES

pan-mat	təm-nəp	mot-pon	nup-tum
pan-tam	təm-pən	mot-nop	nup-mut

XXC-XXC = CODA MATCHES

pan-mən	təm-nam	mot-put	nup-top
pan-tən	təm-pam	mot-nut	nup-mop

CVX-CVX = BODY MATCHES

pan-pat	təm-təp	mot-mon	nup-num
pan-pam	təm-tən	mot-mop	nup-nut

XVC-XVC = RIME MATCHES

pan-man	təm-nəm	mot-pot	nup-tup
pan-tan	təm-pəm	mot-not	nup-mup

CXC-CXC = MARGIN MATCHES

pan-pən	təm-tam	mot-mut	nup-nop
pan-pən	təm-tam	mot-mut	nup-nop

CVC-CVC = CONTROLS

pan-pan	təm-təm	mot-mot	nup-nup
pan-pan	təm-təm	mot-mot	nup-nup

APPENDIX 2. STIMULI IN EXPERIMENT 2

(NOTE: Two pairs in the shaded columns are the repetitions of the same pair)

CVX^{P}-CVX^{P} = BODY MATCHES WITH [PLACE] MISMATCH IN THE CODA

pan-pam	təm-tən	mot-mop	nup-nut
pan-pam	təm-tən	mot-mok	nup-nuk

CVX^{m}-CVX^{m} = BODY MATCHES WITH [MANNER] MISMATCH IN THE CODA

pan-pat	təm-təp	mot-mon	nup-num
pan-pal	təm-təp	mot-mol	nup-num

CVX^{B}-CVX^{B} = BODY MATCHES WITH [PLACE] AND [MANNER] MISMATCH IN THE CODA

pan-pap	təm-tək	mot-mom	nup-nun
pan-pak	təm-təl	mot-mom	nup-nul

$X^{P}VC$-$X^{P}VC$ = RIME MATCHES WITH [PLACE] MISMATCH IN THE ONSET

pan-tan	təm-pəm	mot-not	nup-mup
pan-kan	təm-kəm	mot-not	nup-mup

$X^{m}VC$-$X^{m}VC$ = RIME MATCHES WITH [MANNER] MISMATCH IN THE ONSET

pan-man	təm-nəm	mot-pot	nup-tup
pan-man	təm-səm	mot-pot	nup-sup

$X^{B}VC$-$X^{B}VC$ = RIME MATCHES WITH [PLACE] AND [MANNER] MISMATCH IN THE ONSET

pan-nan	təm-məm	mot-tot	nup-pup
pan-san	təm-məm	mot-sot	nup-kup

X^PVX^P-X^PVX^P = VOWEL MATCHES WITH [PLACE] MISMATCH IN IN THE ONSET AND THE CODA

pan-tam	təm-pən	mot-nop	nup-mut
pan-kam	təm-kən	mot-nok	nup-muk

X^mVX^m-X^mVX^m = VOWEL MATCHES WITH [MANNER] MISMATCH IN THE ONSET AND THE CODA

pan-mat	təm-nəp	mot-pon	nup-tum
pan-mal	təm-səp	mot-pol	nup-sum

X^BVX^B-X^BVX^B = VOWEL MATCHES WITH [PLACE] AND [MANNER] MISMATCH IN THE ONSET AND THE CODA

pan-mak	təm-mək	mot-tom	nup-kun
pan-sak	təm-məl	mot-kom	nup-sul

APPENDIX 3. STIMULI IN EXPERIMENT 3

XV.CVC-XV.CVC = C_1 MISMATCHES

ku.saŋ-pu.saŋ	mu.cən-nu.cən	so.kak-co.kak	ko.tok-to.tok

CV.XVC-CV.XVC = C_2 MISMATCHES

co.saŋ-co.caŋ	to.caŋ-to.saŋ	so.kak-so.pak	so.tok-so.pok

CV.CVX-CV.CVX = C_3 MISMATCHES

co.səŋ-co.sək	ca.cəŋ [illegible]	so.kam-so.kaŋ	co.kam-co.kan

XX.CVC-XX.CVC = C_1V_1 MISMATCHES

ku.saŋ-po.saŋ	ca.cəŋ-sə.cəŋ	ku.pak-to.pak	ca.kaŋ-sə.kaŋ

CX.XVC-CX.XVC = V_1C_2 MISMATCHES

ku.saŋ-ko.caŋ	ko.cəŋ-ku.səŋ	so.kak-so.pak	ku.pak-ko.kak

CV.XXC-CV.XXC = C_2V_2 MISMATCHES

co.cəŋ-co.san	ca.cəŋ-ca.san	to.tuk-to.pok	so.tok-so.kuk

CV.CXX-CV.CXX = V_2C_3 MISMATCHES

ku.saŋ-ku.sək	co.cak-co.cəŋ	ko.tok-ko.kuk	ko.taŋ-ko.tok

XX.XVC-XX.XVC = $C_1V_1C_2$ MISMATCHES

su.taŋ-co.cak	ca.səŋ-sə.cəŋ	to.tuk-pu.kuk	su.tan-co.kan

CX.XXC-CX.XXC = $V_1C_2V_2$ MISMATCHES

ku.saŋ-ko.cəŋ	to.cən-tu.san	so.kak-su.tək	ko.tok-ku.kuk

CC.XXX-CC.XXX = $C_2V_2C_3$ MISMATCHES

po.cən-po.saŋ	to.cən-to.saŋ	ko.tok-ko.kuŋ	su.kən-su.tam

XX.XXX-XX.XXX = ALL MISMATCHES

ku.saŋ-po.cən	to.cən-ku.saŋ	ku.pak-to.səŋ	co.kən-su.taŋ

CV.CVC-CV.CVC = CONTROLS

ku.saŋ-ku.saŋ	ca.cən-ca.cən	so.kak-sokak	ko.tok-ko.tok

APPENDIX 4. STIMULI IN PRETEST

TARGETS = a.a

/a.a/

paŋ.hak	hap.p^{h}an	caŋ.kap	kan.caŋ	caŋ.pal
kam.saŋ	cham.cak	pan.tal	tam.p^{h}an	c^{h}aŋ.cak

/Ca.Ca/

pan.pal	pan.pak	caŋ.cak	san.sam	kam.kak
tam.taŋ	yaŋ.yak	t^{h}am.t^{h}ak	maŋ.mak	yəŋ.yək

/aC.aC/

maŋ.saŋ	haŋ.saŋ	taŋ.caŋ	pal.tal	pan.c^{h}an
p^{h}an.tan	kan.tan	kaŋ.taŋ	saŋ.caŋ	c^{h}an.pan

DISTRACTORS

/u.u/

cuŋ.sun	c^{h}uŋ.pun	kun.cuŋ	cuŋ.kuk	c^{h}ul.cuŋ
mun.mul	pul.sun	mul.pul	kuk.kun	kuŋ.cuŋ

/u.a/

c^{h}ul.pal	sun.pak	cuŋ.kan	tun.t^{h}ak	pun.caŋ
sun.kan	hun.paŋ	cuŋ.tan	hun.caŋ	pun.pal

/a.u/

san.pul	caŋ.kun	kaŋ.puk	hap.suk	pal.kul
han.kuk	taŋ.kuk	p^{h}an.kuk	tan.kun	nam.puk

APPENDIX 5. STIMULI IN EXPERIMENT 4A

TARGETS = /ka/

TI = Target in Initial syllable

kaŋ.co	kam.cəŋ	kaŋ.to	kam.toŋ	kaŋ.yən
kan.səp	kan.ho	kan.cəp	kan.so	kan.cu

TF = Target in Final syllable

sæŋ.kaŋ	ye.kam	so.kam	si.kan	cu.kan
cuŋ.kan	mi.kan	su.kap	po.kaŋ	pu.kaŋ

TC = Target Common in Experiments 4A and 4B

kak.pyəl	kak.to	kak.səŋ	kak.coŋ	kak.ca
sim.kak.	sæŋ.kak	co.kak	c^həŋ.kak	si.kak

DISTRACTORS

D0 = Neither /k/ nor /a/

toŋ.yə	cəŋ.sin	no.læ	cəl.tæ	yəŋ.hon
sin.mun	næŋ.su	tæ.hwæ	si.səl	səl.myəŋ

D1 = Onset /k/

koŋ.toŋ	kyəŋ.həm	ki.lim	kəm.so	kɨm.ci
so.kim	poŋ.kɨp	ɨy.kyən	coŋ.kyo	si.kol

D2 = Rime /ak/

cak.sim	pak.su	mak.cuŋ	cak.səŋ	mak.næ
si.cak	so.pak	co.cak	toŋ.cak	su.pak

APPENDIX 6. STIMULI IN EXPERIMENT 4B

TARGETS = /ak/

TI = Target in Initial syllable

cak.sim	pak.su	mak.cuŋ	mak.yən	cak.əp
cak.pyəl	cak.səŋ	cak.yoŋ	cak.cən	hak.kyo

TF = Target in Final syllable

si.cak	su.pak	co.cak	ho.pak	so.pak
to.pak	mi.kan	su.hak	yu.hak	tæ.hak

TC = Target Common in Experiments 4A and 4B

kak.pyəl	kak.to	kak.səŋ	kak.coŋ	kak.ca
sim.kak.	sæŋ.kak	co.kak	c^{h}əŋ.kak	si.kak

DISTRACTORS

D0 = Neither /k/ nor /a/

toŋ.yo	cəŋ.sin	no.læ	cəl.tæ	yəŋ.hon
sin.mun	næŋ.su	tæ.hwæ	si.səl	səl.myəŋ

D1 = Coda /k/

cik.cəp	sok.to	suk.ce	tok.sin	c^{h}uk.so
so.tɨk	cə.c^{h}uk	hæŋ.pok	to.tuk	yən.lak

D2 = Body /ka/

kaŋ.co	kam.toŋ	kaŋ.to	kaŋ.yən	kam.cəŋ
sæŋ.kaŋ	si.kan	ye.kam	so.kam	cu.kan

APPENDIX 7. STIMULI IN EXPERIMENT 5A

TARGETS = /k/

TI = Target in Initial position [k]

kil	kam	kaŋ	kəp	kəm
kop	kom	kol	kup	kul

TF = Target in Final position [k°]

pak	mak	pyək	mək	sək
pok	mok	sok	puk	muk

TC = Target in Middle position [g]

ma.kam	to.kul	sa.kwan	po.kəm	si.kan
ci.kim	iy.kyən	pu.kaŋ	so.kyən	si.kol

DISTRACTORS

D0 = No velars

pan	pəp	nap	non	səl
sol	cul	cim	mom	ton

D1 = Velar nasal /ŋ/

paŋ	saŋ	məŋ	cuŋ	tiŋ
so.paŋ	to.maŋ	pu.saŋ	cə.caŋ	su.cəŋ

D2 = Velar stops /k^h/ or /k'/

k^hal	k^hi	k'ol	k'um	k'ul
k^hin.cip	k^hal.nal	kwæ.cæ	k'o.li	k'ol.c'i

APPENDIX 8. STIMULI IN EXPERIMENT 5B

TARGETS = /kak/

TI = Target in Initial position (kak.-)

kak.pyəl	kak.to	kak.ca	kak.sə	kak.səŋ
kak.coŋ	kak.cu	kak.o	kak.un	kak.pon

TF = Target in Final position (-.kak)

sim.kak	sæŋ.kak	ye.kak	cəŋ.kak	so.kak
sam.kak	ho.kak	co.kak	cʰəŋ.kak	si.kak

TC = Target in Middle position (-.kak.-)

ta.kak.to	sam.kak.ci	kyəŋ.kak.sim	sam.kak.hyəŋ	so.kak.caŋ
co.kak.ka	sa.kak.mo	hwan.kak.ce	ci.kak.sæŋ	sən.kak.ca

DISTRACTORS

D0 = Neither /k/ nor /a/

toŋ.yo	cəŋ.sin	no.læ	cəl.tæ	yəŋ.hon
sin.mun	næŋ.su	ma.nil	tæ.hwæ	si.səl

D1 = Either /k/ or /a/

koŋ.toŋ	ki.lim	cəŋ.kyo	sən.cʰaŋ	cə.caŋ
ku.səŋ.won	ca.cən.kə	sik.mul.won	chəŋ.sa.cin	sən.ku.ca

D2 = Either /ka/ or /ak/

kam.sok	kaŋ.to	pan.kam	mak.tæ	caŋ.mak
kan.ho.won	si.kan.pʰyo	yən.mak.tʰan	to.pak.sa	u.su.cak

END

13-03-96

FIN