THE UNIVERSITY OF ALBERTA

The Psychological Reality of Sub-syllabic Units ©

by

Maureen Louise Dow

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF Doctor of Philosophy

IN

Psycholinguistics

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA

FALL 1987

ISBN 0-315-40865-0

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Maureen Louise Dow

~~TITLE OF THESIS:~~ The Psychological Reality of Sub-syllabic Units

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1987

Kolbenacker 14/30

Ch-8052 Zürich

Date: September 25, 1987

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled THE PSYCHOLOGICAL REALITY OF SUB-SYLLABIC UNITS, submitted by Maureen L. Dow in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Psycholinguistics.

Bruce L. Derwing
Supervisor

John T. Hogan

Frederick J. Boyne

Terrance M. Nearey

Rebecca Treiman
External Examiner

Date 13 October 1987

To my father, James Dow, who sparked my interest in Linguistics many years ago, and to my husband, Heinz Ehrensberger, who supported me through this latest effort.

Abstract

This study tests the psychological reality of sub-syllabic units using a variety of tasks. In classical segmental phonology syllables are viewed as strings of equally weighted phonemes; this stand predicts that one segment ought to be as easy to manipulate as another and that groups of segments ought to present added difficulties in proportion to the number of segments involved. Alternate models of the syllable include various kinds of intermediate units between the syllable and the segment (such as onsets, nuclei, codas, rhymes). Proponents of such models would differences in the ease of manipulation of the various units, depending on what the particular units are, where their boundaries occur, and how they are organized within the syllable.

Children and high school students were asked to count and delete various phonological units within words (Experiments 1 and 2) and their results were analysed using different models of syllable structure. An onset-rhyme model of syllable structure best accounted for the results for the children, whereas the high school students seemed to be sensitive to finer divisions, specifically the nucleus and the coda.

A substitution-by-analogy task was used to test adults on several hypothesized sub-syllabic units. Of the six units assessed, both the onset and rhyme emerged as clear units, since they were most easily manipulated, the nucleus and coda slightly less so, and the head (onset plus nucleus) and margins (onset plus coda) were the most difficult to manipulate.

The phonetic similarity judgment task of the fourth set of experiments

confirmed the psychological reality of the onset, nucleus, and coda. Predictions for similarity judgments with these units were consistently better than with a model based on a string of independent segments or a simple onset-rhyme or onset-vowel-coda structure.

The results reported here clearly support a model of syllable structure for English with a major branch or break between the onset and the rhyme and also provide considerable support for a secondary division or branch between the vowel nucleus and the coda. The reality of these sub-syllabic units indicates the need to re-evaluate traditional notions of phonological organization.

Acknowledgments

The decision to move to Switzerland after the second year of my PhD program demanded great patience and understanding from the members of my supervisory committee. Dr. Bruce Derwing, my supervisor, not only demonstrated these, he was a constant source of advice, inspiration, and support, and continually challenged me to explain what these ideas <u>really</u> meant. His ability to assimilate bits and pieces of experimental results into the broader context of linguistic theory was humbling. Dr. Terrance Nearey took on the role of co-supervisor and essentially provided the phonetic relevance of this thesis. His interest and advice on my sporadic visits kept me from giving up on the last experiment. Dr. Fred Morrison's enthusiasm about the implications for reading research encouraged me and contributed greatly to the first two experiments. I would also like to thank Dr. John Hogan and Dr. Rebecca Treiman, my external examiner, for their time and effort in reviewing my thesis in its final stages.

The further along I got on my thesis, the more I appreciated how essential colleagues are in the process. Tracey Derwing was not only always willing to argue with me on theoretical points, she helped in the design, carrying-out, and scoring of the experiments, willingly read and edited endless drafts, was a source of advice and my long-distance manager for the administrative aspects of being a student, and probably most importantly, generously provided me with a warm home away from home. It is difficult to imagine how I could have finished without her constant support and friendship.

Peter Assmann encouraged me, kept me from making some wild claims, and provided much-needed advice on statistics at a couple of crucial points when I was feeling very lost. My discussions with Sam Wang on phonology were invaluable, as were the bibliography exchanges we often made. Murray Munro willingly helped with designing experiments, scoring results, setting up equipment, running pilot studies, and developing my arguments and explanations into a coherent form. Thanks also to Janet Schwegel for the formatting and printing of this thesis and for related errands that went beyond the call of duty.

Without subjects there would have been no thesis. Thanks to the Department of Linguistics graduate students and Diane and Karen Dow for taking part in the pilot studies and to the following for taking part in the experiments: the students and staff from the Edmonton Public School Board, the students at the University of Alberta, the students in English Phonetics at the University of Zürich, the students and staff at the American School at Kilchberg, the students from the Kantonschule Hohe Promenade in Zürich, and friends and friends of friends. Special thanks are due to Dr. Spörri of the Phonetisches Labor at the University of Zürich for arranging for subjects and for the use of the University computer. Regula Krüger became my linguistics contact in Zürich, assuming the multiple roles of native speaker and instructor of Züritüütsch, research assistant, and interpreter.

Thanks are also due to the Heritage Foundation of Alberta and the Humanities and Social Sciences Research Council of Canada for support throughout my program. I also gratefully acknowledge financial assistance from the University of Alberta Faculty of Graduate Studies and Research and the Alma Mater Fund, which enabled me to attend the

Deutsche Gesellschaft für Sprachwissenschaft at Heidelberg University in February 1986 and the University of Wisconsin at Milwaukee Conference on Linguistic Categorization in April 1987, where material related to this thesis was presented.

The support, both financial and emotional, that my husband, Heinz Ehrensberger, has provided me with throughout the past four years is lovingly acknowledged. He is even happier than I am that trips to Canada for thesis work are no longer necessary.

Once again, heartfelt thanks to Tracey, Bruce, and, of course, Chloe, for their support, friendship, and generous hospitality.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

The notion that syllables consist of strings of phonemes has a long tradition and is widely accepted today. Phonemes have proven to be extremely useful constructs for the study of language, with isolation and description of a language's phonemes forming the first crucial step in field work. The International Phonetic Alphabet, based on segments, has been and will continue to be an invaluable tool with which to record speech utterances. Although phonemic or phonetic transcription is useful for the study of any language's phonology, it is particularly helpful in languages for which no orthography exists, or for which the orthography is not based on sounds. The phoneme has also proven to be a useful concept for use in the description of phonological processes and, although somewhat overshadowed by feature-based systems in recent years, phonemes are still used to describe morphophonemic alternations. As a cover term for a number of features and possible allophones, phonemes, along with allophonic rules, simplify the inventory of possible sounds in a language.

Despite the apparent value of the phoneme in much linguistic work, a number of phonological theories have recently emerged (e.g., metrical phonology, grid-based phonology, autosegmental phonology) which claim to deal with a variety of language phenomena that a model of segmental phonology cannot handle simply. For example, Japanese may be better described in terms of syllables or moras than phonemes (McCawley, 1968), English phrase and word stress has been explained with hierarchical structures (Liberman, 1975; Liberman & Prince, 1977) and Goldsmith (1976) argued for autosegmental tiers of features to describe tone languages. Although theoretically all languages can be transcribed with IPA symbols, it is very difficult in a segmental system to capture prosodic features such as stress and tone, which may overlap segments or change over the course of a single segment.

These problems aside, phonemes do seem eminently reasonable constructs with which to describe most Indo-European languages. There are a number of linguistic similarities among these languages, not the least of which are the presumed common origin and the use of an alphabetic orthography. The latter may seem irrelevant, since writing is often viewed as merely a secondary code to record spoken language, but it has also been argued that the existence of the alphabet proves the basic nature of phonemes, that is, that speech is interpreted as a sequence of phonemes and the alphabet merely reflects this fact of nature (Warren, 1982).

According to Lüdtke (1969), the circular argument of whether the alphabet suggested phonemes or whether phonemes suggested the alphabet can be resolved either by finding physical reality for phonemes or a historical (as opposed to psychological) explanation for the alphabet. In the absence of physical evidence he argued that the monogenesis of the alphabet indicated its arbitrariness and concluded that phonemes were artificial constructs. Our present-day view of separate phonemic segments may thus be an accidental result of historical developments such as the Greeks placing vowel marks between consonants rather than as part of the consonant symbol.

Lüdtke was not alone in arguing that the orthography has colored phonological representations or theory. Moskowitz (1973) claimed that

English orthography has probably imposed some regularities onto phonological representations and Jaeger (1980a; 1980b), with evidence from Japanese and English speakers, suggested that the level for conceptualizing sounds in a language is the one represented by the orthography, because that is the one most often dealt with by speakers using their language. If these linguists are correct, then phonemes may still be the basic units of the English language's sound system because the orthography reflects this level. Such reasoning, however, retains the circularity decried by Lüdtke because most English language speaking countries have high rates of literacy. Learning the alphabet may result in learning about phonemes.

Although native speaker intuitions and behavior often seem phoneme-oriented, which provides support for the reality of phonemes, knowledge of letter-sized units through orthography may influence this. It is difficult to separate this knowledge from phonological knowledge. One must look farther afield than adult literate English speakers to argue for the reality of phonological units. Of special interest is the form of phonological representations before literacy. If phonemes are basic forms of sound representation, then literacy should produce no fundamental change in phonological judgments, assuming that these are tapping underlying knowledge directly and do not depend on knowledge of spelling for responses. If phonemes are not basic, but somehow learned or imposed on phonological representations, then there should be differences among pre-literate children, adult illiterates, and adult literate speakers. The literature reviewed in the next chapter clearly indicates that important differences exist. The evidence accumulated from experimental phonology, reading research, psychology, and spontaneous speech provides some support for the psychological reality of phonemes but forms a strong case for the reality of units intermediate in size between phonemes and syllables.

Although a number of theorists have suggested that some unit larger than the phoneme may be more appropriate to describe phonological knowledge, there has been little testing of this possibility. In fact, a variety of possible models of phonological structure have been proposed. The word is a basic unit in language, but is rarely credited with being a basic sound unit. Rather, the syllable is generally labelled as such. There are a number of theoretical arguments against a syllable-based phonological representation, economy being the most notable, particularly in light of the severe restrictions on consonantal sequences that exist in most languages. Sub-syllabic units have been raised as an alternative solution. In contrast to syllables, there is not so much difficulty in terms of the number that would need to be stored and there is evidence that many combinations of phonemes function as units. It is therefore the focus of this thesis to test various models of phonological structure, from syllables down to phonemes, with different populations of speakers and different tasks.

A finding that a phonological unit other than the syllable or phoneme is psychologically real would have implications which extend beyond phonological theory. A general assumption of most current reading theories, for example, is that the phoneme is the basic unit of sound, and that it is already available to children when they first approach the reading task. If this is not the case, then the learning of reading may be unnecessarily complicated by first having to break down grapheme units into phonemes and then to 'blend' them back into words.

The following chapter is a review of the evidence that some representation other than phonemic can be real to speakers. The third chapter outlines the various models of syllable structure that have been proposed and suggests ways to test them experimentally. Chapter Four is comprised of four experiments: a unit counting task; a deletion-by-analogy task, a substitution-by-analogy task, and a phonetic similiarity judgment task. The final chapter presents conclusions and outlines how the tasks in this thesis could be extended to other languages, especially those without alphabetic orthographies.

CHAPTER 2: SOUND UNITS OF SPEECH

The thesis to be put forward here is that segmental phonology has been driven and then justified, at least in part, by prior conventions of alphabetic orthographies and that there may be more appropriate models to describe pre- or non-literate phonological knowledge. There is considerable evidence from a variety of language phenomena that a model of phonology based on a string of independent segments cannot explain. It is therefore necessary to determine whether sufficient justification exists to abandon segmental phonology entirely and, if so, to garner the evidence that already exists to establish the general form that a revised model might take. The first section of this chapter deals with the interdependence of orthography and segmental phonology. Topics ranging from experimental investigations of phonology to spelling pronunciations indicate that orthography and segmental phonology are inextricably linked in English. The second section deals with evidence that can be considered relatively uncontaminated by orthographic knowledge to try to establish what pre- or non-literate speakers' phonological representations are like. Evidence from pre-schoolers, beginning readers, reading-disabled children, illiterate adults, and speakers of languages with non-alphabetic orthographies is presented in this section. The search in the literature for a single, basic unit of speech perception is discussed in the third section. Several areas of study are relevant here. Children's phonetic perceptions differ from adults' in ways that are consistent with a new model of phonological representation which relates, in turn, to the units that appear in psychological experiments, speech errors, word games, and speech synthesis and recognition work. Some experimental results from an earlier phonological study (Dow, 1981) are also re-examined in light of the new model.

2.1 The Influence of Orthography on Phonology

The influence of orthographic knowledge has been identified in a number of experimental investigations of phonological questions, including the featural and segmental status of various English sounds and the reality of phonological rules. A variety of techniques have been used, but in most cases the investigators found that they could not explain their results in purely phonological terms in isolation from other types of knowledge.

In addition to their role in theoretical linguistics and language description, phonemes have been used to predict behaviour in language tasks. To determine the basis of sound similarity judgments, Nelson and Nelson (1970) compared predictions based on phonemic distances to judged similarity of pairs of words. Generally, predictions were very good, although there did seem to be some influence of spelling. They proposed that both the number of shared letters and phonemes influenced the ratings for at least one group of word pairs which systematically varied in both acoustic and formal (spelling) similarity, because both factors were found to be significant. Vitz and Winkler (1973) formalized a predictor of phonemic distance (Predicted Phonemic Distance or PPD) and used aural presentation in four similarity of sound studies. The correlations they reported were very good (on the order of r=-.9), but orthographic influence could not be ruled out.[1] The Vitz and Winkler replication of Nelson and

Nelson's study, using the word pairs that varied in both acoustic and formal similarity, resulted in a lower correlation (*r*=-.7), which they attributed to the focus on spelling due to visual presentation, but an aural replicate produced only a marginal improvement (*r*=-.77). It may have been that the subjects' knowldege of orthographic similarity overrode the importance of phonemic similarity in the ratings, thus lowering the correlation.

Derwing (1976) used Vitz and Winkler's (1973) PPD index as well as his own grapheme index in another similarity of sound study with the stimuli presented visually. Not surprisingly (since letters are highly correlated to phonemes), both indices yielded high correlations. Using partial correlations, Derwing found that the correlation for phoneme index controlling for graphemic similarity was much higher (*r*=.63) than that for grapheme index controlling for phonemic similarity (*r*=.34). He concluded that although orthographic similarity had an influence, the phonetic dimension was more important and he suggested that the task be repeated with an aural presentation. A high correlation (*r*=.97) was found between the mean scores for words given by subjects in the visual presentation experiment and those of an aural replicate, which suggested that both groups of subjects performed similarly; however, the possibility of an orthographic influence under both conditions was not excluded (Derwing & Nearey, 1986).

In an attempt to isolate the effects of orthography in phonological judgments, two tasks were designed to tap phonological knowledge using

[1]Since Vitz and Winker's correlation was of distance (PPD) and similarity, negative values are expected.

both aural and written presentation (reported in Dow, 1981). First, when subjects were asked to provide a rhyme for test words, they gave rhymes that matched the spelling at greater than chance levels if they had seen and heard the words (written presentation), but not if the words were only heard and not seen (aural presentation). The second task was a direct request to provide the number of 'speech sounds' for each test word. Phonologically non-controversial examples were provided to clarify the term. Overt presentation of the spelling had no apparent effect on responses, since no significant differences were found between the aural and written presentation groups. Further, a higher correlation was obtained for responses with a phoneme index (*r*=.93) than with a grapheme index (*r*=.80), and partial correlations again resulted in a wider separation between the two indices (*r*=.86 for the phoneme index, *r*=.57 for the grapheme index). It should be noted, however, that all of the correlations were statistically significant (at the .01 level), which suggests that the possibility of orthographic influence cannot be ignored. In fact, there was some evidence at the level of individual items that spelling may have influenced judgments. For example, the words cower, bower, and glower were consistently judged as having more speech sounds than their rhymes sour, dour, and scour, presumably because of the presence of the orthographic w or the sequence er in the former set (see Derwing, Nearey, & Dow, 1986).

An expanded version of the segment count task was carried out by Derwing (1985) in an attempt to evaluate the influence of instructions, presentation mode, knowledge of linguistics, and spelling. Of these four factors, only spelling was significantly related to judgments. The overall correlation of the number of speech sounds with the phoneme count for more than a hundred subjects was very high (r=.95) while the correlation with grapheme count was somewhat lower (r=.80). On analogy with a study by Ehri and Wilce (1979) for elementary school children (see below for further details), Derwing also compared the segment counts provided by adults for words with affricates spelled tch and dge to those spelled ch and ge. Significantly higher counts were provided for words spelled with the additional letter. The same phenomenon was noted for words with the velar nasal, which were consistently judged as containing more segments than words with [n]. Although it has been argued that the underlying form of the velar nasal is /ng/, which would account for higher counts than /n/, there is no corresponding argument for the affricates, which differed only in spelling. It is also questionable that the /ng/ analysis is the appropriate one, given the criticisms, historical and behavioural, reviewed by Smith (1982a; 1982b). More likely, the higher counts in these and other cases reflect the influence of orthographic knowledge.

The perception of phonetically ambiguous segments may be guided by knowledge of conventional spelling. In an investigation of adults' and children's perception of stops following /s/, for example, Fink (1974) used the very simple technique of asking three groups of subjects (grades 2 and 3 children and adults) to spell nonsense words with word-medial /s/+stop clusters which they heard spoken by a native speaker of English. He found that children responded differently than adults in their choice of representing the stop after /s/. Specifically, adults chose the voiceless stops /p t k/, even ~~when the tokens~~ they heard were voiced, whereas children preferred the voiced representation (b) for labial stops, whether the tokens were voiced or not. Fink suggested that there was a shift to spelling conventions as children got older (e.g., sp but not sb occurs word-initially in English spelling), which might result in actual shifts in perceptual categorization. Treiman (1985c) also asked children and adults to indicate what the sound following /s/ was and found an increase in voiceless responses concomitant with age. Most of the Kindergarten children showed no clear preference for voiceless or voiced stops while most Grade One children and adults preferred the voiceless stops. Treiman found that the preference for voiceless stops correlated significantly with the children's reading level (as measured by WRAT scores).

Jaeger (1980a; 1980b) used a concept formation task to test the membership of unaspirated and unreleased velar stop allophones in the category formed by the aspirated allophone [k^h]. She found that all of these allophones were categorized as /k/, despite alternate spellings (e.g., k, c, ch, q), although there was a high error rate with words spelled with q. Jaeger pointed out (p. 351) that knowledge from literacy is as real as any other knowledge and it is impossible to exclude the possibility of its influence here. Indeed, Derwing and Nearey (1986), in an investigation of the categorization of English stops following /s/, also found evidence of such an influence. They had subjects judge the similarity between stop-initial and /s/+stop-initial words and found that generally there was greater similarity between [p^h]-[p] than [b]-[p], for example. However, they found that this was only consistently true when the spelling clearly supported the voiceless interpretation. In cases where the spelling representation was not the conventional sk, sp, st (e.g., cot-squat), this similarity difference was reduced. In another study, Derwing and Nearey

(1981) found that the alveolar stops of nonsense words (but not real words) with initial /s/-stop clusters were judged as being both /t/- and /d/-like; the clear /t/ interpretation for real words might then have been because of knowledge of the words' spelling.

It may be that orthography is only used by speakers when they are asked to disambiguate cases of phonemic overlap. Ehri (1985) described an experiment she conducted with Wilce (Ehri & Wilce, in press) in which they taught second graders to read words containing alveolar flaps, while a control group said the words but did not see them. In a rhyme judgment task children were asked which of two words rhymed with the first part of a word they had learned (e.g., "meteor rhymes with meet or meed?"). They found the judgments matched the spelling for the children who had seen the words, whereas the controls judged more flaps as matching /d/ words than /t/ words. Ambiguous results for interpretation of the flap by adults were also noted by Derwing and Nearey (1981). Judgments for nonsense words definitely leaned to a /d/ interpretation of the flap and the word buddy was always classed with the /d/ sound, but words with the t spelling for the flap caused problems. They found that 56% of the group trained on the phoneme /t/ agreed that butter contained the initial sound in tough while 58% of the /d/-trained group agreed that butter contained the initial sound in duck!

While the classification of neutralized segments is one potential area of disagreement betweeen traditional linguistic analyses and speaker judgments, the segmental status of some sounds is another. Using the concept formation technique, Jaeger (1980a) tested whether affricates were considered one sound or two by native speakers. Although she found that the affricates seemed to be perceived as one sound, she could not dismiss the possibility that this might also be orthographic influence. For example, the voiced affricate is typically represented by only one letter (j), a fact which perhaps influences its perception as a single sound. Another possible influence is the analogy of ch to clear cases of single phonemes represented by two letters (such as sh and th). Jaeger suggested that word-final affricates spelled in such a way that the two-segment interpretation is more likely (e.g., dge and tch) should be tested before conclusions as to their status are made.

Ehri and Wilce (1979) investigated just such cases in a study of the influence of orthography on children's notions of phonemic structure. They trained fourth graders to segment words into phonemes by tapping out sounds with a wooden dowel. Although the children noted that /tʃ/ (spelled ch) has two letters but is only one sound, and ignored 'silent' letters (like the e of home), the majority of them detected an 'extra' segment in each of the words catch, pitch, and badge. In a visual word learning task, Ehri and Wilce found that nonsense words with extra letters were easier to learn than those without (e.g., zitch *versus* zich) and children who had learned the spellings with an extra letter detected an extra phoneme in the segmentation task consistently more often than children who had been exposed to the other spellings. For the two nonsense words with affricates (tadge and zitch), 83% of the children exposed to the spellings with an extra letter detected an extra segment.[2]

The influence of non-phonetic knowledge can be very strong, leading speakers to perceive differences that are not realized phonetically (e.g.,

[2] After reading Ehri and Wilce's (1979) article and discussing their results, several members of our department became convinced that they were starting to hear the alveolar stops, too!

Sapir's (1933) example of a Sarcee speaker's perception of a stem-final [t] in some cases that had no phonetic realization; cf. the analogous case of English [n] in hymnal, but not hymn) and even to producing distinctions that are not normally part of the language. Fourakis and Iverson (1984) criticized studies by Mitleb (1981), Port, Mitleb, and O'Dell (1981), and O'Dell and Port (1983) which reported acoustic differences and native speaker discriminability between 'underlying' voiced and voiceless final obstruents in German (which are all phonetically voiceless), a finding that seemed to contradict the general assumption of word-final obstruent neutralization in German. Fourakis and Iverson argued that the acoustic differences found were due to hypercorrect spelling pronunciations and, using a technique designed to disguise the focus of study (students were asked to conjugate strong verbs rather than simply read lists of words), they found no acoustic differences related to the underlying voice of the word-final obstruent.[3]

Although the generality of word-final obstruent neutralization in German has arguably been re-established by excluding the influence of orthography, there is good evidence that another 'phonological' rule is in fact orthographically based. The vowel shift rule (VSR) described in Chomsky and Halle (1968) is pivotal to their conception of abstract phonology. Pairs such as divine/divinity, extreme/extremity, and

[3]Their partial replication of O'Dell and Port (1983) did find trends for longer vocalic nuclei for underlying voiced consonants which may have been the result of morphophonemic knowledge, but the authors argued against this on the grounds that one of their words, the adjective weg, never inflects. This is not entirely convincing, however, since weg [vɛk] is related to the noun Weg, which has the plural form of [veːgɛ]. See also Fourakis (1984) and Dinnsen (1985) for discussions of neutralization with respect to experimental evidence.

pronounce/pronunciation are assumed to be derived from morpheme invariant underlying forms through a number of phonological rules, including the [illegible] Moskowitz (1973) found that children did have some knowledge of VSR, but that its acquisition was gradual. Any kind of vowel alternation (VSR, anti-VSR, tense-lax) was tolerated by 7 and some 9 and 12 year-olds, while other 9 and 12 year-olds accepted only VSR alternations. The vowel alternations of Chomsky and Halle's VSR include those that are preserved in English spelling and relate directly to the 'long' and 'short' vowel quality distinctions taught in school. Moskowitz concluded that what Chomsky and Halle claimed was a rule of abstract phonology was actually learned from the spelling system of English.

This interpretation was supported by Jaeger (1980a; 1986), who claimed that speakers feel that the five alternations related to the 'long' and 'short' spelling pattern of vowels (A [ey]-[æ], E [i]-[ɛ], I [ay]-[ɪ], O [ow]-[ɔ], U [u]-[ʌ]) go together. In a concept formation experiment on university graduates she tested the pairs predicted by the Chomsky-Halle version of the VSR (the same except [u]-[ʌ] is replaced by [aw]-[ʌ]) against the spelling rule and found that subjects rejected the VSR [aw]-[ʌ] pair but accepted the [u]-[ʌ] spelling rule pair as part of the category. Jaeger pointed out that the speakers' vowel alternation rule is psychologically real but is probably derived from orthography rather than an assumed deep phonology. Templeton and Scarborough-Franks (1985) found that grade ten students were significantly better than grade six students at pronouncing VSR pseudowords. Similarly, Taft and Hambly (1985) compared people's judgments of morphologically supported and non-supported monosyllabic primes which were spelled the same as (consistent) or differently from (inconsistent) the first syllable of target words and found that orthographic, rather than morphological, similarity was the important factor, again suggesting an orthographic interpretation. Finally, using a concept formation technique and a variety of other techniques with high school students, Wang (1985; Wang & Derwing, 1986) convincingly demonstrated that the critical back vowel cases which distinguish the effects of the VSR (under three different linguistic formulations) from the long-short vowel spelling rule clearly support the psychological reality of the spelling rule.

The VSR can perhaps be better described as a historical change that has been preserved in the English orthography. Although spoken rather than written language has traditionally been linguists' principle domain of inquiry, the importance of writing must be acknowledged since writing clearly influences language users. As Kerek (1976) pointed out, writing has always enjoyed popular prestige and authority. This is especially true today, with the value most societies place on literacy and the disappearance of oral-only cultures. In some such latter cultures, three or four generations of language can result in such widespread changes that great-grandparents and grandchildren cannot communicate (Gelb, 1952). In contrast, English written 400 or 500 years ago is still comprehensible, a fact which Gelb credited to the restraining influence of the written form. Older forms no longer generally used in speech are often preserved in writings and may emerge again later in oral language, a process which Householder (1971) described as a type of dialect borrowing.

Spelling pronunciations can also be viewed as examples of such borrowing and Householder (1971) claimed there were 'countless' occurrences of these. Learned borrowings into English from Latin and Greek are essentially spelling pronunciations (Levitt, 1978), as are probably the pronunciations of most words used primarily in the written form. Spelling can not only dictate the pronunciation of infrequently used words, but also of words borrowed into English from other languages. Kerek (1976) cited the examples of chef, chalet, chic, chassis, chaise which some English speakers produce in accord with the English rather than the French spelling conventions (i.e., with initial /tʃ/ rather than /ʃ/). He suggested that the standard /θ/ pronunciation of th is now also exerting pressure on words such as thyme [tʰaym], Thames [tʰɛmz], and Waltham [waltʰəm] and that the h of herb, homage, and heir and the t of often, pestle, and chasten are gradually reappearing. As every schoolchild quickly learns, the dictionary is generally taken as the ultimate reference source for pronunciation, but Levitt actually argued that dictionaries encourage spelling pronunciations because they use pronunciation codes that are often difficult for the phonetically untrained user to decipher. Another interesting example of the influence of spelling on pronunciation is the so-called standard form of German, which Levitt called a 'paper language', because the entire vocabulary is pronounced based on the norm derived from the spelling. This serves the useful purpose of enabling speakers of sometimes very different dialects to have a common base with which to clarify intended messages.

Spelling pronunciations are of some interest to linguistics because of the effect they can have at slowing or even reversing language change. Of more interest to the study of psycholinguistics are the perceptions language users have about the priority of the written form. Householder

(1971) claimed that naïve speakers intuitively feel that speech is derived from writing, not the converse, and cited the familiar phenomenon of requesting the spelling of a new word or name in order to remember or to be able to reproduce it. In a similar vein, careful pronunciations, produced perhaps to clarify potentially ambiguous words (e.g., ladder/latter), are only perceived to be correct if they agree with the spelling. Skousen (1982) discussed careful pronunciations in the resolution of this and other cases of phonemic overlap and found that, generally, these were guided by the spelling. There have been some cases, however, where the spelling reflects a phonemic ambiguity: protest (with a medial [t^h]) is related to Protestant (with a medial flap) which is sometimes shortened to Prod (a clear [d], example from Skousen). Careful pronunciations can lead spellers astray, as Skousen pointed out. Although the medial /ʌ/ of general, mystery, camera, and interest can be deleted even in careful pronunciation, speakers know that it is 'there', perhaps partly or solely due to the spelling; thus, on analogy, fabric and athlete have been mistakenly written as athelete and faberic, which 'sound' right with careful pronunciation but do not 'look' right.

Skousen (1982) also suggested that it may be orthography that determines which consonant clusters are permissible in English. For example, tm and bn are not permissible clusters, although it might be argued that tmoro, tmato, and bnana are perfectly reasonable representations of English words (all are taken from Skousen's examples of children's spellings). Nasalization of vowels is not considered distinctive in English, but Read (1971) found that children's invented spellings commonly omitted nasals whenever another consonant followed, suggesting that the vowel carried the nasality for these children, and Ehri and Wilce (1987a) reported that pre-literate children who had been trained to do simple letter-sound decoding of words also omitted pre-consonantal nasals in their spellings.[4] It may be that judgments about the presence of a nasal segment *versus* a nasal vowel are coloured by knowledge of the spelling of words. Skousen has also pointed out that many linguistic arguments are based on adult perceptions of phonemic representations, and these may be different from children's because of maturity and/or a broader knowledge base, including knowledge of conventional spellings.[5]

It may be that both phonological and orthographic knowledge are automatically activated in certain language-related behaviours. Words' spellings may guide phonological judgments in ambiguous cases and there might be orthographic influence in tasks such as rhyme production and segment counting, especially when stimuli are presented visually. Seidenberg and Tanenhaus (1979) investigated the effects of orthography on rhyme detection (rather than production) with spoken words and found consistently shorter times to detect orthographically similar rhymes than differently spelled rhymes (e.g., pie-tie was faster than rye-tie). In a later study using a different technique (Tanenhaus, Flanigan, & Seidenberg, 1980), they found that this orthographic influence was present even when it had a negative effect on performance. In this study, they used the Stroop (1935) technique whereby various combinations of priming and target words are presented embedded in a colour naming task in order to judge the relative amount of interference on the task. Colour naming latencies were longer when the target word was preceded by a phonologically and/or

[4]See Derwing (1973, Chapter 6) for further theoretical discussion of this point.

[5]See also Linell (1982) for an even broader account that encompasses syntactic effects.

orthographically similar prime than a control prime word. They concluded that visual word recognition entails activation of multiple codes. This interaction of phonological and orthographic codes also has neurophysiological support (see Kramer & Donchin [1987] for a discussion of brain potentials).

Further support for the hypothesis of interacting orthographic and phonological codes comes from a rather unusual phenomenon - backward talking. Cowan and Leavitt (1982) described two boys who could talk backwards. One reversed words as though they were spelled backwards and the other reversed according to phonemes. Interestingly, the former backward talker could not reverse a word if he did not know the spelling, and the other boy occasionally used some forms that were related to the words' spellings rather than phonemes (e.g., <u>trash</u> /træʃ/ was [hʌ'sært], not [ʃært]). Apparently the ability to speak backwards is not too unusual, since Cowan and Leavitt noted that 27 fluent adults have also been found and studied. Approximately half of these reversed according to spelling and half according to phonemes.

Obviously, both phonological and orthographic codes are accessible for various types of language behaviour. However, it is difficult to separate the role of different types of knowledge in any specific language behaviour. One possibility is to examine the behaviour of people who have no knowledge of conventional orthography.

2.2 Literacy and Evidence of Phonemic Knowledge

Sapir's writings are probably most often cited in discussions of the psychological reality of the phoneme. In his 1933 article, he claimed that people hear phonemes (as opposed to phones) and he found some evidence for the psychological reality of phonemes. One example Sapir discussed was that his assistant transcribed a word-final /t/ when there was no phonetic evidence of that /t/. It may seem insignificant that this example had to do with a written rather than an oral report, but informal accounts suggest that it is very difficult for native speakers to label what is different about two near-homophonous words without recourse to the spelling. Explanations are usually given in general, qualitative terms without specific reference to the differing phonemes, or by the use of simple repetitions of the word to provide the correct model. Explanations from non-linguists are seldom made with respect to individual sounds, although some refer to letters (e.g., "Think begins with a 'th'-sound"). On first reading, Sapir's article might convince the reader that phonemes are very natural units because native speakers of languages without orthographic systems were able to act as linguistic assistants, learn the phonetic alphabet, and transcribe their own languages. However, the assistant Sapir described was someone who could already read and write English reasonably well. The imposition of phonetic transcription, which assumes a segmental representation, may have inadvertently presented a biased picture of the phonological representations of certain languages and it eliminated the possibility of determining whether some other type of phonological representation might be more appropriate for languages without alphabetic orthographies.[6]

For example, the traditional Cherokee syllabary writing system of 85

[6] Although, as discussed in Section 1 with respect to the segment count studies in English, it does seem to be quite an appropriate model for literate English adult speakers.

letters was developed by Sequoyah, who reportedly had no alphabetic or phonetic training (Walker, 1969). The system was used successfully in the 1800's and it has been estimated that in the 1830's 90% of the Cherokee were literate in their language. The Cree syllabary was designed by an English speaker (James Evans) and seemed eminently suited to that language. A similar example of this preference for syllable-sized units was discussed by Jaeger (1980a), who found that Japanese speakers apparently did not believe it was possible to break syllables into smaller units (see also Sakamoto 1980) and Read, Yun-Fei, Hong-Yin, and Bao-Qing (1986) have recently shown that older Chinese speakers untrained in the pinyin romanization scheme have great difficulty in manipulating phonemic segments. Although syllabaries may not be suitable for all languages (e.g., English has hundreds of distinct syllables), segmentally-based orthographies may not be, either.

It can be argued that the principle of speech as a sequence of phonemes is natural in some ways. It is certainly an economical way to represent the sounds of words. In English only about 40 phonemes are required to represent the entire vocabulary of the language, whereas a syllabic or pictoral representation of words would require hundreds, even thousands more symbols. The person who first stumbled on this ecomony and devised the alphabet (although the origin and monogenesis is somewhat in doubt; see Gelb, 1952) must have managed, without instruction, to conceive of speech as a sequence of phonemes. The alphabetic principle was quickly adopted by a variety of languages and cultures, suggesting that there was something very appealing about it. There are also reports of cases in which the alphabetic principle has been spontaneously 're-invented' by children.

The best-known instance of a spontaneous segmental analysis is 'invented spelling' by pre-literate children (Read, 1971; C. Chomsky, 1975), which is, in fact, quite unusual (see Paul, 1976). The motivation for spontaneous spelling seems to be for its own sake, rather than for communication. Children in Paul's study seldom invented the same spelling twice over several months and Read reported that some children could not even read back what they themselves had written. Although there has been some attention directed towards distinguishing spontaneous spellers from children who do not invent spellings, nothing conclusive has emerged except that parents of spontaneous spellers tolerate the activity. There is some evidence that spontaneous spellers are more conscious of phonology, as Read (1973) found that they performed much better at a phonological ABX discrimination task (i.e., "Is X the same as A or B?") than other children.

Rather than proving the innate nature of segmental analysis, it may be that invented spelling is a sign of a heightened skill in 'cracking a code'. All of these children knew the alphabet and had obviously made the connection between sounds and letter names. C. Chomsky (1975) claimed that by age 5 or earlier children can recognize that words can rhyme and can start with the same sound. The spontaneous spellers had simply gone the one (important) step further of breaking sounds down into units that could be isomorphically matched to the tools they had available (letter names). The stages of spontaneous spelling reported in Paul (1976) support this analysis: Children first represent only the initial sound of a word, later both the initial and final sounds, and finally more complete segmentations.

Reading researchers have often assumed that phonemic analysis skills are required in order to understand the alphabetic principle and hence learn to decode the written word. A number of studies have focussed on the development of phonemic analysis skill (e.g., Liberman, Shankweiler, Fischer, & Carter, 1974; Fox & Routh, 1975; Skjelfjord, 1976; Treiman & Baron, 1981). The consensus seems to be that preschoolers can segment words into syllables but not into phonemes, that kindergarten children have some success at phoneme segmentation, and that children at Grade One and above can readily segment words into phonemes.

Bruce (1964) was one of the first to investigate phonemic analysis ability. He asked children to tell him what was left after a particular letter/sound was removed from a word (e.g., taking away the /s/ from <u>spin</u> yields <u>pin</u>). Children with a mental age (MA) below 8 years managed to do fewer than half of the 30 items correctly. Bruce's conclusion, widely accepted and extended to phonemic analysis of any kind, was that a certain level of cognitive development was necessary before words could be analysed phonemically. The task used in this study was a very difficult one, however. Not only did children have to realize that words were made up of individual sound segments and be able to segment the word accordingly, they had to match the target sound, remove it, and reconstruct the word. This may not seem so difficult for the example <u>s</u> from <u>spin</u>, but <u>n</u> from <u>hand</u> is more complicated, even for adults. Furthermore, it is not clear that mental age alone determines this skill. Bruce gave some hint of this in that he reported that children from different teaching environments differed significantly in ability.

With a considerably easier, more child-appropriate task, Zhurova

(1973) found that phonemic analysis skill seemed to develop gradually. The strategies used by children in naming the first sound of a word differed greatly at different ages. Young (3 to 4 years old) children could isolate an initial sound but not separate it (e.g., d-d-doggie but not d oggie). Older children (4 to 5 years old) isolated the sound and then produced it (e.g., d-d-doggie . . . d) and by 5 to 6 years old, half of the children could isolate the first sound of their own name without instruction. Interestingly, children who knew the letters of the alphabet produced the first sound as a letter name, whereas other children pronounced syllables (e.g., in English, /d/ would be produced as [di] as opposed to [dʌ]). Zhurova concluded that the ability to isolate sounds in words is not a single stage act manifested spontaneously and pointed out that until reading and spelling instruction begins, no such artificial discrimination of individual sounds in words is necessary.

The American tradition[7] of phonemic analysis research seems more committed to an explanation of segmental analysis as a naturally developing ability than to one related to specific instruction, with Liberman *et al.*'s (1974) study widely cited in discussions of phonemic analysis ability. Even with some training, the preschoolers and kindergarteners in their study could not do segmental analysis in a tapping task, although they could do some syllabic analysis. By contrast, the children in first grade could do phonemic analysis quite well. Fox and Routh (1975) looked at correlates of phonemic analysis ability and found that it was significantly related to age, receptive vocabulary, vocabulary I.Q., reading

[7]Zhurova is Russian and the examples she used have been given rough English equivalents here.

recognition, reading comprehension, sentence segmentation, and syllable segmentation. Of the factors considered, only parental occupation did not seem to matter! Unfortunately, the amount of previous instruction in reading was not examined as a correlate. It is clear that instruction does make a difference in performance of segmental analysis. McNeil and Stone (1965) successfully trained kindergarteners who had had no previous instruction in reading or phonetic analysis to segment nonsense and familiar words; Skjelfjord's (1976) pretesting showed that children were not able to analyse words spontaneously into phonemes but after one week's teaching the majority of responses were phonemes, not syllables. (Interestingly, one girl in Skjelfjord's sample could read and apparently performed the task with no difficulty and made no errors even without the week's training.)

The various experiments supposedly dealing with the relation between phonemic analysis skills and reading can be criticized in a number of ways. As Golinkoff (1978) noted, some tasks were quite artificial. Tapping in time to phonemes that are not acoustically separable (e.g., stop plus vowel sequences) is a questionable activity. Removing a phoneme from a word to produce a new word is usually only done with word-initial sounds for rhyme, and medial and final sounds are rarely manipulated in English language games or art. It is also doubtful that children in first or second grade are cognitively ready to deal with such operations. In a study by Rozin, Bressman, and Taft (1974), children were asked to indicate which written word related to a spoken word. Children who had not yet attained moderate reading fluency did not even understand that short spoken words generally relate to short written words (e.g., /mo/ is related to mow, not motorcycle). It may be that some problems children have in determining what writing is all about is at the more general level of understanding that different dimensions (space in writing and time in speech) are systematically related.

The relation between phonemic segmentation and reading ability can be understood in two ways. The first, preferred by many reading specialists, seems to be that phonemic segmentation is a natural language ability based on the minimal units of speech (phonemes). In order to understand the alphabetic principle, conscious awareness of phonemes must be realized, and the better this is, the better reading will be. For example, a test designed by Rosner and Simon (1971), similar to Bruce's (1964) design, claimed to predict reading ability based on phonemic analysis skills fairly accurately (up to a correlation of .84). Similarly, Rozin and Gleitman (1977) claimed that accessibility to the phonemic level as represented by the orthography must be taught. It is accessed later than syllables, for example, because it is a lower level.

The second view of the relationship reverses the causality between reading and phonemic analysis and would claim that reading instruction teaches children phonemic segmentation (i.e., the idea that speech can be regarded as a sequence of phonemes). The awareness of speech as a sequence of syllables apparently does not need to be taught, since children as young as 3 years old can segment speech into syllables without training (Fox & Routh, 1975). Children who cannot read and spell can speak, repeat and learn new words, and process speech, so their phonological analysis must be adequate at some level. Perhaps the specific level on which reading researchers have focussed (i.e., the phoneme) is not the appropriate one, or perhaps what has been called phonemic analysis is not really accessing the cognitively fundamental representation at all, but is testing conscious, reported analysis.

Clearly, by school age, most children are cognitively ready to discover that speech can be described as a sequence of segments. If the ability to segment does not develop suddenly or spontaneously, then differences in instruction should result in differences in ability. Alegria, Pignot, and Morais (1982) tested just this hypothesis with Belgian schoolchildren learning to read either by the phonic approach or the whole-word approach. The children in the phonic group performed better at phonemic reversal than the whole word group. The performance of the phonic group was also significantly correlated with reading level but the whole word group's was not, supporting the notion that specific instruction, rather than reading behaviour itself develops segmental analysis. As with the counting studies already cited, Alegria *et al.* found that manipulating syllables was easier than phonemes in a reversal task, with both groups performing equally well with syllables. Fox and Routh (1984) reported that segmenting training helped segmentation ability (phonemic analysis) but not necessarily decoding written words into speech.

Such a learned skill (if that is what phonemic segmentation is) should be related to differences in other learned abilities, specifically reading and/or spelling. Probably the most convincing evidence came from studies by Morais and his colleagues (Morais, Cary, Alegria, & Bertelson, 1979; Morais, Bertelson, Cary, & Alegria, 1986) with Portugese illiterates. In investigations of phonological knowledge it has often been suggested that experiments be performed with illiterates to truly control for the influence of orthography. In the Morais *et al.* (1979) study, the illiterates seemed to understand the task and performed similarly to a literate control group on a syllabic reversal task, but the phonemic analysis task proved to be almost impossible for them (average score of 19% correct). Although the literates performed much better, they showed differences that were related to their reading ability. All had learned to read and write at age 15 or older, but those who had attained a course certificate for literacy performed the phonemic task with a mean of 79% correct, whereas those without a certificate performed with only 55% correct. The later (1986) study also reported a large difference in performance between literates and non-literates on a variety of phonemic-related tasks. Read *et al.* (1986) compared alphabetic and non-alphabetic literates in China and found that the non-alphabetic literates could not segment speech at the phonetic level (alphabetic literates mean=93% correct for words, 83% correct for non-words; non-alphabetic literates mean=37% and 23% correct, respectively) and concluded that segmentation skill must be related to alphabetic literacy, not cognitive maturation, literacy, or knowledge of rhyme.

This difference in ability to perceive segments has also been noted for native English speaking adults and children. Baron, Treiman, Wilf, and Kellman (1980) distinguished children and adults by their use of phoneme-grapheme correspondence rules and found a relation to segmental analysis ability. One group (their 'Phœnicians') was good at these rules and could segment words into phonemes, the other group (their 'Chinese') performed less well with spelling-sound rules and were not as good at segmental analysis. Barton, Miller and Macken (1980) found that young children could isolate the word-initial consonant cluster from the rest of the word, but most were unable to segment the cluster into constituent phonemes. There was a high correlation between segmentation ability and reading ability for these pre-schoolers, leading the authors to suggest that orthography teaches the children how to segment these cluster units.

Fox and Routh (1980) found large differences in segmental analysis for good and poor readers. Reading disabled boys could not segment at any level (words, syllables, phonemes) and severely reading disabled girls could not segment syllables into phonemes. In fact, the test Fox and Routh devised proved to be a perfect predictor of severe reading disability. All children except those with severe reading disability could segment sentences into words and words into syllables perfectly and most could also do phonemic segmentation perfectly (except the mildly reading disabled boys with only 93.7% on phonemic segmentation).

Obviously segmentation ability is strongly related to reading ability, but this may in fact be an incidental product of the relation between reading ability and spelling. Ehri and Wilce (1987b) found that young children trained to spell phonetically were able to segment words much better than untrained children were. It may be that people can read without being able to segment (as suggested by the whole word group of Alegria *et al.*, 1982), although segmental analysis is usually a by-product of learning to read (as argued earlier) and is necessary for good spelling.

Perin (1983) made the strong claim that segmentation ability is a necessary precursor to spelling. She had 668 adolescents do two types of phonemic segmentation tasks: to produce spoonerisms of the names of pop singers and groups, and to determine the number of segments in real and nonsense words. The subjects who read and spelled well (her group A) performed both tasks without difficulty. Both good readers/poor spellers and poor readers/poor spellers (groups B and C) had difficulty with the two tasks and performed poorly. From verbal reports and behaviour, Perin judged that all the students approached the tasks in the same way, but that the poor speller groups were unsuccessful. The fact that the poor spellers could do spoonerisms at all, and that the majority of errors involved phoneme exchanges, suggests that it is only conscious isolation of phonemes that is difficult, not necessarily phonemic awareness. Perin concluded that spelling was more closely related than reading to segmental analysis (since reading ability did not seem to matter here) and further suggested that learning about phonemes may only be possible in the context of spelling and may be a consequence of learning the alphabet.

The English alphabet and specifically grapheme-phoneme correspondences (or lack thereof!) have been much maligned in the literature and are the focus of many complaints by native and non-native English speakers alike. In an attempt to determine the influence that the alphabet plays in reading difficulties, Rozin, Poritsky, and Sotsky (1971) carried out a study on second grade children with reading problems, teaching them Chinese characters to represent English words.[8] In a total of a little over four hours, the children learned to read simple English sentences presented with Chinese characters, although relatively little progress had been made in reading the English alphabet over the same period of time. This result has been replicated (Harrigan, 1976) with offered explanations ranging from increased motivation because of novel materials to the assumed processing of characters on the opposite side of

[8]Gleitman and Rozin (1973) used a syllabary to similar effect. This is discussed in Chapter 3, section 3.1.

the brain from speech.

The reports from countries that use non-alphabetic writing systems also suggest an important relation between orthography and reading ability. Taylor (1981) reported that South Korea has an illiteracy rate of only 5%, which is limited to the mentally retarded and older people who did not take part in the educational system. Japan is noted for its low rate of reading disability (1%, according to Sakamoto & Mikita, 1973) and has a number of writing systems, including Kanji (characters) and Hiragana (syllables). Hiragana consists of 46 Haku which are roughly equivalent to the English notion of syllables (the exception is the alveolar nasal Haku). Children are considered ready to read in Japan when they can divide words into constituent Haku (Sakamoto, 1980). There is some evidence that this is the minimal unit of sound to Japanese speakers. Jaeger (1980a) reported that some of her Japanese subjects claimed that the syllable was the smallest unit of sound possible and could not be broken into smaller sound units. Almost half of the Japanese subjects were unable to perform a task based on a phoneme-sized category, whereas all English subjects could. The question was whether this difficulty with phonemes for Japanese speakers was a function of the sound system or whether it merely reflected the influence of the Japanese writing system. A study by Mann (1986) suggested this influence might be uni-directional. She reported that first grade Japanese students performed much more poorly than American students on tasks requiring a segmental analysis, although their ability with syllables was equivalent.

Children's phonetic judgments differ from adults' in ways that may also be a reflection of different forms of phonological representations. Read

(1971) noticed that children sometimes used CHR to represent /tr/ clusters (e.g., AS CHRAY for ashtray). He tested whether the perception of /tr/ by kindergarteners, first graders, second graders, and adults differed (Read, 1973). The first and second graders classified the majority of /tr/ tokens as more similar to /tʃ/, whereas the adults clearly preferred /t/. The suggestion of orthographic influence on this judgment emerges not only because of the difference between children and adults, but also from Read's observation that first graders with high reading achievement preferred the /tr/-/t/ classification.[9] Treiman (1985b) supported Read's findings with three experimental tasks with kindergarten and first grade children and also found significant negative correlations between reading ability and nonstandard judgments, suggesting that knowledge of orthography forces standard judgments.

Although phonological features like the number of syllables in words and position of stress are probably equally important to children and adults, Vihman (1981; 1982) suggested that other features may be weighted differently, with children's underlying forms closely related to their output forms, rather than to the adult norms. Waterson (1976; 1981) claimed that features are related to syllable position for children, rather than to segments and that these 'schemata' become increasingly differentiated and complex and closer to adults' forms over time.

A direct comparison of children's and adults' judgments of phonetic similarity was carried out by Treiman and Baron (1981) and Treiman and Breaux (1982). They asked adults and preschool pre-reading children to

[9]The results of the kindergarten children did not support a clear developmental trend towards increasing /t/ judgments for /tr/, however, because they preferred the /t/ over the /tʃ/ judgment.

classify syllables by putting the two that matched (out of three) together. They found that children based their classifications on overall similarity of the syllable as calculated by the metric from Singh, Woods, and Becker (1972; e.g., [bɪs]-[diz], not [bɪs]-[bun]) whereas adults generally based theirs on shared phonemes (e.g., [bɪs]-[bun], not [bɪs]-[diz]); they used overall similarity only when there were no shared phonemes to exploit. Children responded to training in overall similarity by increases in that classification, but did not improve phoneme classifications even after training. In fact, phoneme training seemed to confuse the children, as overall similarity judgments decreased due to an increase in the number of anomalous judgments. The authors concluded that children and adults may differ in the initial encoding of spoken syllables, but they did not discount the possibility that these may still be represented as strings of phonemes. There has been some suggestion in the psychological literature (Cooper, 1982) that people may differ in the type of strategy used, with a contrast drawn between holistic and analytic processing.

Descriptions of sound in terms of individual phonemes cannot explain why children attend preferentially to the overall similarity of syllables. Perhaps children's phonological representations are more appropriately described in terms of syllables, or a level between syllables and phonemes, than phonemes. Menn (1978) suggested this in her discussion of characteristics of phonological development. One characteristic is cluster reduction, whereby children reduce clusters but apparently retain as many features from the adult cluster as possible. It may be that children view clusters as single units and therefore omit some characteristics rather than delete a whole segment. Barton *et al.* (1980) found that pre-reading children treated initial consonant clusters as units and pointed out that young children typically failed to produce all the features of many adult consonant clusters, yet phonologists still treat these as separate phonemes. If the sibilance of an /st/ cluster is omitted, for example, a segmental analysis would conclude that the segment /s/ was omitted, whereas an analysis that viewed clusters as units would conclude that only one feature of the cluster was omitted.

Children may not be the only ones with phonological perceptions that differ from traditional phonological analyses. Marcel (1980) investigated spelling problems of children, of adults in literacy classes, and of neurological patients. These subjects tended to omit liquids and nasals, especially in clusters with unvoiced consonants. In clusters with voiced consonants there was some trend to devoice consonants along with omission of the liquid. Marcel suggested that these spellers were using a set of features to distinguish stop plus liquid clusters that did not correspond to individual letters (or phonemes). Although these spellers could repeat the words accurately aloud, they did not seem to be aware that the phonemes they consistently omitted in spelling were present in the spoken form. Segment counts showed that the number of phonemes in initial stop-liquid clusters were consistently underestimated, as compared to single phoneme items. For these spellers it seems more appropriate to describe clusters as units than as groups of segments.

2.3 Units of Speech Perception

The move away from the phoneme to a different basic level of representation has been made to describe children's phonological representations. It has also been suggested that phonemes may not be accessible to non-literates or users of non-alphabetic orthographies. A series of psychological experiments in the 1970's attempted to define the unit of speech perception. Although these experiments demonstrate the need for careful design more than the reality of any linguistic unit, they do point out some interesting facets of speech sound perception.

Savin and Bever (1970) sought to compare search times for auditorially presented targets of phonemes (/s, b, æ/) or syllables (/sæb, bæb/). They found that times for syllables were fastest, followed by initial /s/, then initial /b/, and finally medial /æ/. From this they concluded that syllables were perceived before their constituent phonemes. They did not deny the psychological reality of phonemes, claiming instead that they were abstract entities, whereas syllables were 'real'. The search times were all considerably less than the duration of a syllable (300 msec. *versus* 600 msec.), so listeners were obviously not waiting for the whole syllable before making their decisions. Foss and Swinney (1973) were concerned about this, and also questioned the economy of representing English syllables instead of phonemes in the lexicon. They replicated Savin and Bever's experiment and included two-syllable words as stimuli, as well as phoneme and syllable stimuli with search lists made up of two-syllable words (Savin and Bever had used search lists of single syllables). Foss and Swinney found that the two-syllable target was fastest and the phoneme slowest. Since it seemed to them unlikely that disyllables were the basic perceptual entities, they suggested that the results reflected listeners' uncertainty about the identity, rather than the perception, of the target. Under this interpretation, phonemes are perceived first, then syllables and words are identified, and finally, phonemes are identified.

This was a fairly sophisticated explanation for an effect that was soon shown to be quite straightforward. McNeill and Lindig (1973) matched the target and search list in terms of size of unit (phoneme or syllable) and found that listeners responded faster when the two (target and list) were on the same linguistic level. Savin and Bever found syllables fastest because they used syllable lists and Foss and Swinney found two-syllable words fastest with disyllabic search lists. According to McNeill and Lindig, what is perceptually real is whatever one pays attention to, therefore monitoring techniques cannot reveal the ultimate perceptual units of speech.[10] This conclusion was indirectly supported by Healy and Cutting (1976), who found no difference between phonemes and syllables when hetereogenous lists were used. They did not reject the usefulness of monitoring tasks, instead claiming that since it was impossible to say whether phonemes or syllables were faster, both must be linguistic entities and equally basic to speech perception.

There is, however, more support for processing of sequences of phonemes as units. Although Mehler, Sequi, and Frauenfelder (1981) agreed that monitoring tasks results were difficult to interpret clearly, they favoured Savin and Bever's (1970) suggestion that syllables are perceptually real. They argued from monitoring type studies and infant discrimination results that sequences of phonemes corresponding to syllables are processed as units. They found that identical CVC sequences were treated differently in a reaction time experiment with adults

[10]Note that Foss, Harwood, and Blank (1980) did not accept McNeill and Lindig's results or conclusions, although they did agree that one cannot accept the pattern of reaction times as necessarily reflecting perceptual processing difficulty.

dependent on whether they formed a syllable (e.g., CVC#...) or crossed syllable boundaries (e.g., CV#C...). Infants, too, seemed sensitive to syllabicity. Mehler *et al.* reported that infants showed different habituation patterns to CCC sequences than to CVC sequences or to the same consonant sequences in a vocalic context (e.g., VCCCV was more like CVC than CCC). There is evidence that Wood and Day's (1975) interpretation of the consonant and vowel being processed together (independent of syllabicity) is correct. Jaeger (1980a) found, to her surprise, that some English speakers who were asked to divide simple words into constituent phonemes divided cat into two, either /kæ-t/ or /k-æt/. It may be that phonemes do group together into larger units that are not necessarily syllabic.

Sub-syllabic units have been investigated in the context of word recognition studies. Santa, Santa, and Smith (1977) used a reaction-time paradigm whereby a target word and various probes (such as single letters, adjacent double letters, adjacent triples, and the whole word) were visually presented simultaneously. All words were five letters long, of the form CCVCC or CCVCe. They found that single letters, the initial consonant cluster, the initial and final triples, and the whole word were all recognized equally quickly. These results were surprising to them because, under an assumption of sequential letter-by-letter processing, determining that two letters of a probe match a target should take twice as long as a single letter match. The fact that the word probe was as fast as single letters suggests that simple sequential processing was not an adequate explanation, and simple bi- and tri-gram frequencies could also not explain the results. The advantage for the initial consonant cluster and triples probes is interesting for the present discussion and not out of keeping with previous results. Initial consonant clusters have been identified as possible units in language games, as well as in children's phonologies, speech errors, and speech synthesis and recognition. The advantage of the initial triple might be related to the integrity of the onset cluster, while the final triple relates to the rhyme element - a very natural, moveable unit in English. The fact that there was no advantage for the medial triples (despite their having the 'prototypical' CVC syllable shape) argues against a simple letter frequency explanation and for an explanation based on syllable structure.

There is some evidence that syllab[illegible]s play a role in confusions in memory. Conrad (1964; 1972) found that letter names sharing a vowel sound were confused in a recall task (e.g., /-i/ for B, C, P, T, V and /ɛ-/ for F, M, N, S, X, but no confusions between the two sets). Recall of lists of rhyming words (e.g., cat, hat, mat, fat) was much more subject to confusions than non-rhyming words. Treiman and Danis (in press) also found that groups of segments related to the syllable onset and rhyme functioned as units in short-term memory errors for spoken syllables.

Another interesting source of evidence about the segmental organization of speech is from speech errors and related phenomena. These include 'tip of the tongue' phenomena, 'slips of the ear', and 'slips of the tongue' (spoonerisms). The units that are involved in these speech phenomena can tell us something about the processes involved in speech production and perception. For example, when people cannot remember the whole of a target word (tip of the tongue phenomenon), they tend to recall the number of syllables, syllabic characteristics, and/or consonant clusters (initial clusters primarily, finals secondarily) of the word, while single segments are least accessible (Browman, 1978).

Groups of segments and single segments are both involved in misperceptions. These occur in everyday language situations, when someone tries to make sense of what they have heard (such as the example from Garnes and Bond (1980) - "slip of the year" for "slip of the ear") or when they do not hear clearly (e.g., "Did you say 'grow up' or 'throw up'?"). The error likelihood for misperceptions is highest for the initial segments of syllables (Browman, 1978; 1980). Stemberger's (1983) analysis of speech errors involving /r/ and /l/ argues quite strongly for a segmental, specifically phonemic, basis for speech errors. The second position of word-initial consonant clusters seemed particularly susceptible to errors (relative to the word-initial segment), which Stemberger and Treiman (1986) argued was a result of the importance of syllable position in speech production. Relatively few misperceptions seem to be based on confusions across syllables, indicating the integrity of the syllable as a unit. If segments are independent and not bound more strongly to adjacent segments than to any others, one would not expect consonant clusters or syllables to emerge as units, yet MacKay (1972) found strong support for the integrity of the syllable and the onset and rhyme portions of the syllable in synonymic blends (e.g., shout and yell results in shell). The majority of such breaks fell between syllables and 70% of the breaks within a syllable fell immediately before the vowel, separating the onset and rhyme. Although the original data MacKay discussed were German, he claimed to have found similar results for English, with little separation of same-syllable consonant clusters and preferred divisions between syllable onsets and rhymes.

The possibility of different sized units participating in speech phenomena was specifically examined by Shattuck-Hufnagel (1983) with spontaneous exchange errors. She found that whole syllable and syllable rhyme errors were rare, as were single feature exchanges. Close to half of the several thousand entries in the MIT corpus of spontaneous speech errors involved exchanges within the word and the majority of these sublexical errors involved single segments. However, as Shattuck-Hufnagel pointed out, it is difficult to distinguish featural, segmental, and sublexical unit exchanges in the absence of phoneme clusters (e.g., bat for pat can be considered an error of the voice feature, the initial segment, or the onset). The consonants of onset clusters tended to function as units, as did sequences of vowel plus liquid. She found that whereas 66% of sublexical errors were accounted for with a segmental analysis, 71% followed from an assumption of onset and rhyme as units, and fully 81% were accounted for by assuming that onset, nucleus (vowel plus liquid), and coda were the processing units. In an experiment designed to elicit speech errors, MacKay (1978) also found that errors could be best accounted for by differences between consonant clusters and vowel nuclei (vowel plus liquid) in the prime and target words. Buckingham (1980) reported that neologistic jargon aphasia speech errors involved perservations of syllabic units such as onsets, nuclei, and rhymes.

The question of what units make up syllables was addressed more directly by Treiman (1983, 1985a, 1986), who capitalized on the possibility of altered speech with novel word games. Subjects were asked to transpose various parts of words and non-word syllables in order to assess which were the more 'natural' or preferred units. Children and adults could learn a variety of rules, but consistently preferred solutions based on the division of syllables into onsets and rhymes. Some evidence indicated that the nucleus and coda were also units. The vowel was grouped with the coda rather than with the onset and the division of the rhyme was before the coda rather than before the last phoneme, but the status of the coda as a unit remained somewhat unclear. Treiman (1983) tried a number of variations to determine whether the coda was separable but could not come to a conclusion. Part of the difficulty might have been with the words she used, which included rhymes with post-vocalic liquids that were analysed as vowel plus consonant cluster coda (V+LC) rather than as complex vowel nucleus and simple coda (VL+C). The results of MacKay (1978) and Shattuck-Hufnagel (1983) support the latter analysis and, with it, the coda might have emerged as an unambiguous unit in Treiman's study.

Indeed, Treiman (1984) found that post-vocalic liquids and non-liquids were treated differently, with post-vocalic non-liquids grouped more closely with the final consonant of the coda than with the vowel, and liquids grouped as often with the vowel as the final consonant. In a subsequent experiment, she found that post-vocalic liquids did seem to be treated as part of the peak (or syllable nucleus) whereas nasals were something between obstruents and liquids with respect to the closeness to the vowel. This pattern held not only for the spontaneous blending of two syllables to form new ones, but also for the relative difficulty of learning a rule requiring breaks after the vowel *versus* one with breaks between the two final consonants for syllables with post-vocalic liquids, nasals, and obstruents.

It would be interesting to know how people would divide up syllables if they were simply asked to do so. In Jaeger's (1980a) study, some Japanese subjects refused to divide syllables, as though to them a syllable was the smallest possible unit of sound in speech. Jaeger's English subjects could divide syllables, but there did not seem to be a consistent basis for the division (cf. the cat example mentioned earlier). If onset and rhyme, or even onset, nucleus, and coda are natural units of speech, then one would expect syllable divisions to be based on these.[11] Of course, segmental phonology (and the segment count experiments discussed in section 2.1) would predict segments as the basis of syllable division, with no preferences for any particular subgroupings.

Some information (far from conclusive) bearing on the question of syllable division and possible sub-syllabic units may have been incidentally collected in the course of an earlier study designed to examine the influence of orthography in phonological experiments. The second experiment of Dow (1981) was a simple segment count task (i.e., "How many speech sounds are there in each word?") completed by 40 subjects. No definition of what was meant by the term 'speech sounds' was given in order to determine what 'speech sounds' meant to the listeners (e.g., syllables, phonemes, or letters). The stimuli were such that syllable-counting would have resulted in scores of one or two for each word, phoneme-counting from two to five, and letter-counting from two to seven. The results were somewhat surprising in that, in addition to a group of subjects who had given mostly one's and two's as responses (syllabic criterion) and another with responses between two and six (phonemic criterion), there was a relatively large group (n=17) that had scores from one to four, which did not fit with any of the expected count criteria.[12] At

[11] The lack of preference between division before and after the vowel in CVC words could simply reflect equal weighting of the sub-syllabic units of onset, nucleus, and coda.

[12] The subjects were grouped based on counts for each word with the hierarchical clustering program of Wishart (1978).

the time, this group was designated as a Mixed Criterion group because it was suspected that the basis for the decision of the number of speech sounds might have been variable, that is, syllables for some words and some other criterion (perhaps segments) for the rest.

In the context of the evidence discussed earlier, however, the Mixed Criterion group begs for a re-evaluation. In phonological experiments it has been repeatedly found that naïve speakers can give judgments about such things as overall string similarity ratings in their language and that such data are highly reliable (see Derwing & Nearey, 1986). The presence of a large group (over one-third of the subjects) that seemed to be acting randomly when the rest of the subjects apparently had principled bases for their judgments should have warned the experimenter that the assumed taxonomy of possible 'speech sounds' might have been at fault. The notion of sub-syllabic phonological units provides a broader frameword within which to re-examine the Mixed Criterion group's responses.

Although the results were not as clear as might be desired, there is at least a hint that the subjects of this group were using something like onset and rhyme as a criterion.[13] For example, words of the form CVC had an overall mean score close to two, which could mean an onset-rhyme criterion. The results were less clear for other syllable shapes, however. Words with no coda (CV) had an overall mean score of only 1.38, as though about two-thirds of the cases were viewed as single units and the other third were treated as onset+rhyme. For words with codas, consonant clusters in the onset also seemed to affect the criterion, as the overall mean

[13] Only the aural presentation subjects (n=9) of the Mixed Criterion group are discussed. Written presentation may have had other influences on responses; 35 of the 120 words had significantly different aural and written means.

score for words of the form CCVC (2.27) was higher than for CVC words (1.95). Some of the subjects may have divided the rhyme portion of the words with complex syllable structure into nucleus and coda. There is some support for this in the words that had consonant clusters for codas (mean of 2.78).

There is some suggestive information here, but test cases to distinguish complex vowel nuclei from consonant cluster codas, for example, are missing. A more complete word list would therefore be required to test various syllable shapes against theories of syllable structure. The word list for the segment count task of Dow (1981) was designed for the rhyme task in the same thesis. Because of this, there were only 20 different rhyme shapes out of 120 words, and the form of these was further restricted because of the constraints imposed in that particular study. There is evidence that simple CVC syllables are much easier to segment phonemically than more complex syllable shapes. Perin (1983) reported that the more phonemes in a word, the lower the segmentation accuracy, and Hohn and Ehri (1983) reported that young children trained to segment CV, VC, and CVC syllables were unable to segment syllables with consonant clusters (CVCC and CCVC).

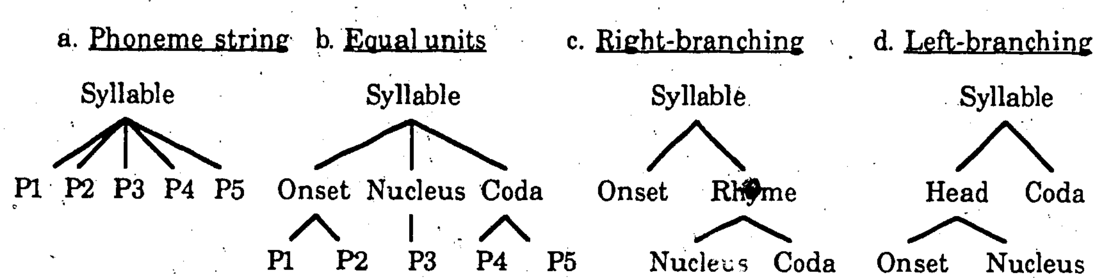
To adequately assess the reality of various sub-syllabic units, experiments would have to be designed specifically to include syllable shapes that would distinguish among various interpretations. For example, it is impossible to distinguish between an onset-rhyme and onset-nucleus-coda analysis with a counting task and words of the form CVL (where L is a liquid). Various models of syllable structure and ways to test them will be described in the next chapter.

CHAPTER 3: MODELS OF SOUND REPRESENTATION

One view of phonological organization at the sub-word level is that syllables are indivisible wholes. Alternatively, classical segmental phonology (e.g, Chomsky & Halle, 1968) views syllables as composed of strings of equally weighted phonemes (see Figure 3.1a). A third, hierarchical view of the syllable posits the existence of various kinds of intermediate units between the syllable and the segment. In one version of this model, only a single intervening level of equally weighted constituents is proposed (see Figure 3.1b), while others involve additional levels, either right-branching (as in the 'rhyme' model of Figure 3.1c) or left-branching (as in the 'head' model of Figure 3.1d).

<u>Figure 3.1</u>

Various models to describe syllable structure



The favored hierarchical structure for syllables from the metrical phonology tradition (i.e., Liberman & Prince, 1977; McCarthy, 1979; Selkirk, 1980;1984; Hayes, 1981) has a major branch between the onset and rhyme, with the rhyme branching further into the nucleus and the coda (see Figure 3.1c). A left-branching tree (see Figure 3.1d) has been proposed by Iverson and Wheeler (1987) to describe syllable structure and the non-hierarchical structure elaborated by Vennemann (in press) based on bonds and affinities also refers to the head and coda units. Each of these models of syllable structure will be discussed in detail in the sections below.

3.1. Syllables

The notion of syllables as units is one way to conceive of phonological organization at the sub-word level and has a long tradition, from Pike's (1947) distinction between 'chest pulse' syllables (phonetic) and phonological units (phonemic) through to Kahn's (1976) claim that the syllable is essential to an adequate explication of phonological processes in languages. Arguments have been presented for and against physiological (e.g., Pike, 1947; Kloster Jensen, 1963; Fry, 1964), acoustic (e.g., Malmberg, 1955; Bondarko, 1969; Scholes, 1967; Hoard, 1966), distributional (e.g., O'Connor & Trim, 1953; Pulgram, 1970), and other theoretical definitions of the syllable (e.g., Von Essen; 1951, Hooper, 1972; Anderson & Jones, 1974; 1977; Kahn, 1976; Lowenstamm, 1981; Clements & Keyser, 1983). Alternatively, it has been argued that the syllable is either unnecessary, impossible, or misleading ("harmful") and has no place in phonology (Kohler, 1966) and the major work of English segmental phonology (SPE, Chomsky & Halle, 1968) did not define syllables, or even include the term in the index. Fortunately for the syllable, Anderson (1969) convincingly rebutted Kohler's arguments, pointing out that they all related to difficulties of segmentation rather than to the existence (or usefulness) of units of syllable size. The problem of segmentation is not a small one, however, and will be discussed after the evidence for the importance and psychological reality of syllables is presented.

Mehler *et al.* (1981) argued on the basis of monitoring task results and infant discrimination that the syllable is the basic processing unit for language sounds. The evidence discussed earlier with respect to children (Liberman *et al.*, 1974; Fox & Routh, 1975) and illiterate adults (Morais *et al.*, 1979; 1986) supports Householder's (1971) claim that the most natural form of segmentation is syllabic rather than segmental. The meaning of this evidence is questionable, however, because, as Bell and Hooper (1978a) noted, although speakers are able to count syllables readily (and segment words into syllables), they often disagree as to the number.

Linguistically, syllables seem to be the most appropriate unit with which to describe prosodic processes in languages (Fudge, 1969). Although another unit, the mora, has been traditionally used to describe Japanese, McCawley (1968) claimed that the best way to predict accent placement in that language was actually in terms of the syllable. A variety of non-prosodic phonological processes can also appear "simpler, intuitively more meaningful and descriptively more adequate" if reference is made to the syllable (Kahn, 1976, p. 9). Some examples of rules that appear simpler when described with reference to the syllable are aspiration and tenseness of consonants (Hoard, 1971; Kahn, 1976), distribution of /s/ and /z/ in German (Basbøll, 1972), consonant devoicing (Vennemann, 1972), distribution of alveolar stops in English (Kahn, 1976), /r/-loss in English dialects (Kahn, 1976), and a variety of other assimilation processes (Hooper, 1972). This list is probably limited only by the interest of the investigators. The point is obvious that a wide variety of phonological processes seem to make more sense when the syllable is acknowledged as a unit.

Syllables have also proven to be very useful for modelling speech production and perception. The acoustic separation of phonemes, at least for CV syllables, is often very ambiguous, especially with regard to co-articulation effects.[1] Phone-sized units are difficult to use for synthesis because they require complex combination rules (Fujimura, 1975; 1976; Fujimura & Lovins, 1978), whereas syllables have been used to synthesize relatively good quality speech. As Gresser and Mercier (1975) and Zwicker, Terhardt, and Paulus (1979) found, syllables are also very useful in speech recognition, with automatic segmentation of speech into syllables quite good and simple.

Fudge (1969) observed that syllables have never been totally dismissed as possible units simply because we cannot do without them. One obvious place where syllables manifest themselves is in the world's orthographies. There are numerous syllabaries in use today and, according to Gelb (1952), syllabaries are a necessary step in the development of alphabetic systems. In keeping with Gelb's claim that syllabaries are precursors to alphabet systems and assuming that the syllabic level is more accessible than the phonemic, Gleitman and Rozin (1973) taught kindergarten children to read using a syllable system. They had great success, even with children who had not done well in alphabetic reading, and recommended using syllables as a natural introduction to reading. The information available from cultures with syllabaries also supports the naturalness of syllable

[1]There is little separation at the production level, either. Fromkin (1966) claimed that there was no simple correspondence between phonemes and motor commands because neural muscular correlates to phonemes differ with phonetic context. One explanation she suggested for this was that the minimal linguistic unit was larger than the phoneme.

units for symbol-sound correspondence. Japan has both a syllabary and logographic system and has a low level of illiteracy (Sakamoto & Makita, 1973), as did the Cherokee Indians with the Sequoyah syllabary (Walker, 1969).

A syllable-like unit (the vocalic center group or VCG) has also been proposed as the basic unit of visual word recognition. Hansen and Rodgers (1968) suggested the VCG as the minimal pronunciation unit, with one vocalic element and zero to three consonants before and after the vowel. Since they claimed that no valid definition of the syllable exists, the VCG, which was distributionally defined, was a more reasonable unit to explain perceptual parsing of printed English. In a test of report accuracy, Spoehr and Smith (1973) found that letters which comprised something similar to a syllable had correlated accuracy scores (as would be expected if the syllable functioned as a perceptual unit). The results did not clearly support a syllabic interpretation, however, and Spoehr and Smith argued instead for the importance of the VCG in perceptual processing.

The notion of the VCG nicely avoids the problem (and probably the reality) of the indeterminacy of syllable boundaries. This is the segmentation problem Kohler (1966) used to argue against the syllable as a phonological entity. Pulgram (1970) proposed a number of principles (including minimal coda and maximal onset) to determine syllable boundaries, but many linguists seem to prefer leaving the boundaries unclear. Thus, Anderson and Jones (1974; 1977) proposed that the boundaries are pushed out from the center syllabic as far as possible to be compatible with constraints on well-formed syllables with overlap allowed. Kahn (1976) strongly advocated ambisyllabicity, but provided certain defining conditions on interlude possibilities, and Jones (1976) claimed that medial clusters should be interpreted as having simultaneous membership in both the preceding and following syllables, or else they would be highly marked.

In her experimental investigation of syllabifications, however, Fallows (1981) found little evidence for ambisyllabicity except for a small tendency in that direction on the part of liquids and nasals. Bailey (1978) argued that there is no validity to the term 'ambisyllabic' because syllabic membership was dependent on rhythm and tempo. According to him, syllabification constraints are dropped with rapid tempo, making it appear that medial consonants belong to both preceding and following syllables. Native speakers' intuitions about syllable boundaries will be difficult to determine, if these are dependent on some combination of phonotactic constraints, tempo, and rhythm. It may be exactly as Anderson and Jones (1974; 1977) suggested: overlap is allowed and all consonants are subordinate to the syllabic unit, which is central.

The accessibility of the syllable to adult native speakers has certainly been well documented - by segmentation studies, by Dow's (1981) syllabic criterion group, and by numerous general intuitive comments about speech sounds.[2] It may be that the notion of a syllable is simply derived by analogy from monosyllabic words and subparts of compound words. This is supported by the history of syllabic writing systems, nearly all of which developed from logo-syllabic systems (Gelb, 1952). One could just as reasonably argue, though, that words of syllable length exist by virtue of

[2]Although note that very young children do not segment syllables accurately with consistency (e.g., Liberman *et al.*, 1974; Fox & Routh, 1975) and illiterate adults do not segment speech into syllables as well as literates do (Morais *et al.*, 1986).

the fact that syllables are natural units and are therefore the early words.[3] If the syllable emerges as a unit, the reason may remain indeterminable. All that can be assessed is whether the syllable functions as a more central phonological unit than some of the others (e.g., phoneme, coda, rhyme) that have been proposed and for which empirical evidence has been found. Tasks which require segmentation of words and manipulation of various parts of multisyllabic words could be used to compare the relative accessibility of these units.

3.2 Sub-syllabic Units

3.2.1 Onset-rhyme. Sub-syllabic units were discussed above with respect to speech production errors and synthetic speech. Three different possibilities for syllable division will be discussed in this section. Probably the most intuitively appealing division of the English syllable is into the initial onset (consonant or consonant cluster) and rhyme portions. Bell and Hooper (1978a) pointed out that groups of segments are more accessible than individual segments for language use, language play, and language art. English 'Pig Latin', for example, requires the movement of the word-initial consonant or consonant cluster (i.e., the onset) to the end of the word, plus the addition of the suffix /ey/ (e.g., sprint becomes intspray /ɪntsprey/). Rhyme in English involves the nucleus and coda of the word-final syllable. Rhyme is an important part of our culture and children produce rhymes relatively early (some say as early as age 3).[4]

[3]The same arguments might be made for the notion of 'rhyme', since there are certainly real words that consist only of rhymes (i.e., no onsets).

[4]There is some disagreement in the reading literature as to the naturalness of the ability to rhyme and indeed whether all children can rhyme. For example, Bradley and Bryant (1978; 1983) reported that reading disabled children lack the ability to rhyme, a difficulty

People apparently attend to rhyme preferentially. In Nelson and Nelson's (1970) study, pairs of words with identical rhyme elements were rated as more similar than other pairs that had the same number of phonemes in common. Of the subjects in Vitz and Winkler's (1973) similarity of sound study, 25% reported that the presence of rhyme formed the basis of their similarity judgments. Treiman (1983, 1985a; 1986) found that onset and rhyme were the preferred units for language games in an experimental setting for both adults and children. According to MacKay (1972), synonymic blends in both German and English tend to have breaks immediately before the vowel, thereby separating the onset and rhyme (as in the example previously cited of shout and yell becoming shell). Morais *et al.* (1986) found that both Portugese literates and illiterates generalized single segment (vowel and consonant) deletions to complex onset and rhyme deletions. There is also evidence that people visually process words in terms of onset and rhyme (Santa *et al.* 1977; Treiman & Chafetz, to appear).

The appeal of onset and rhyme as sub-syllabic units probably also relates to the ease of describing the division (i.e., between the pre-vocalic consonant and the vowel), at least for monosyllabic words, and speakers have no difficulty producing rhymes for these (see Dow, 1981). For polysyllabic words, the problem becomes more difficult; it is not always clear where the first syllable's rhyme ends and the second syllable's onset begins, although some linguistic guidelines do exist. Most phonologists who have discussed syllabic structure are in favor of the notion of maximal

with sound categorization which they claimed was the cause of their reading problem, whereas Rack (1985) found no difference between dyslexics and control subjects in the effect of rhyme on recall of auditorially presented words or on a rhyme judgment task.

onsets consistent with constraints on word-initial consonant clusters (e.g., Vennemann, 1972; Kahn, 1976; Bailey, 1978; Pulgram, 1970). Kahn claimed that slow, over-precise syllable-by-syllable speech shows a strong tendency to syllabify in this way, and Fallows (1981) supported this experimentally, reporting that maximal onsets occurred in 71% of syllable divisions in her study. However, the principle of maximal onset is less useful for unstressed syllables which seem to lose their consonants to stressed syllables (see Hoard, 1971; Bailey, 1978).

3.2.2 Onset-vowel-coda. A second possibility for syllabic divisions partially overcomes the problem of segmenting polysyllabic words because it divides the rhyme portion into vowel plus coda (or interlude). In a monosyllabic word the coda consists of all consonants after the vowel, just as the onset consists of all consonants before the vowel. For polysyllabic words, the division would be into word-initial onset, first vowel, interlude (consisting of the coda of the first syllable and onset of the next), second vowel, and so on until the word-final coda. According to Hockett (1955; 1958), who is apparently responsible for the term 'interlude', the interludes are onset-like and coda-like at the same time. The notion that inter-vocalic consonants may belong simultaneously to two adjacent syllables is also advocated by Jones (1976), who claimed that this bracketing of segments can be motivated by the role such units play in phonological processes.[5]

It may be that native speakers also perceive interludes as belonging to both preceding and following syllables. Under such as assumption, a CVCCVC word would be composed of 6 units (i.e., C_1-V_1-$C_2C_3\$C_2C_3$-$V_2$-$C_4$

[5]The arguments for phonological rules and syllable units were presented in the discussion of the syllable in section 3.1.

where \$ is a syllable boundary and the interlude cluster (C_2C_3) counts twice, once for each syllable). Alternatively, the interlude may be broken down into coda and subsequent onset (i.e., C_1-V_1-C_2-C_3-V_2-C_4), again with 6 units. Notice that both of these counts are indistinguishable from a phoneme count, which is why words with consonant clusters in the onset and coda must be tested in some other way to distinguish the two theories. Finally, it may be that the interlude is treated as ambisyllabic but only counted once, thereby resulting in a count of 5 for a CVCCVC word.

Clarification is needed not only on the matter of how interludes are interpreted but also on the question of whether consonant clusters are perceived in the same way independent of their position in the syllable. The distributional properties of consonant clusters differ with syllable and word position (although often syllable and word position constraints are the same). For example, it has been suggested for a number of reasons that /s/ plus stop clusters in English should be treated as single or special units (e.g., Davidsen-Nielsen, 1974; Basbøll, 1974; Fujimura, 1975; Cairns & Feinstein, 1982; Selkirk, 1984). Whether these clusters function as units to break up onsets, interludes, or codas (e.g., st-rike, textile /tek-st-ayl/, glimpsed /glɪmp-st/), or whether they combine with other consonants to form larger units that fit the onset-vowel-coda model should also be investigated.

3.2.3 Onset-nucleus-coda. The distributional arguments presented for the division into onset-vowel-coda also hold for the third possibility of syllable division with one important qualification. There are reasons to suggest that the nucleus can consist of more than just the vowel (or vowel plus glide, as suggested by Hockett, 1955). Pilch (1966) argued for English syllable nuclei of vowel plus /w/, /y/, or /r/ on distributional grounds. He did not include /l/ in the nucleus because it did not pattern the same as /r/ and left pre-vocalic /y/ unspecified because he claimed that there was no principled reason to assign it to either the onset or the nucleus. The same analysis of the nucleus was also advocated by Hultzén (1965) on distributional grounds, and by Kahn (1976) and Clements and Keyser (1983) for reasons of rule simplification. The classic phonemic discovery handbook of Pike (1947) was vague on the details of the syllabic nucleus, saying only (in the Appendix) that it was convenient to use the term nucleus to include the vocoid and a consonant when "this sequence serves as an inner structural entity distinct in distribution from marginal elements which precede and/or follow it". A later version (Pike, 1967) suggested that voiced sonorants such as /l/ and /n/ could be included as part of the nucleus, but surprisingly, in light of other discussions of the problem, not /r/.

There are more than purely formal, distributional reasons to suggest that complex vowel nuclei are more natural units than simple vowel segments, separate from post-vocalic glides and liquids. Shattuck-Hufnagel (1983) found that 81% of sublexical speech errors could be accounted for by assuming that the onset, the vowel-plus-liquid nucleus, and the coda were units, and she confirmed the importance of complex nuclei in a later study (1986). Experimentally induced errors (MacKay, 1978) also involved syllable nuclei (vowels plus liquids) and consonant clusters (onsets and codas).

In fact, an analysis of syllable structure with a nucleus rather than a simple vowel seems to explain some unclear results with respect to syllable codas. Treiman (1983), for example, found some evidence that the coda functioned as a unit, but the results were not as clear as for the onset. The explanation for this probably relates to the words that she used, which included post-vocalic liquids which were uniformly treated as part of the final consonant cluster. A later experiment (Treiman, 1984) showed that subjects treated liquids as grouping closely with the vowel, leaving a consonant coda. Words with post-vocalic liquids were treated as though the division was between the liquid and the following consonant whereas words with post-vocalic non-liquids were treated as though the division was immediately the vowel. A similar, apparently anomalous result in Santa *et al.*'s (1977) experiment with visual word recognition was that initial consonant clusters (onsets) but not finals (codas) functioned as units. Some of their words also contained liquid plus consonant 'codas' that were treated as simple consonant clusters. If these had been analysed differently, as complex nuclei plus consonant, then the expected unitization might have emerged.

Zwicker *et al.* (1979) claimed that segmentation problems in speech recognition could be reduced considerably by using sub-syllabic units. German city names were segmented into syllable onset, nucleus, and coda and, although the results were far from perfect, the use of consonant clusters and vowel nuclei reduced segmentation problems and facilitated the use of a look-up lexicon. Ruske and Shotola (1978) also reported high recognition accuracy (75%) using a similar segmentation for German city names.

There is some controversy as to what should be included as part of the vowel nucleus. For example, pre-vocalic glides (e.g., the /y/ of pew /pyu/)

might be thought of as either part of the onset or part of the nucleus, as Hofmann (1966) suggested. Segment counts from Dow's (1981) speech sound task were consistently lower for words with pre-vocalic glides than the count predicted by treating the glide as a separate segment, and rhymes did not necessarily include the glide, leading to the suggestion at the time that the pre-vocalic glide belonged to the preceding consonant. Shattuck-Hufnagel (1986) reported that speech error evidence did not support a clear interpretation of the pre-vocalic glide.

Post-vocalic segments also pose certain problems, as suggested in the previous discussions. There are a number of arguments for post-vocalic /r/ as part of the nucleus and although some people have included the general category of liquids in the nucleus (e.g., MacKay, 1978; Shattuck-Hufnagel, 1983), the status of /l/ needs to be clarified. Stemberger's (1983) analysis of speech errors showed that /r/ and /l/ distributed differently from both consonants and vowels and he suggested that post-vocalic liquids were best considered part of the nucleus, whereas post-vocalic glides were part of the vowel. Rarely mentioned in this context are the nasals. Although distributionally they may be considered part of the coda, it may be that post-vocalic nasals are also perceived as part of the nucleus. In some dialects of English, at least before voiceless stops, nasals are deleted, leaving the length and nasalization information on the vowel. Certainly Read's (1971) evidence of children's invented spellings supported this, as the nasals were consistently omitted pre-consonantally. Treiman (1984) found that nasals were intermediate between liquids and obstruents in terms of closeness to the vowel. On the other hand, Vitz and Winkler (1973) relatively successfully used a simple consonant cluster index (which grouped post-vocalic nasals with the following consonant) to predict phonetic similarity judgments.

All of these possibilities of sub-syllabic analyses need further evaluation using a variety of tasks. To test the psychological reality of sub-syllabic units in production, monosyllabic and polysyllabic words would have to be included in tasks such as word segmentation (specifying number and location of divisions), counting the number of units in words, and manipulation of various units to compare possible units. Combinations of segments which do not form constituents would also have to be evaluated. Speakers' judgments of words with complex syllables (as opposed to simple CVC syllables) with a variety of different pre- and post-vocalic consonant clusters would have to be assessed to determine which of the three descriptions of sub-syllabic units is correct. A non-production-oriented task (such as phonetic similarity judgments) would also have to be used to lend support to the suggestion that sub-syllabic units are important in speech perception.

3.3 Structural Models

Perhaps most deserving of the descriptor 'model' is the form of representation proposed by metrical phonology. Syllables and hierarchical structure are integral parts of this theory, which was originally motivated to explain phrasal and word stress in English (see Liberman, 1975 or Liberman & Prince, 1977). The hierarchical structure organizes syllables, words, and syntactic phrases into binary branching trees. Stress is assigned, nodes are labelled, and then the tree is built up. Syllables with 'long' vowels or final consonant clusters are stressed and the [illegible] are labelled strong (S); syllables with lax vowels and only single final consonants are unstressed and the relevant nodes are labelled weak (W). The relative prominence of each syllable and word is therefore represented in the hierarchical structure (i.e., a direct line of S[trong] nodes projects to the primary stressed syllable).

Although neither the syllable nor the foot (minimal unit with some prominence, either mono- or disyllabic) were considered primitive categories by Liberman and Prince (1977), both were by Selkirk (1980). She went further, arguing for onset and rhyme constituents of the syllable and for the internal structure of the rhyme. Having the foot as a primitive category meant that the stress feature Liberman and Prince used was no longer required and reference to syllable structure rather than vowels and consonants eliminated disjunction in the stress rule.[6] As McCarthy (1979) noted, having units arranged in hierarchical fashion with rules referring to the hierarchy eliminated the need to refer directly to the constituents. Thus the apparent disjunction of Liberman and Prince's (and Chomsky and Halle's) main stress rule could be described in terms of whether the rhyme portion of the syllable branches.[7]

The end nodes of metrical syllable structures are the onset and rhyme units described in the last section, or perhaps even the onset, nucleus, and coda units, since Selkirk (1980) referred necessarily to those constituents in her explanation of stress assignment. An important, potentially testable difference between a simple string of sub-syllabic units and the metrical hierarchy was introduced by Hayes (1981). He supported the branching

[6]Recall that only syllables with long vowels or final consonant clusters receive stress.
[7]Note that long vowels must be treated as underlying geminates for this to work.

rhyme notion of stress assignment and, to account for the stress of certain exceptions, proposed that some types of words had a final "extrametrical" unit (rhyme, morphological affix, and segment, respectively, for each of nouns, adjectives, and words with certain final consonants). This notion was expanded upon by Giegerich (1985), who defined the structure of the syllable more precisely. According to him, onset and rhyme were compulsory constituents with the minimal syllable structure of a weak and a strong branch. Only the conditions of increasing and decreasing sonority from the vowel nucleus had to be met (after Kiparsky, 1981), with any amount of overlap allowed between syllables. Thus, bolt would consist of only one syllable whereas bottle would consist of two, because the /lt/ of the former word is in the direction of decreasing sonority whereas the /tl/ of the latter word is increasing sonority.

Sonority has also been used to describe syllable structure with no hierachical implications (e.g., Vennemann, 1972; Hooper, 1972; Basbøll, 1973; Anderson & Jones, 1974; Selkirk, 1984). Generally, vowels or vowel-like segments constitute the nucleus of a syllable, with segments ordered with decreasing sonority out to the syllable margins (e.g., vowel-glides-sonorants-fricatives-obstruents is the order discussed by Selkirk, 1984). However, the principle of decreasing sonority to define the end of the syllable coda does not work for words like ox /ɔks/ or sixths /sɪksθs/, which have a shift in the direction of sonority in the coda /ks/ yet are clearly mono-syllabic.

Selkirk (1984; after Halle & Vergnaud, 1980) tried to circumvent this problem by claiming that syllable-final coronal obstruents are best considered appendices and not part of the basic template for English syllables. Giegerich (1985) took Hayes's (1981) notion of extra-metricality even further by suggesting that *all* word-final consonants were extrametrical and became part of the next syllable's onset. (If there was no following word, the consonant was still extra-metrical and not considered part of the metrical syllable.) This ensured that the condition of obligatory onset was met and at least for words like *ox* solved the problem of sonority shift. It is not clear, however, what Giegerich would claim about words like *sixths* which have a shift in sonority before the last consonant of the syllable.

Another problem with Giegerich's version of extra-metricality is that the onset cluster of the following word that results from the addition of the extrametrical consonant may not meet the phonotactic constraints of word-initial clusters (e.g., *tilt the* /tɪlt θʌ/ becomes /tɪl tθʌ/, with the second word's onset phonotactically impermissible). Fallows' (1981) results argued against this analysis, since phonotactic constraints were met 98% of the time in subjects' responses for a syllabification task . Selkirk's (1984) notion of appendices (coronal obstruents) accounts for some of the problem cases for vowel sonority which motivated Giegerich's version but does not have the weakness of phonotactically impermissible onsets and is stated clearly enough to be testable. The basic division of syllables into onset and rhyme would be expected, along with some indication that syllable-final coronal obstruents were not part of the coda.

There is at least one other non-hierarchical model of syllable structure besides the one based on sonority. Vennemann (in press) has proposed that bonds and affinities between adjacent segments (described in terms of laws) account for the grouping of segments within the syllable.

Affinities between /s/ and plosives are very strong, for example, whereas those between /t/, /d/, /θ/ and /l/ in English are very weak. He proposed a 'body bond' which tied the onset to the nucleus (resulting in two units - the head [Vennemann's 'body'] and the coda). Iverson and Wheeler (1987) have also propounded a theory based on the pre-eminence of the head, as opposed to the rhyme, with a left-branching structure describing the head and coda at one level and further dividing the head into onset and nucleus. Although this is very different from the onset-rhyme analysis discussed above, there is some external support for such a description of syllables, at least for some languages, and Taylor (1981) claimed that old written Egyptian (between 3,000 BC and 400 AD) had head-type signs.

3.4 Core plus Affix

The core plus affix model has similarities to some of the others, although its proponents (Fujimura, 1975; 1976; Fujimura & Lovins, 1978) motivated the core-affix distinction by speech synthesis and co-articulation effects, rather than on the basis of distributional or psychological arguments. The syllable core was defined as the vocalic nucleus together with all consonants having a "inherent cohesion" to it. The principle of vowel affinity (similar to the sonority hierarchy discussed above) determined the ordering of consonants out from the nucleus. 'Affixes' were apical segments that agreed in voicing with the core-final elements. This handled such morphological endings as /θ, z, s, d, t/ in English, which were apparent exceptions to the vowel affinity principle. Fujimura's category of affix differed from the morphological term bearing the same name in that is also included final consonants for which there was no morphological justification (e.g., the final /t/ of tent would belong with the core but the /d/ of tend would be an affix). Initial /s/ plus stop clusters were also considered to be single consonantal elements, with one place feature and a spirantization feature. (Recall that Davidsen-Nielsen (1974), Basbøll (1974), and Barton *et al.* (1980) also suggested that /s/ plus stop clusters might be more appropriately analysed as single elements).

Fujimura's affix was similar in some ways to Selkirk's (1984) appendix, although the motivation was different. Fudge (1969) also treated word-final apical segments separately from other consonants of the coda, based on distributional arguments. Thus, some linguistic motivation exists for Fujimura's model and, given its apparent usefulness for speech synthesis, it is obviously worth testing as a candidate for phonological representation. Predictions for word segmentations with the core plus affix model differ from any of the other models discussed so far. Of course, words with final apical consonants that do and do not agree in voicing with the core-final consonant would have to be included in a stimulus list, as would words with /s/ plus stop clusters. Except for the special treatment of affixes and these clusters, segmentation of words should be into syllables. (An example given by Fujimura and Lovins (1978) was the English word sixths, which they analyse as having a core /sɪk/ and the three affixes /s/, /θ/, and /s/, which would presumably receive a unit count of four.)

The difference among the models of syllable structure would lie in the treatment of word-final consonants. Most sub-syllabic models would predict that the word-final consonant would be part of the last syllable's coda, whereas a simple syllable model would predict no special treatment for any consonants (except perhaps ambisyllabic medials). The metrical model would predict that all word-final coronal obstruents would be separate from the rest of the syllable. The core plus affix model would predict that apical segments which agree in voicing with the core-final element would be separate from the core.

3.5 Phonemes

The reality of phonemes for native speakers was assumed by all of the studies reviewed in the preceding chapter. The phoneme was probably 'discovered' by the inventors of the Greek alphabet and was described as such by Henry Sweet or Baudoin deCourtenay in the nineteenth century (see Krámsky [1974] for a review of the history of the phoneme) and has been viewed as the minimal unit of sound throughout the 20th century, at least by European and North American linguists (e.g., Trubetzkoy, Jakobson, Hjelmslev, Sapir, Bloomfield, Chomsky and Halle). The economy of phonemes to capture distinctive differences in language sounds and to describe sound distribution, and the relative ease of describing phonological processes with phonemes (and later with features) is well-known and argues for their inclusion in a description of the sound structure of language.

However, a number of the results reported in preceding sections call into question just how basic a unit the phoneme is. Certainly it seems to be real to adult literate speakers when they are asked to perform tasks that focus on the level of differences between phonemes. Too often phonemes have been 'proven' with clear cases, where there is little chance to analyse speech into anything but phonemes. For example, people asked to segment CVC words might normally not be inclined to think of CVC's as divisible, but when task instructions suggest that it can be done, people may comply by using a segmental or phonemic criterion.[8]

The other difficulty with claiming that phonemes are the basic unit of sound representation is that of distinguishing them from letters. Rather than being derivative of phonemes, English orthography may be responsible for leading speakers to the perception of speech as a sequence of phonemes. Using standard pen and paper tasks with adults, therefore, is somewhat questionable for phonological experiments and the evidence from such tasks to date clearly shows the influence of orthography on certain types of judgments, particularly in the case of segments whose phonological status is not clear. All such tasks, however, have assumed that speech is analysable as a string of phonemes and have therefore guided the subjects to perform in that way. Very little study has been directed to determining how naïve speakers spontaneously perceive the sound structure of their language.

Morais (1985) pointed out that although phonemic analysis did not seem to be natural for illiterates and non-alphabetic literates, most individuals could be taught to perform phonetic manipulations with relatively little training. He suggested that it was difficult to believe that this would be possible if phonemic analysis was simply an artifact of linguistic analysis. Speech error data (e.g., Shattuck-Hufnagel, 1986) and phonetic similarity judgment results (e.g, Derwing & Nearey, 1986) also suggest that the phoneme may have a role to play in such tasks, regardless of whether there is overt awareness of the concept on the part of the speaker/listener.

[8]Of course, this is not always so, as Jaeger's (1980a) discussion and Experiment 2 of Dow (1981) show. Syllable and sub-syllabic divisions seem possible even for CVC words.

To further investigate the psychological reality of phonemes, therefore, a task should be used which does not force a phonemic interpretation and which is minimally influenced by orthography. A candidate task is the phonetic similarity judgment task described in Derwing and Nearey (1986). In other words, it may be that only the experimental techniques need to change, not the speakers or the language. It would be preferable to have evidence from different subject and language groups (e.g., non-alphabetic, pre-literate, literate alphabetic) but, failing that, a variety of production-oriented and perception-oriented tasks could be used with children and adults to contrast the accessibility of phonemes with that of other proposed sub-syllabic units.

CHAPTER 4: EXPERIMENTAL WORK

Experiment 1

Counting Task

The current evidence supports the notion of sub-syllabic units but does not definitively discriminate among the five models presented in Chapter 3. Recall that the first model is that words are composed of non-analysable syllables that lack any psychologically real subdivisions. The second type of model proposes that syllables consist of linear sequences of intermediate units (e.g., onset-rhyme, onset-vowel-coda, or onset-nucleus-coda) which can be described with reference to the the vocalic element or syllabic nucleus (which has been described as consisting of the vowel and the post-vocalic sonorant). The third type of model assumes that sub-syllabic units such as the onset, rhyme, vowel or nucleus, and coda are arranged in some kind of structural or hierarchical fashion. Some versions of the metrical model also treat syllable-final coronal obstruents as separate 'appendices' (Selkirk, 1984). Although the right-branching model is more prevalent in the literature, a left-branching model consisting of a head (the onset and vowel) and coda has also been proposed (Iverson & Wheeler, 1987) and a flat structure of head and coda has been proposed by Vennemann (in press). The fourth model comes from the field of speech perception and treats a syllable as a core plus an affix, where the latter is an apical consonant which agrees in voicing with the core-final segment (Fujimura, 1975; 1976). Finally, the last model, long accepted as the standard by traditional phonology, is that words are strings of equally-weighted phonemes. Examples of words segmented as predicted by each of these models can be found in Table 4.1.1

Table 4.1.1

Segmentations according to different models of sound representation

Models	Sample Words and Segmentations		
Syllable	drift	art	so-da
Sub-syllabic units:			
Onset-rhyme	dr-ift	art	s-o-d-a
Onset-vowel-coda (OVC)	dr-i-ft	a-rt	s-o-d-a
Onset-nucleus-coda	dr-i-ft	ar-t	s-o-d-a
Structural:			
OVC+Appendix*	dr-i-f-t	a-r-t	s-o-d-a
Head-coda	dri-ft	a-rt	so-da
Core+affix	drif-t	art	so-da
Phonemes	d-r-i-f-t	a-r-t	s-o-d-a

* the 'OVC+Appendix' analysis refers to the metrical model of onset-vowel-coda and a separate appendix of word-final coronal obstruents

The main criticisms of the phoneme model do not have to do with either the usefulness or efficiency of describing words in terms of phonemes. Rather, the evidence suggests that phonemes may be learned or externally-imposed representations, rather than natural units ~~determined by the brain's chunking of~~ information or the phonotactic patterns of the language itself. One likely form of external influence is the particular orthographic tradition that may be associated with a given language. It has already been established that non-literate individuals seem to perceive and/or manipulate syllables more readily than phonemes (Morais *et al.*, 1979), which may reflect the relatively greater accessibility of syllables as phonological units. Testing pre-literate children's phonological judgments is one way to address the issue of the form of phonological knowledge before formal instruction in the traditional representation of language sounds begins. The kinds of tasks that have been used to compare syllable and phoneme perception could be used to assess the other possibilities for phonological units proposed by the models discussed so far.

One of the studies often referred to in discussions of young children's phonemic awareness is that of Liberman *et al.* (1974). They had children use a wooden dowel to tap out the number of phonemes or syllables in words. The children were trained with four triads of stimuli and the test was presented only after the children could do the training trials correctly. The test items consisted of words of simple canonical structure (i.e., no complex vowel nuclei or consonant clusters), containing one to three phonemes or one to three syllables. The tapping task of Liberman *et al.*'s study has been criticized because it may have biased the results in favor of a syllable count over a phoneme count, given the natural association of pulses of air with the motor action of rhythmic tapping. Fox and Routh (1975) coaxed children to make increasingly finer segmentations with the request to "Say a little bit of (word)", but their procedure proved to be quite tedious for both the children and the experimenter. The technique of setting out poker chips or counters to match the number of segments would seem to be a more neutral response than dowel tapping and was used successfully by Ehri and Wilce (1980) and by Treiman and Baron (1981).

All of the segmentation studies reviewed here included training sessions before the test sessions, partly to teach the procedure of segmentation, but also to focus attention on the desired level of analysis (phoneme or syllable). In the present experiment, training was carried out for syllable segmentation, onset and rhyme segmentation with open syllables,[1] and phoneme segmentation. None of the test cases for the models (coda-closed syllables, complex vowel nuclei, or apical final consonants) were used for the training trials, but words with consonant cluster onsets were included in all training sets in order to differentiate between onset-vowel and phonemic segmentation (e.g., st-ay *vs* s-t-ay). The response measure was the number of correct segmentations (as determined by the number of counters set out by the children) according to a syllabic criterion, different sub-syllabic criteria, and a phonemic criterion.

If syllables and sub-syllabic units are natural and phonemes learned, there should be differences in performance on the phoneme counts between pre-literate children and children who have learned phoneme-grapheme correspondences. There may also be developmental differences in the way children treat sound units in their language. To control for the possibility of developmental *versus* experiential differences, a very special type of control group was required. In fact, just such a group existed by virtue of the cut-off dates for eligibility to enter Grade One.[2] Specifically, in the Edmonton Public School System, children with a birthday before March 1st were eligible to enter Grade One the year they were five, but

[1]The decision was made to do an onset-vowel division in order to provide some guidance to the subjects as to the level of analysis required (that is, not syllabic), although this was biased against the head-coda and core+affix models.

[2]This particular selection of subjects was part of a larger experiment that Dr. F. Morrison was carrying out in Edmonton to investigate the general question of developmental *versus* experiential changes with schooling. The data from the children for this and the following experiment (Experiments 1 and 2) were obtained as part of a battery of tests given to these children as a pre-test. The post-test phase necessitated the second control group.

those born on or after March 1st had to remain in Kindergarten. Thus, if children two to three months on either side of the cut-off date were included, two groups of children resulted that were very close in age but with different educational experiences. One more group was needed to complete the picture: children exactly one year older than the Kindergarten group but in Grade One, who would therefore have had the same experience as the younger Grade One group but we[illegible] year older than the Kindergarten group. High School students served as an adult comparison sample.

Method

Subjects. There were 16 Kindergarten children (7 girls, 9 boys, mean age of 5 years, 8 months), 16 younger Grade One children (10 girls, 6 boys, mean age of 5 years, 10 months), 16 older Grade One children tested (11 girls, 5 boys, mean age of 6 years, 8 months), and 16 High School students (9 girls, 7 boys, aged 16 to 18). The study was explained to the parents of the children and consent was obtained before any testing was carried out. All of the Kindergarten and Grade One children were enrolled in the Edmonton Public School System and were native speakers of Canadian English. The High School students were enrolled in the American High School in Kilchberg, Switzerland and were native speakers of American (n=10), British (n=5), or Canadian English (n=1).

Materials. Two word lists for the test portion were prepared, including words of the form shown in Table 4.1.2 (see Appendix A for the complete lists). The words were chosen such that there was a variety of different consonant clusters and post-vocalic consonants (i.e., including /r/ and /n/ which have been described as part of the vowel nucleus). All words were of a frequency greater than one occurrence in a million tokens of written text (with standard frequency index (SFI) values of greater than 45 from the Carroll *et al.*, 1971 frequency list).

Table 4.1.2

Predicted number of units for various canonical syllable types

	Sample syllable types and words						
	VR	VRC	VC(C)	[illegible]	(C)CVC	(C)CVRC	(C)CVCC
Models:	ear	ink	ask	[illegible]	gas	fence	drift
Syllable	1	1	1	1	1	1	1
Sub-syllabic units:							
Onset-rime	1	1	1	2	2	2	2
Onset-vowel-coda	2	2	2	3	3	3	3
Onset-nucleus-coda	1	2	2	2	3	3	3
Structural:							
OVC+Appendix *	2	2	2	3	3	4	4
Head-coda	2	2	2	2	2	2	2
Core+affix	1	1	1	1	1	1	2
Phonemes	2	3	3	4	3	4	5

* the 'OVC+Appendix' analysis refers to the metrical model of onset-vowel-coda and a separate appendix of word-final coronal obstruents.
C=Consonant V=Vowel R=/r/ or /n/

One of the word lists was used for the syllable counting task, with ten each of one-, two-, and three-syllable words, randomly ordered. The other word list was used for the sub-syllabic unit and phoneme counting tasks; this list included 15 words (ten monosyllabic and five disyllabic) from the first list, as well as 15 other monosyllabic words with various consonant clusters. The words ranged from two to five phonemes in length or, depending on the model, from one to four sub-syllabic units. Two separat randomly ordered lists were prepared from these 30 words, one for the su syllabic task and one for the phoneme task.

There were training sets for each segmentation task consisting of fo triads of words (see Appendix B). Each triad was ordered in terms of a decreasing number of units. For the syllable counting there was first a trisyllabic, then a disyllabic, then a monosyllabic example given (e.g., ba na-na, he-ro, my); for the sub-syllabic counting, first a disyllabic word segmented with open syllables, then a monosyllabic word with no coda, then a word consisting only of a vowel (e.g., fr-i-sk-y, fl-y, a); for the phoneme counting, a four- or three-phoneme word, then a two-phoneme word, then a one-phoneme word (e.g., s-p-r-y, b-ay, owe). All of the training words were segmented as open syllables (i.e., no ambisyllabic consonants), and many had initial consonant clusters. Large poker chip were used as counters for the children. A puppet was available if boredo became a problem, but it was not an integral part of the task.

Procedure. The children were tested individually in a quiet room in their school building.[3] The three counting tasks were part of a battery done over three sessions and were given at the beginning of the sessions. In order to minimize any effects due to order of presentation, the three counting tasks were given in separate sessions at least a day apart. The syllable counting was given in the first session, then the sub-syllabic, an finally the phoneme counting task.[4]

[3]Thanks to MargoLee Horn for helping with the testing.

[4]The lists were presented in this order to prevent discouraging the children with the hardest task first, since it was known from other studies that the phoneme counting

A similar procedure to Liberman *et al.* (1974) was followed for the training trials. The four triads of training stimuli, consisting of one to four units, were presented to the children with modelling. The children were asked to put out a poker chip for each unit (syllable, sub-syllabic unit, or phoneme) in the word. The experimenter said the word, let the children try the segmentation, and provided feedback and modelling until at least two trial words were spontaneously segmented correctly. The children were encouraged to say the sounds as they marked them with the chips. The four training triads were presented once, then repeated once if necessary, (i.e., a possible total of 24 word segmentations as training), and any child who still could not segment spontaneously was given a score of zero. The test session procedure was the same as the trials except that there was no correction or feedback (i.e. no indication was given to the children as to whether they had segmented the words according to the experimenter's expectation or not). When children seemed to falter in the task or indicated a loss of interest, reminders were given of what they should be doing (i.e., tapping out sounds, although there was no retraining) and an indication was made as to how many words remained. At times the puppet was used to re-stimulate interest.

The procedure for the High School students was similar, although they were simply asked to state the number of units in each word and to say the word slowly, pausing to separate the units. The task was explained

hardest task first, since it was known from other studies that the phoneme counting would be more difficult than the syllable counting. Since performance on the sub-syllabic unit counting was of primary interest in this study, the risk of effects due to blocked presentation was taken. The results of a post-test with unblocked presentation justified this procedure, since there was no difference in the order of difficulty found (see also note 7).

as a counting task of different sound units, and the experimenter divided the first two or three example triads (as necessary) until the students could do the segmentations correctly on their own. For the high school students, the three lists were presented in one session in the order syllable, sub-syllabic, and phoneme. The students were encouraged to count out the units on their fingers.

Analyses. The number of chips selected was noted or, in the case of the High School students, the number given, as well as apparent boundary locations where possible. For the syllable and phoneme lists, the number correct was calculated according to traditional analyses of syllables and phonemes, with post-vocalic glides treated as part of the vowel (Derwing, 1973). The sub-syllabic unit list was scored three different ways in order to determine which analysis yielded numbers that related most closely to empirical unit counts. The three analyses were onset-rhyme ('rhyme'), onset-vowel-coda ('vowel'), and onset-nucleus-coda ('nucleus'). (The nucleus criterion differed from the vowel analysis in that it included post-vocalic /r/ and /n/ in the nucleus rather than as part of the coda; see Table 4.1.2 for counts that would be considered correct according to various criteria and see Appendix A for the complete word list.) Because of the onset-based criterion used for the training, criteria based on the structural models and the core+affix model were not used to score the sub-syllabic list, but individual comparisons were made between words with and without appendices and affixes. As part of the battery of tests for the children, I.Q. (Stanford-Binet; Terman & Merrill, 1973) and WRAT reading scores (reading subtest of Wide Range Achievement Test, Jastak *et al.*, 1978) were determined.

Results

All the children and High School students reached criterion on the training trials and went on to the test session. Because there was no significant difference in performance between males (n=27) and females (n=37) on any of the measures, the results were collapsed across sex. Overall, there was significantly better performance on syllable counting (70%) than on any other measure (the means for sub-syllabic counting with the three different analyses were rhyme=38%, nucleus=35%, and vowel=28%). There was significantly lower performance on phoneme counting (22%) than on all other measures except the score on sub-syllabic counting using the vowel criterion. All comparisons were paired sample *t*-tests using the Bonferroni $t(63)>3.35$ for k=7 tests, $p=.01$ (see Myers, 1979). The overall means and group means for percentage correct are given in Table 4.1.3.

Table 4.1.3

Mean percentage correct for each type of counting and analysis (n=16)

	Syllable	Sub-syllabic			Phoneme
Group:		Rhyme	Vowel	Nucleus	
Children	61	36	24	30	12
Kindergarten	58	35	14	21	7
Younger Grade One	61	35	25	29	12
Older Grade One	65	36	35	40	16
High School	96	45	40	51	51
Mean	**70**	**38**	**28**	**35**	**22**

To compare the different analyses of the sub-syllabic counting, sign tests (Woods, Fletcher, & Hughes, 1986) were carried out comparing the sub-syllabic scores under each criterion. Overall, both the rhyme and the nucleus criteria were better than the vowel criterion (Z=2.77, p=.01 and Z=5.40, p=.001, respectively) and not significantly different from each other. For the three older groups (younger Grade One, older Grade One, High School) the nucleus analysis scores were consistently higher than the vowel scores (p<.05), although the rhyme scores did not differ from the other two analyses. The Kindergarten children differed from the other groups in that their scores with the rhyme analysis were significantly better than with either the nucleus analysis or the vowel analysis, [illegible]e nucleus analysis scores were significantly better than with the vowel analysis (p<.002).

Partially repeated measures analyses of variance (ANOVA) were carried out separately for each of the three sub-syllabic analyses with the raw scores on each of the counting tasks (syllable, sub-syllabic, and phoneme) as a within subjects factor (3 levels), and with group (4 levels) as the between subjects factor.

Rhyme analysis for the sub-syllabic list

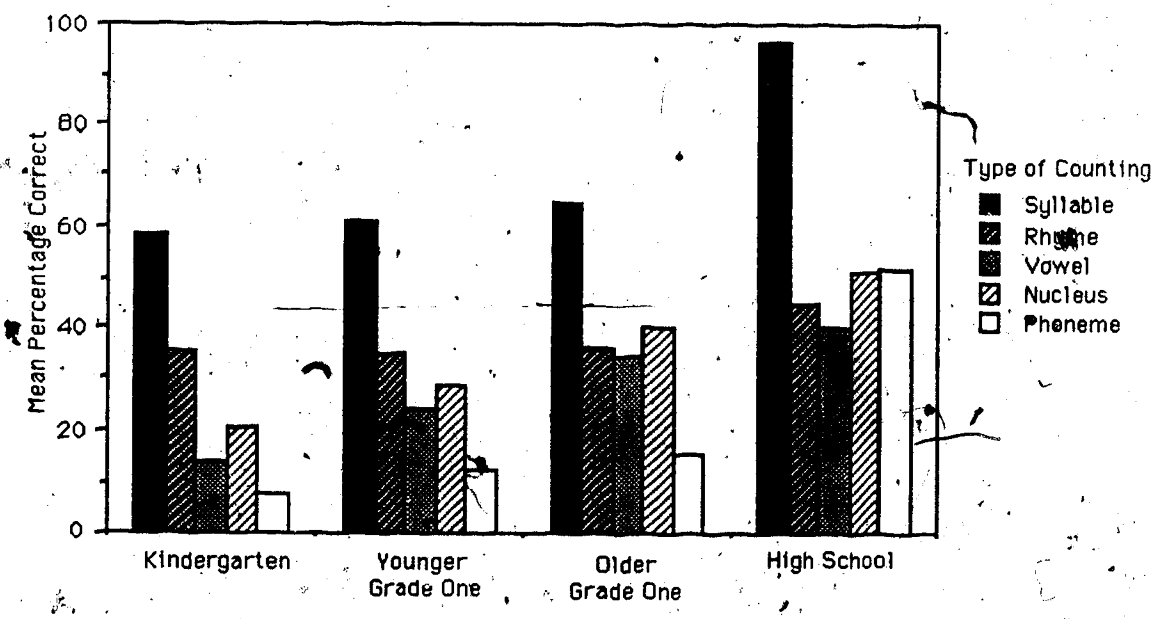
With the rhyme analysis for the sub-syllabic counting task, there was a significant subject group effect (F(3,60)=45.41, p<.0001), a significant effect of type of counting (F(2,120)=115.63, p<.0001), and a significant interaction (F(6,120)=3.29, p<.01). The scores on syllable counting were best overall, then sub-syllabic counting, and lowest was the phoneme counting (all t(15)>4.00, p<.0002).[5] The High School students did

[5]Tukey HSD tests were used for all the between subject comparisons in the analyses.

significantly better on the syllable and phoneme counting tasks than the other groups did (Tukey $t(60)>3.74$, $p<.05$), although there was no difference between groups on the sub-syllabic counting scores. This pattern held for all three groups of children, but for the High School students only the syllable counting score was significantly better than the other two (using Bonferroni's criterion $t(15)>4.28$ for k=12 tests and $p=.01$; see Figure 4.1.1).

Figure 4.1.1

Mean percentage correct for syllable, sub-syllabic (with the rhyme, vowel, and nucleus analyses), and phoneme counting (n=16 per group)



Vowel analysis for the sub-syllabic list

The ANOVA with the vowel analysis for the sub-syllabic counting also showed significant effects of subject group ($F(3,60)=28.59$, $p<.0001$), and type of counting ($F(2,120)=173.53$, $p<.0001$) and, at a lower level of significance, a significant interaction ($F(6,120)=3.11$, $p<.02$). The High School students did not differ from either Grade One group in their performance on sub-syllabic list counting using the vowel analysis but both the High School and older Grade One students performed significantly better than the Kindergarten children (Tukey $t(60)>3.74$, $p<.05$). For the Kindergarten, younger Grade One, and High School students, performance on syllable counting was significantly better than that on the sub-syllabic counting and phoneme counting, which were not significantly different from one another. (Although the phoneme counting score was higher than the sub-syllabic vowel score for the High School students and lower for the children's groups, these differences were not significant.) On the other hand, the older Grade One students' scores on the syllable and sub-syllabic vowel counting did not differ significantly, but both were higher than the phoneme scores (using the same Bonferroni $t(15)>4.28$ for k=12 tests and $p<.01$; see Figure 4.1.1).

Nucleus analysis for the sub-syllabic list

The nucleus analysis for sub-syllabic counting yielded yet another pattern for the four groups. The group and counting type effects were both significant ($F(3,60)=39.11$, $p<.0001$; $F(2,120)=170.78$, $p<.0001$, respectively), although the interaction was not ($F(6,120)=2.22$). Recall that the High School students performed significantly better than the other three groups on syllable and phoneme counting. They also performed better than the younger Grade One and Kindergarten children on sub-syllabic counting, using the nucleus analysis. The older Grade One children were also significantly better than the Kindergarten children

(Tukey $t(60)>3.74$, $p<.05$). The pattern across counting types differed for the four groups and from the other two sub-syllabic analyses discussed. The Kindergarten, younger Grade One and High School groups had significantly better scores on syllable counting than on sub-syllabic counting with the nucleus analysis, but the older Grade One children did not. Syllable counting scores were significantly higher than the phoneme counting scores for all four groups. The three young groups had significantly higher scores on the sub-syllabic counting than phoneme counting, but the High School students did not (although the difference was in the same direction; all Bonferroni $t(15)>4.28$ for k=12 tests and $p=.01$; see Figure 4.1.1).

Correlations

There were no significant correlations for any of the groups between performance on the syllable counting and any of the sub-syllabic analyses scores or phoneme counting scores. The correlations between performance with the vowel analysis for sub-syllabic counting and that of phoneme counting were significant for all four groups (i.e., $r(16)>.64$, $p<.01$). The scores with the nucleus analysis were significantly correlated with the phoneme scores for all four groups (i.e., $r(16)>.54$, $p<.05$), and the phoneme and rhyme analysis scores were negatively correlated for the two older groups (the older Grade One and High School students, $r(16)>-.62$; $p<.05$).

The correlation between traditional phoneme counts (see Appendix A) and the mean number of phonemes given for each word by the High School group was highly significant ($r=.95$, $p<.001$). Corresponding correlations were not calculated for the children groups because the mean phoneme counts were very close to one with very little variance.

For the children, both I.Q. and WRAT scores were obtained and correlations were calculated for a variety of measures. I.Q. was significantly correlated with WRAT scores only for the Kindergarten children (r=.68, p<.005) and with phoneme scores only for the younger Grade One children (r=.79, p<.001). Although WRAT scores were significantly correlated with the vowel and nucleus analyses and phoneme list collapsing across the three young groups (r=.56, r=.51, r=.38, respectively, p<.01), the only correlations with WRAT that were significant at the individual group level were for the younger Grade One children with vowel (r=.73,p<.005) and nucleus (r=.80, p<.001) analyses.

<u>Additional analyses</u>

To test the predictions of specific models, the mean numbers of units given in the sub-syllabic list for words that had different analyses were compared (using paired comparison t-tests and Bonferroni's criterion of $t(63)$> 4.59 for k=9 tests and p=.05). None of the words which could be analysed as having appendices (according to Selkirk's 1984 model) were treated differently from those words with the same canonical syllable shape which did not have syllable-final coronal obstruents (i.e., the means for words with appendices [in italics] and without them were spi*t*=2.02 *vs.* break=2.05; ar*t*=1.91 *vs.* ink=1.72; spen*d*=2.17 *vs.* drink=2.00; drif*t*=2.14 *vs.* grasp=2.30). Although the training trials required an onset-based analysis, the words that fit the definition for core+affix syllables were compared to core-only words. There were no significant differences. Finally, to test Vennemann's (in press) and Iverson & Wheeler's (1987) claims that syllables consist of heads and codas, the counts for syllables of the form VCC (mean=1.75) were compared to those for CCV syllables (mean=1.89). This difference was also not significant.

Discussion

The advantage of syllables over other types of counting was expected from previous work (e.g., Liberman *et al.*, 1974; Fox & Routh, 1975; Blachman, 1984) and was confirmed here. The mean score for syllable counting for the Kindergarten children was 58% and for the Grade One children 63% (compared to Blachman's 56% and 69%, respectively). That all groups could count syllables relatively well supports the notion that the syllable is a natural unit. Phoneme counting, on the other hand, was not very good. The children were tested at the beginning of the school year, before much reading instruction had occurred or ability had developed, so it is not surprising that the Grade One children were not significantly better than the Kindergarten children at counting phonemes. What is interesting is that the older Grade One children did not differ significantly from the younger Grade One children or the Kindergarten children. This argues against the notion that phonemic awareness is simply developmental, since there was a mean of 10 months' difference in age between the older and younger Grade One children and of 12 months' between the older Grade One and Kindergarten children. A year's difference should show some difference in ability, if such an ability is related to maturation, although it may have been below threshold for the present experiment.

The relatively poor performance of the High School students on the phoneme list was a surprising result, since phonemes have been assumed to be accessible to literate adults. It does not seem likely that the data were biased, since the high correlation of mean phoneme counts for each word and traditional phoneme counts for the High School group is in line with previous results (e.g., Dow, 1981; Derwing & Nearey, 1986). Other studies with adults (e.g., Perin, 1983) have found that people can count phonemes relatively well in short words but that accuracy decreases with the length of the word. The apparent contradiction between the high correlation and low score for phoneme counting in the present study may have been because long words were included (up to five phonemes) which were consistently underestimated. A sign test comparison between percentage correct for the high school students on short words (two- and three-phoneme; mean=64%) and long words (four- and five-phoneme; mean=43%) proved to be significant (Z=2.58, p=.007), which supports this underestimation explanation for the low phoneme scores.

The mean score for phoneme counting for the children was much lower than that reported by Blachman for her Grade One group (1984; 48%). However, she used only one, two, and three segment words for her phoneme counting task, whereas the majority of words in the present study were longer. The children here gave mostly one's, two's and three's as responses in the phoneme counting task. It may have been that the children were counting phoneme clusters or sub-syllabic units instead of phonemes. In fact, the nucleus (mean=31%) and rhyme criteria (mean=38%) proved to be significantly better predictors (using the sign test, Z>4.6, p<.001) of the children's responses on the phoneme counting list than the phoneme criterion (mean=12%).

Onset-rhyme

Of the three analyses for the sub-syllabic counting, the rhyme analysis (36%) accounted for the results best overall, especially for the three young groups (*versus* 30% for the nucleus analysis and 24% for the vowel analysis). This is perhaps not surprising, since rhyme is important in English song and poetry and children are 'taught' rhyme at home when they learn nursery rhymes. Recall, however, that all of the training words for sub-syllabic counting had open syllables, which would allow for either a rhyme analysis or an analysis dividing the rhyme into a vowel nucleus plus coda for the test words with codas. The rhyme score is similar to the syllable and phoneme scores in that there are no group differences for the young children, and it falls between the two in terms of the number correct. The syllable and rhyme scores are also similar in that neither are significantly correlated to WRAT or I.Q. scores, whereas some of the other scores are. Interestingly, the High School students did not differ from the children in performance using the rhyme criterion. This could simply mean that rhyme was used as the criterion whenever a sub-syllabic analysis was required (dependent on task demands).

Comparing the scores on the three types of counting with the rhyme analysis for sub-syllabic counting might lead to the suggestion that what is portrayed is simply the increasing difficulty of making finer distinctions. Syllables are largest and most salient, sub-syllabic units next, and phonemes smallest. This view ignores the fact that the categories can overlap, such that a syllable can be a single phoneme, or a single sub-syllabic unit (a rhyme). Thus, students might be simply counting syllables in every case. This would predict a maximum score of seven on the sub-syllabic counting task, because there are only seven items in which the number of syllables matched the number of sub-syllabic units and a maximum score of zero on the phoneme counting task, since no item had the same number of phonemes and syllables.

To test whether the scores for sub-syllabic counting could be explained by carry-over of the syllabic criterion, the responses to the sub-syllabic counting task were also scored with a syllabic criterion. The mean score for the Kindergarten children was higher than with any of the sub-syllabic analyses (51%), as was that for the younger Grade One children (38%), but the mean score for the older Grade One children was lower than on any of the other analyses (24%), as was that for the High School students (21%). There is thus some evidence that the syllabic criterion from the first task might have carried over to the sub-syllabic counting task, at least for the Kindergarten and younger Grade One children. The fact that none of the sub-syllabic scores or the phoneme score correlated significantly with the syllable score for any of the groups, though, suggests that performance on these tasks was fairly independent in general.

<u>Onset-vowel-coda</u>

The results are even more interesting when the scores for the vowel analysis of the sub-syllabic counting are examined. There is a clear difference between the groups here, with the most experienced/oldest (older Grade One and High School) groups' performance significantly higher than the least experienced/youngest (Kindergarten) group's. Although the means suggest a progression (Kindergarten=14%, younger Grade One=25%, older Grade One=35%, High School=40%), the differences between adjacent means were not significant. There was no difference between the sub-syllabic counting scores with the vowel analysis and the phoneme scores for the Kindergarten, younger Grade One and High School students, but for the older Grade One children the score with the vowel analysis was different from the score on phoneme counting, but not from syllable counting. There was some parallel patterning, that is, an increase in scores with age for the phoneme counting and vowel analysis, and this was reflected in the significant correlations between phoneme scores and vowel analysis scores for all four groups.

The suggestion that the pattern of scores for the vowel analysis reflects learning to make increasingly finer phonological distinctions is supported by the significant correlation between WRAT and vowel analysis scores for the younger Grade One group. One could argue that Kindergarten children could not discriminate below the syllable level (no difference in performance between vowel analysis and phoneme counting, but scores as high as those of the other two children groups on syllable counting); younger Grade One children were learning to make finer distinctions, possibly from learning about reading (high correlation of vowel analysis scores with WRAT); and older Grade One children were already making these distinctions (since the vowel analysis scores were not significantly different from syllable counting scores but were signficantly higher than the phoneme counting scores). However, the High School students' scores for the vowel analysis were not significantly different from their phoneme counting scores or from the Grade One children's scores. Both of these results conflict with the notion of increasingly finer discriminations with age.

Onset-nucleus-coda

An explanation of increasing discrimination with age is appropriate only if the nucleus analysis is used for the sub-syllabic counting. The nucleus analysis was consistently better than the vowel analysis as a predictor of scores on the sub-syllabic counting for all four groups. In addition, the high school students were better than the children at all three types of counting (syllable, sub-syllabic with the nucleus analysis, and phoneme) and the three young groups performed better at nucleus analysis than phonemic analysis (as did the High School group, although the difference between nucleus and phoneme scores was not significant for them). The older Grade One children's performance on the syllable list was not significantly different from the nucleus analysis score, but this apparent equivalent naturalness of syllable and nucleus analysis was not found in any other group (including the High School group, which was obviously performing at a very high level, since they achieved over 96% correct on the syllable list).

The younger Grade One children's scores on nucleus analysis for sub-syllabic counting were significantly correlated with reading level (as measured by the WRAT), suggesting some kind of discrimination improvement as a concomitant of learning to read (i.e., learning that words can be analysed into component letters may guide children to the realization that words have component sounds). The nucleus scores were not related to phoneme scores, so it is unlikely that it was merely learning about phoneme-sized units (letters) that was related to nucleus analysis or one would have expected parallel increases in phoneme counting ability. The increase across groups for the nucleus analysis contrasted with the rhyme analysis, which showed minimal variation across groups.

General discussion

Significant correlations have been reported between performance (percent correct) on phoneme and syllable counting (Treiman & Baron, 1981; Blachman, 1984) and between phoneme counting and I.Q. (Fox & Routh, 1975), but similar correlations were not significant in the present study. This might have been because the phoneme counting scores here were very low, with so little variation that the validity of the correlations is questionable. The children in the present study were simply unable to count phonemes. Although Treiman and Baron reported that WRAT scores correlated significantly with syllable counting for their Grade One group, no significant correlation was found here. It is not clear what to make of this discrepency, although the non-significance of the correlation in the present study was not unexpected, since syllables seemed to be accessible even to the youngest children (i.e., the Kindergarten group).

It would be worthwhile to do a post-test with these children one year later. Because of the unique population (same age, different school experience), some interesting predictions can be made at this point. If syllable and rhyme[6] are available to children before Kindergarten, then testing a year later should reveal no differences, even if that year included training in reading. If the ability to count sub-syllabic units such as vowel/nucleus and coda results from learning to make finer discriminations, one might expect a developmental trend, with a simple elevation of scores on sub-syllabic counting using the vowel or nucleus analysis because the children would be a year older; thus these scores should

[6] The rhyme analysis for the sub-syllabic counting reflects this.

approach the syllable counting scores (depressed from 100% because of boredom, carelessness, fatigue, etc.). The phoneme counting scores should therefore show the greatest difference among groups a year later because the Kindergarten children would have had no training in reading as yet (thought to be related to phonemic awareness), whereas the other two groups would have had a year's training by then. The present group of Kindergarten children should therefore obtain much lower scores on phoneme counting than the other two groups.[7]

Even without such a post-test, it is possible to evaluate the models discussed in the previous chapters in light of the present results. The claim that syllables are easily accessible units is certainly supported. Some types of sub-syllabic units also seemed to be fairly accessible, although it is not immediately apparent what kind of sub-syllabic analysis was being done. The very low scores on the phoneme list suggest at least that subjects did not perform a phonemic analysis. The rhyme analysis seems to be the best predictor overall (38%), but does not seem as good when one realizes that the (inappropriate) syllabic criterion for the sub-syllabic list yielded only a slightly lower overall mean (34%).

[7]The post-test phase of the study that Dr. F. Morrison carried out used slightly different methodology and had different subjects than in this study, so that no comparisons are possible. However, for the subjects he tested, the syllable scores a year later were not much different (pre=65% *vs* post=60%), the sub-syllabic scores (using the nucleus analysis) improved somewhat (pre=33% *vs* post=41%), and the phoneme scores improved dramatically (pre=11% *vs* post=38%). The post-test phoneme scores for the Grade One students were as high as the sub-syllabic scores (phoneme=45% *vs* sub-syllabic=42%), although the Kindergarten children's post-test scores for the phoneme list (24%) were lower, and roughly equal to the older Grade One's pre-test scores (17%). As far as they can be trusted, these results thus seem to support the predictions suggested above.

The significant group differences for the vowel and nucleus analyses for sub-syllabic counting (as opposed to the rhyme analysis, which did not differ among groups) and the trend to increasing scores with age and/or experience suggest that these analyses are worth examining in further detail. The vowel analysis here happens to make predictions identical to those of the metrical model suggested by Giegerich (1985), who posited an onset, rhyme, and extrametrical final consonant (not limited to Selkirk's [1984] narrower notion of an appendix as a syllable-final obstruent), yet was a poorer predictor of performance than either the rhyme or nucleus analyses. The performance on sub-syllabic counting using either the rhyme or nucleus analysis was not very different, which suggests that further work is necessary to discriminate between them. None of the comparison pairs designed to test the predictions of Selkirk's theory of extra-syllabic appendices proved to be significantly different. Due to the nature of the experiment (with training trials presented first), the sub-syllabic counting was biased to some kind of onset-oriented model. Thus, neither the core+affix model proposed by Fujimura (1975; 1976) nor the head and coda models proposed by Iverson and Wheeler (1987) and Vennemann (in press) could be adequately assessed here, although the tests made of specific cases showed no evidence of the use of either of these analyses. In the following sections some new tests are described which were intended to overcome some of these limitations of the unit counting procedure.

Experiment 2

Deletion-by-Analogy Task

The counting technique of the previous experiment, though prominent in the literature on phonemic awareness, proved to have only limited application to the case at hand, largely because of uncertainty as to what units the subjects were actually counting. It was assumed from subjects' correct responses that they were operating with the criteria presented in the examples, and their scores were taken as an indication of how consistently or well they could do this. A technique that more clearly shows what units are being accessed is the deletion task first reported in Bruce (1964) and later formalized by Rosner and Simon (1971) to predict reading difficulties. The test requires children to repeat a given word, then to say it again, omitting a specified phonological unit (e.g., deletion of /s/ from spin leaves pin).

In both of the above-mentioned studies, it was found that deletion of medial units, whether syllables or phonemes, was more difficult than deletion of non-medial units. Content, Kolinsky, Morais, and Bertelson (1986) found that young native French speakers could delete initial and final consonants from CVC syllables equally well (about 35% correct). The advantage of initial and final phonemes over medial phonemes found in the English studies (medial phonemes meant vowels or parts of consonant clusters in these studies) has also been noted for tip of the tongue phenomena (Browman, 1978; 1980) and slips of the tongue (Garnes & Bond, 1980; Shattuck-Hufnagel, 1983). Syllables are rarely mentioned in this context; however, the disadvantage for medial syllables in Rosner and

Simon's (R&S, 1971) test may have occurred because most of the medial syllables in their word list were unstressed reduced syllables (or perhaps not even present, as a word such as offering can be thought of as bi- rather than tri-syllabic). They found that deletion of a whole consonant cluster was easier than deletion of only part of a cluster, but there was only one example of this type in their test.

It was decided to use a task similar to R&S's to contrast the units discussed so far (e.g., onset, rhyme, nucleus, coda, head, appendix, affix) with incomplete units (such as parts of consonant clusters and parts of vowel nuclei), which have no consistent status in any of the models. A few changes were in order, however, before the R&S technique could be used here. First of all, the R&S list did not include enough instances of complex syllable structure to test these, so a new list had to be designed. Other changes were required to make the task easier to administer and carry out. In the original version, for example, the experimenter had to try to articulate isolated phonemes accurately (impossible in the case of stops), leaving the children to recognize the desired unit, somehow isolate it in the stimulus, and then pronounce the word with that unit left out.

A new technique, called deletion-by-analogy,[1] was therefore proposed that involved identifying the unit to be deleted more indirectly by analogy (e.g., flat:at=spin:*in*). Theoretical considerations then dictated the type of stimuli to employ. A model of syllables as strings of phonemes would predict no differences in difficulty based on position in the syllable. Deletion of clusters of phonemes, though, would presumably be considered more difficult than single deletions, since two or more phonemes would

[1]Thanks to Dr. B. L. Derwing for this suggestion.

have to be recognized, isolated, and deleted, instead of only one. A model of syllabic structure that posited intermediate units (such as onset, rhyme, nucleus and coda) would predict that complete unit deletions would be easier than incomplete unit deletions, regardless of the number of phonemes involved. Units comprised of both cluster and single phonemes were included in the stimulus list to discriminate among these possibilities.

If onset-rhyme is a syllabic division learned before Kindergarten, we would expect no difference between Kindergarten and Grade One children in their ability to delete onsets. If a sub-syllabic structure with codas becomes evident around the time of beginning Grade One (as suggested by Experiment 1), then the coda should be accessible to children in Grade One, but not to those in Kindergarten. Incomplete unit deletions should be difficult for preliterate children, as pilot studies with adults suggested that deletion of incomplete units was accomplished by using a spelling strategy. (People reported that they first imagined the spelling of the word and then imagined the sequence of letters with the requested deletion.)

The errors made in carrying out the deletions might also provide an indication of what was being manipulated. Suppose the task was performed by deleting the most appropriate unit accessible and producing what was left, (or, alternatively, just producing the most salient part of the stimulus word). If an onset-rhyme structure is the appropriate one to describe syllables, the rhyme portion of the word would most likely be produced as an error, since onsets are difficult to produce in isolation (and those with pre-vocalic stops are impossible to produce without, for example, a following schwa), but rhymes are word-like. If the sub-syllabic structure is onset-nucleus-coda, any of the three might be produced as errors (although, again, onsets and codas without vowels are difficult to produce). If none of these sub-syllabic units are particularly accessible, then the errors ought to have little relation to any of these categories and should be distributed in essentially random fashion among the various possible positions.

Method

Subjects. The same 16 Kindergarten children (7 girls, 9 boys, mean age of 5 years 8 months), 16 younger Grade One children (10 girls, 6 boys, mean age of 5 years, 10 months), and 16 older Grade One children (11 girls, 5 boys, mean age of 6 years, 8 months) who took part in Experiment 1 also performed the deletion task.[2] The study was explained to parents and permission was obtained before testing began. The children were all native speakers of English enrolled in the Edmonton Public School System. The 16 High School students from the American School in Kilchberg, Switzerland who took part in Experiment 1 also performed this task. All were native speakers of American (n=10), British (n=5), or Canadian English (n=1) between 16 and 18 years of age.

Materials. Ten sets of analogy pairs and four sets of trial pairs were designed for each of onset, coda, and incomplete unit deletions, such that the deletion of the target unit from the first word of each pair resulted in another real English word. Each analogy set consisted of a sample pair demonstrating the analogy and a stimulus word which the child was expected to change to the target word by deletion of the appropriate

[2]Thanks to MargoLee Horn for helping with the data collection.

element(s). There were five sets with single consonant deletions in each category and five sets with cluster deletions. Of the incomplete unit deletions, there were five deletions involving "external" segments (i.e., word-initial or -final) and five involving word-internal segments. The four trial sets had single and cluster deletions with the same phoneme(s) deleted in the example and target. All words (example and target words) were of a frequency of more than one occurrence in a million tokens (with standard frequency index (SFI) values of 40 or over from the Carroll *et al.*, 1971 frequency list). The onset, coda, and incomplete unit lists were prepared by randomly ordering the ten sets of analogy pairs (see Table 4.2.1).

Pictures were used to train the children on the task. There were pictures of a cowboy, a cow, a boy, a toothbrush, a tooth, and a brush. The puppet used in Experiment 1 was again available to entertain the children between lists.

Procedure. The children were tested individually in a quiet room in their school building. The first step of the procedure was a prescreen for the general ability to deal with linguistic analogies. The experimenter showed the cowboy picture, labelled it, then said, "If I say part of cowboy, I get cow (or boy)." The analogy pair consisted of the presentation of the toothbrush picture, the request for the child to name it, and then to choose between the pictures of a tooth and a brush to match the example pair. The training was repeated with various combinations of example and analogy pairs (e.g., toothbrush:brush::cowboy:[boy], cowboy:cow: :toothbrush:[tooth]). A minimum of two sets was presented and any children who were not successful in choosing the correct picture after four presentations were excluded from the experiment.

Table 4.2.1

Stimuli for the deletion-by-analogy task

Training Items:

Onset	Coda	Incomplete Unit
1. smile:aisle smart:(art)	1. makes:may rocks:(raw)*	1. sprang:rang spread:(red)
2. leg:egg love:(of)	2. beet:bee hate:(hay)	2. desk:deck task:(tack)
3. spoil:oil spout:(out)	3. seeds:see tides:(tie)	3. smell:sell smash:(sash)
4. cough:off cod:(odd)	4. choose:chew lies:(lie)	4. trend:tread bend:(bed)

Test Items (excluding analogy pairs):**

Onset	Coda	Incomplete Unit
Single Phoneme Deletions:		
1. mold:(old)	1. tease:(tea)	3. smack:(sack)
2. can:(an)	3. base:(bay)	6. shrug:(rug)
7. tar:(are)	5. grate:(gray)	7. hand:(had)
8. both:(oath)	6. juice:(Jew)	8. trail:(rail)
10. meat:(eat)	8. keep:(key)	9. please:(peas)
Cluster Deletions:		
3. flow:(owe)	2. loft:(law)	1. spring:(sing)
4. break:(ache)	4. toast:(toe)	2. sixth:(sick)
5. split:(it)	7. traced:(tray)	4. strange:(range)
6. spin:(in)	9. most:(mow)	5. splash:(lash)
9. trend:(end)	10. grouped:(grew)	10. strip:(sip)

*Same vowel in local dialect.
**Numbers refer to the order in which the pairs were presented.

The test session consisted of three lists of ten analogy sets, each preceded by four relevant training trials (see Table 4.2.1 for details), always presented in the order onset, coda, incomplete unit deletions.[3] Feedback

[3]The lists were presented in blocked order to prevent the young children from becoming discouraged early in the testing. It was known from other studies that onsets were accessible, whereas it was not known whether codas and incomplete units would be. The results of a replication with the High School students with completely randomized presentation showed no differences from the blocked presentation.

was provided for each of the training trials and, whether a child was successful or not, the test trials were then presented.

Each analogy word pair was presented to the child for comparison and then the stimulus word in the form of a game (e.g., "If I say part of flat, I get at. If you say part of spin, you get [in]"). If the child failed to respond to an item, it was repeated. If there was still no response from the child, a null response was recorded for the item. After five consecutive sets in a list with no response the list was discontinued (with null responses recorded for all non-presented items). A short break was taken after each list before the trial sets for the next deletion list were presented. Each list was preceded by the caution that it was a new game and the child would have to do something a little different. The puppet was sometimes used to explain the new game. The procedure for the High School students was essentially the same except that the cowboy and toothbrush examples were explained without using the pictures.

Analysis. The overall number correct for each of the onset, coda, and incomplete unit deletions was determined, with separate tabulations for the cluster and single phoneme deletions. Errors were coded either as a null response, an inappropriate deletion (i.e., deletion of something other than what was intended), a rhyme (which can also be thought of as an onset substitution), or as some other change (e.g., repetition of the word or the production of an unrelated word). As part of the battery of tests given to the children, I.Q. (Stanford-Binet; Terman & Merrill, 1973) and WRAT scores (reading subtest of Wide Range Achievement Test, Jastak *et al.*, 1978) were also determined.

Results

<u>Accuracy</u>

The proportion correct was calculated for each subject by deletion size and type. Because there were no significant differences between males and females on any of the lists or sub-lists, results were collapsed across sex.

A partially repeated measures analysis of variance (ANOVA) was carried out on the accuracy scores (proportion correct)[4] with group (4 levels) as the independent factor and type of deletion (onset, coda, incomplete unit) and size of deletion (single phoneme or cluster) as the repeated, within subject factors (see means, Table 4.2.2; the means for individual pairs can be found in Appendix C). The effect of group was significant ($F(3,60)=64.27$, $p<.0001$), because the High School group was significantly better than the three young groups, which did not differ significantly from each other on any of the dependent measures (all t-tests less than the critical Tukey HSD $t(60)=3.74$ for $p=.05$).[5] The type of deletion was also significant ($F(2,120)=44.28$, $p<.0001$) with onsets highest overall (54% correct), then codas (33%), then incomplete units (21%). Size of deletion was also significant ($F(1,60)=14.31$, $p<.0004$), with single deletions more often correct (39%) than cluster deletions (33%).

The interaction of type by group was significant ($F(6,120)=4.25$, $p<.002$), since the groups differed in their ability to perform the deletion types (see Figure 4.2.1). Kindergarten children did onset deletions

---

[4]Although proportional scores can have a bimodal distribution, a violation of the normality assumption of ANOVA, Myers (1979) suggests that the mean square ratio (and Type I error rate) is little affected.

[5]All between subject group comparisons were done using Tukey's HSD criterion.

significantly better than either coda or incomplete unit deletions, which were not significantly different from each other. Although that pattern also held for the younger Grade One children, the differences were not significant, and for the older Grade One children, the only significant difference was that the score for onset deletions was significantly higher than that for incomplete unit deletions. For the High School students, coda deletions were slightly better than onsets, although not significantly so, and both were significantly better than incomplete unit deletions (all significant differences were based on Bonferroni's $t(15)>4.27$ for k=12 tests and $p=.01$; see Myers, 1979).

Table 4.2.2

Mean percentage correct by group, deletion type, and size (n=16/group)

Deletion Type:	Onset			Coda			Incomplete Unit		
Deletion Size:	Cl	S	**Mean**	Cl	S	**Mean**	Cl	S	**Mean**
Group:									
Children	42	42	**42**	11	14	**12**	6	8	**7**
Kindergarten	43	36	**39**	3	1	**2**	5	6	**6**
Younger Grade One	36	43	**39**	18	24	**21**	5	8	**6**
Older Grade One	48	49	**48**	13	16	**14**	9	10	**9**
High School	78	100	**89**	93	96	**94**	46	80	**63**
Mean	**51**	**57**	**54**	**31**	**34**	**33**	**16**	**26**	**21**

Cl=Cluster deletion S=Single phoneme deletion

The interaction of size by group was also significant ($F(3,60)=8.47$, $p<.0001$), since only the High School group showed a significant advantage of single deletions (92% correct) over cluster deletions (72%; $t(15)=5.72$, $p<.0001$). The interaction of type by size by group was also significant ($F(3,120)=3.10$, $p<.05$, see Figure 4.2.2). For the Kindergarten group the only significant differences in the pairwise comparisons were between the cluster onset and both the cluster and single incomplete unit deletions and the cluster and single coda deletions. There were no significant differences for the younger Grade One children. For the older Grade One children only the differences between single onset deletions and single and cluster incomplete unit deletions were significant, and for the High School students the cluster incomplete unit deletions score was lower than all of the others and the single incomplete unit deletions score was lower than that of single onset deletions (all pairwise comparisons were significant using Bonferroni's $t(15)>4.41$ for $k=60$ tests and $p=.05$).

Figure 4.2.1

Percentage correct for each group by deletion type

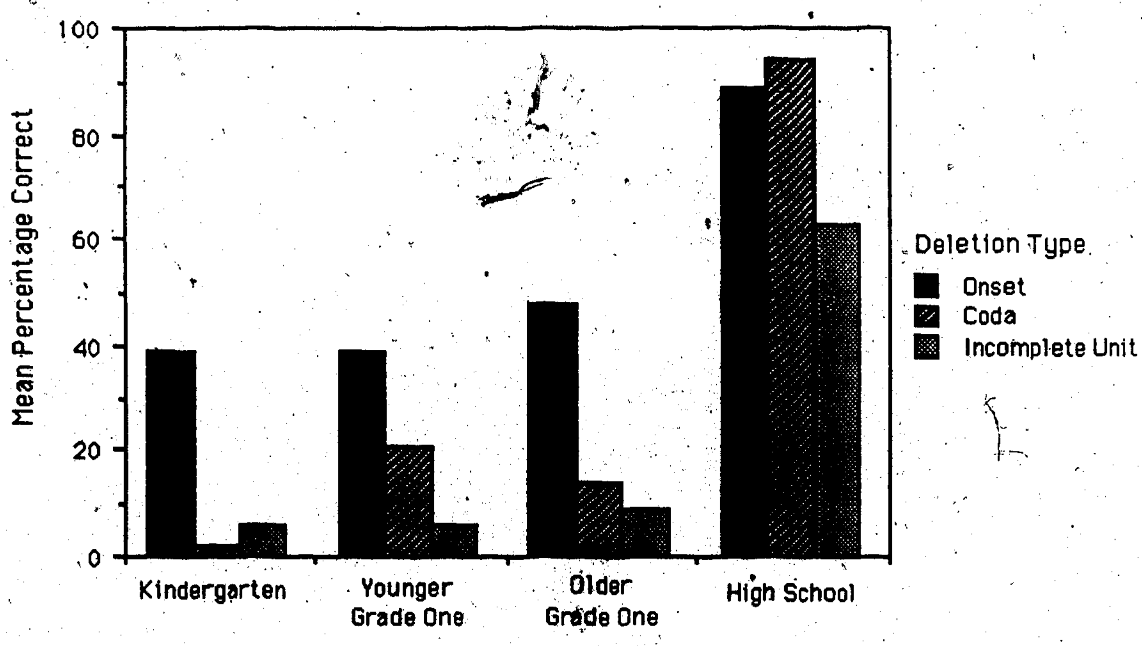
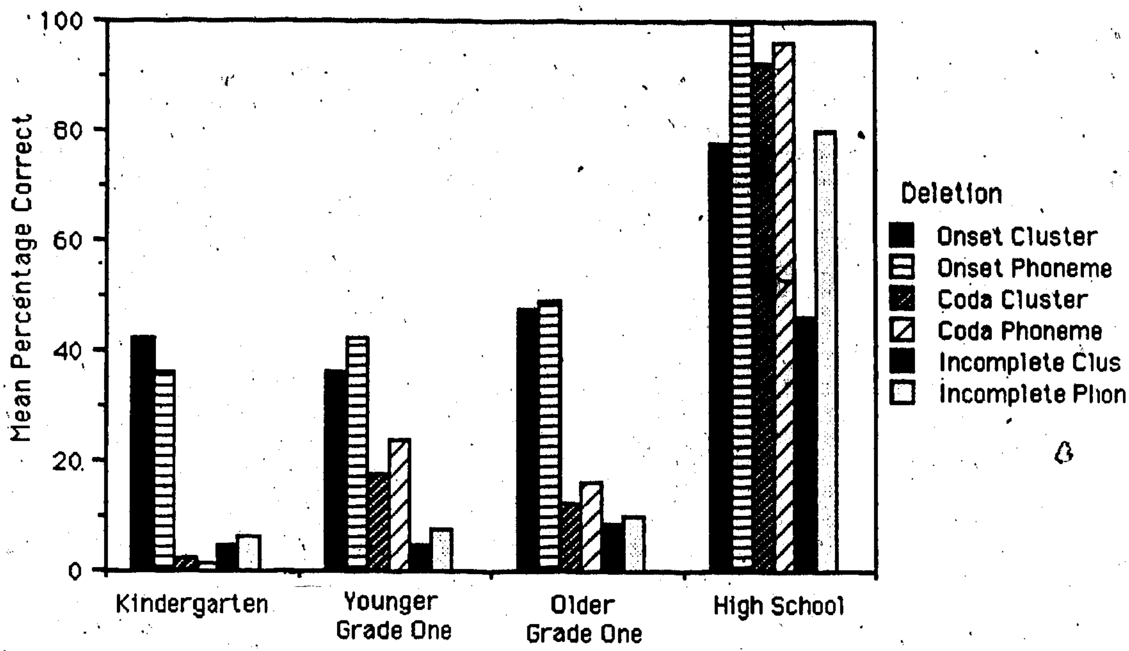


Figure 4.2.2

Percentage correct for each group by size and type of deletion



Correlations

To examine the possible relationship between I.Q. and WRAT scores and each type of deletion, correlations were calculated for the children by group and combined.[6] The overall correlations were significant between I.Q. and onset deletion scores ($r(48)=.336$) and between I.Q. and coda deletion scores ($r(48)=.340$), both at the .05 level. These were not significant for the Kindergarten children. For the younger Grade One children only the I.Q. and coda deletion scores were significantly correlated ($r(16)=.614$), and for the older Grade One children only the I.Q. and onset deletion scores were ($r(16)=.681$).

[6] The WRAT and Stanford-Binet were administered to the children as part of a battery of tests but were not administered to the High School students.

Errors

Incorrect responses were classified as one of four error types (see Table 4.2.3) and a number of planned comparisons were undertaken between different error types. For example, for an intended coda deletion (e.g., tease:tea) there might be no response (null response), there might be an inappropriate deletion (e.g., tease:ease), a rhyming word might be produced (e.g., tease:fees), or there might be some other change made (e.g., tease:dog). Of the total number of responses, significantly more were rhyme errors (26%) than null responses (10%, $t(63)=5.65$, $p<.0001$) and other changes (13%, $t(63)=3.85$, $p<.0003$), but not inappropriate deletions (15%). Comparisons were also made for each list. (All tests were paired *t*-tests using Bonferroni's criterion of $t(63)>3.39$ for $k=36$ tests and $p=.05$). For onset deletions, there were significantly more rhyme errors than null responses, inappropriate deletions, and other changes, and significantly more other changes than null responses. None of the errors for coda deletions were significantly different from one another, and for the incomplete unit list the only significant differences were that there were more inappropriate deletions and rhyme errors than null responses. There were significantly more null responses for coda deletions than either onset or incomplete unit deletions. There were significantly fewer inappropriate deletions with onset deletions than with coda deletions or incomplete unit deletions, and significantly fewer rhyme errors with the codas than with either onsets or incomplete units. No other comparisons were significant.

The inappropriate deletion category was actually a combination of onset and coda deletion errors for the children, of which the majority were onset deletions (14%) rather than coda deletions (1%). These were distributed between the coda and incomplete unit deletion lists. For the High School students, onset and coda deletion errors were quite rare (3% and 1%, respectively) but inappropriate partial onset and coda deletions were quite common (11%), especially for the incomplete unit deletion list. It was as though the students knew that a partial deletion was required for the incomplete unit deletion list but were unable to isolate the correct one.

Table 4.2.3

Mean percentage of each response type for each deletion type (n=64 subjects)

Error: Deletion Type:	Null Response	Inappropriate Deletion	Other Change	Rhyme	Correct
Onset	0	4	11	31	54
Coda	21	15	13	18	33
Incomplete Unit	8	26	16	29	21
Mean	**10**	**15**	**13**	**26**	**36**

Discussion

Although the children could perform the deletion task, it was certainly difficult for all of them. The High School students had no difficulty with the onset and coda deletions, but their performance with the incomplete unit deletions was significantly lower than with the other two types. That onset deletions were most often performed correctly overall fits with the notion of a readily accessible onset-rhyme structure. The hypothesis discussed in Experiment 1 that Kindergarten children can only do onset-rhyme analysis but Grade One children can do finer analysis predicts some differences in coda scores between these two groups. The Kindergarten children's scores for coda deletions were as low as for incomplete unit deletions, whereas both Grade One groups had higher coda deletion scores than incomplete unit scores (although these differences were not significant).

Criterion carry-over

A possible explanation for the unexpectedly low scores for the coda deletions for the Grade One groups is a carry-over of the onset task from the first list (since the three deletion lists were always presented in the same order of onset, coda, incomplete unit). There were significantly more inappropriate deletions with codas than with onsets (simply incorrect deletions). It could also be argued that there was a carry-over of both the onset and the coda criteria to the incomplete unit list, since the number of inappropriate deletions was highest for this list (although not significantly different from the coda list).

The problem with a criterion carry-over explanation for the coda list is that the most common error was a null response - the children simply did not produce a response. This happened significantly more often with the coda deletion list than with the other two lists. It seemed that children did not know what to do with this list, but their failure to respond suggests that they recognized that it was not simply onset deletions that were required. The task was designed so that making the intended deletion would result in an English word, but this was not necessarily true with an incorrect deletion. It may be that word awareness combined with failure avoidance contributed to the increased number of null responses for the coda list. In other words, if the children determined that the game was to produce part of the word and still have a word, but did not have access to the coda, then there would have been little possibility of success with the coda deletion list, so they said nothing. With the incomplete unit deletion list, however, it happened that a different deletion than the intended one often still resulted in a real word (e.g., the response for spring:sing (intended) was sometimes the inappropriate ring), which might account for why there were more inappropriate deletions and fewer null responses for the incomplete unit deletion list than for the coda deletion list.

The High School students could not do the incomplete unit deletions very well, though they performed almost perfectly on the onset and coda deletions. It did not seem that the problem was with the location of the deletions *per se* (i.e., word-initial or -final *vs.* word-internal segments) since the performance with word-initial incomplete deletions (77%) was lower than complete onsets (89%) and not much better than word-internal deletions (62%), and the performance with word-final incomplete unit deletions (19%) was much lower than with complete codas (94%).

One problem with the design of this task was that the deletions were presented in ordered blocks, as indicated above. This was done to minimize confusion in the young children, but it may have biased the results, because the incomplete unit list was always presented last. Thus, the low performance on the incomplete unit list may have been caused, at least in part, by lagging attention or carry-over of the onset and/or coda criteria. The high number of incorrect partial deletion errors on this list made by the High School students suggested that they realized that a new criterion was in effect, but the possibility still exists that order contributed to their poorer performance. Of more concern, however, is that the incomplete unit list included partial deletions of both the coda and the onset, so the students' attention had to shift from one part of the syllable to another with each analogy pair.

To determine whether order of presentation could have affected the results, an additional group of High School students from the same school at Kilchberg, Zürich (n=16) was tested. They were trained with the same example sets as in this study, but all examples (onset, coda, and incomplete unit deletions) were given before any test items, and the test pairs were presented in a different random order for each student. There were no significant differences between the blocked presentation and random presentation groups on any of the dependent measures (using independent *t*-tests and an alpha level of .05), which provides more support for the claim that the incomplete unit deletions were the most difficult because they were unnatural chunks of the syllable.

<u>Effect of size of deletion</u>

The significant difference between single and cluster deletions seemed to be due only to the High School group, which suggests that the difficulty with phonemic awareness for younger children results not from the size of the unit *per se* but from the role that unit plays in the syllable. The apparent disadvantage of phoneme-sized deletions found in other studies (e.g., Rosner & Simon, 1971; Bruce, 1964) may have resulted because these were sometimes also incomplete units (e.g., deletion of /n/ from <u>hand</u> disrupts either the complex nucleus or cluster coda, depending on the structural analysis, whereas deletion of /h/ should be easy, since it is a complete onset unit). These other studies cited medial deletions as more difficult than initial or final, but in all cases the deletions involved were smaller than a syllable; what they called medial deletions, therefore, would be treated here as incomplete unit deletions (parts of initial or final consonant clusters or the nucleus), which are predicted to be more difficult because they do not relate to any syllable structure constituent.

The question of why the High School students performed better with single phoneme deletions than cluster deletions is difficult to answer in light of the clear evidence for the psychological reality of sub-syllabic units. Unless it is argued that the High Students used a spelling strategy and the advantage of single phoneme deletions simply reflects a lower processing load relative to cluster deletions, this result provides support for the psychological reality of phonemes for literate speakers.

Onset-rhyme structure

As further support for the primacy of an onset-rhyme sub-syllabic analysis, the most frequent error type was production of a rhyme. These errors did not involve a simple deletion, since deletion errors resulting in a rhyme were classified as inappropriate deletions. It can be argued that production of a rhyming word involves a more difficult analysis than simple deletion, since the onset must first be isolated, then replaced with another to produce a new word. More likely, however, since rhyme errors were produced by the children as frequently as correct responses for the onset list, production of rhyme is familiar to the children and is the default strategy for a task requiring onset-rhyme analysis. Perhaps words are lexically accessible in rhyming sets, or perhaps manipulation of rhyme is learned so early that it becomes virtually automatic and is used whenever a task demands sub-syllabic unit manipulation (although sometimes children simply refused to do the task, as for the coda deletions).

General discussion

The significant overall correlation of I.Q. score with onset and coda deletion scores for the children may have been related to the difficulty of the task itself. Children with higher I.Q.'s were probably able to do the deletion-by-analogy task better simply because it was a new intellectual activity and they learned it more quickly. The overall correlations are not particularly interesting, however, since they were only on the order of $r=.34$, accounting for less than 12% of the variance. The individual group correlations isolated the relative contributions. The younger Grade One children's I.Q. scores were related to coda scores; whereas the older group's I.Q. scores were related to onset deletions. Even with these higher correlations, less than 40% of the variance was accounted for, so it is difficult to make strong claims about these relationships.

In terms of the models of syllable structure, the results of the present experiment support a major division or weak bond between onset and rhyme. Neither a secondary branch between the nucleus and the coda nor a bond between the onset and the vowel nucleus (Vennemann's 'body' bond) were supported by the results for the three groups of children (as demonstrated by the low scores for the coda deletion list), although both the onset and coda seemed equally accessible to the High School students. Even for preschool children, therefore, we have found support for the idea that a syllable does not consist of a mere string of phonemic segments, but contains constituents of a higher order, since single phonemes and clusters did not differ significantly for the children but onset and incomplete unit deletions did.

Testing these children after they learn to read and write would be interesting. Just as Rosner and Simon (1971) found that their test predicted reading difficulties, one would expect that children's performance on this task would improve after they become literate. If learning to read results in the ability to make finer discriminations, then codas should become more accessible and children should be able to perform much better on the coda deletion task. Performance on onset deletions should also improve, simply because of increased attention with age and intellectual maturity. The incomplete unit deletions should remain difficult, however, since they require breaking up natural constituents. Thus, the predicted pattern would be (in order of decreasing performance) onset, coda, incomplete unit deletions. There should be a group difference between Grade One children (both younger and older groups) and Kindergarten children, with the Kindergarten children still unable to do either coda or incomplete unit deletions very well.[7]

---

[7]The procedure and subjects in the post-test phase of Dr. Morrison's study were somewhat different than the present study (the lists were presented in different orders for each subject and there were different subjects), so reliable comparisons were not possible. His subjects tested a year later showed some improvement with the onset deletions (pre=48% *vs.* post=57%) and the incomplete unit deletions (pre=8% *vs.* post=12%) and, as predicted here, showed considerable improvement with the coda deletions (pre=12% *vs.* post=35%). The Kindergarten children were not as good as the two Grade One groups with coda deletions in the post-test (Kindergarten=19% *vs.* Grade One=43%), but were as good as the Grade One children had been on the pre-test (17%).

Experiment 3

Substitution-by-Analogy Task

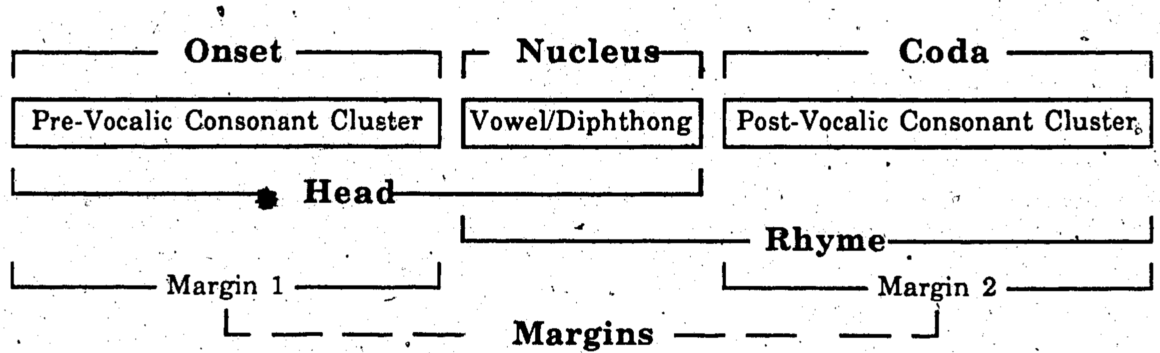
In order to claim that sub-syllabic units are psychologically real, it must be possible to demonstrate that the units discussed so far are accessible in a variety of manipulations. Examples of these are rhyme activities or language games like Pig Latin. Most adults can perfectly carry out requests to rhyme words or move word-initial consonant clusters to the end of words, so there is little point in asking them to do so. A manipulation that is not common in traditional language activities but which has been used successfully as a reading diagnostic is deletion of parts of words to form new words (Bruce, 1964; Rosner & Simon, 1971). However, adults seem to find a simple deletion task too easy to provide investigators with much information about syllable structure. In trial runs adults could do even the least likely deletions relatively quickly and well. It seemed that they relied on a spelling strategy, spelling the words to themselves, removing the letter(s) corresponding to the target deletion and then 'reading off' the remaining sequence of letters. The deletion-by-analogy task of the previous experiment (Experiment 2) may have minimized the use of this strategy since the focus of the task was directed to phonemes or letters. Nonetheless, it was still relatively easy for the high school students (who achieved an overall score of 82% correct). It is important to evaluate adults' abilities with a task that requires some cognitive processing in order to avoid such ceiling effects. Treiman (1983; 1985a; 1986) found that the onset and rhyme, and to a lesser extent, the coda emerged as units in novel word games. One difficulty with using her results to argue strongly for a particular model of syllable structure, however, is that the tasks and subjects varied across studies, which meant that the relative accessibility of various sub-syllabic units was not directly compared in a single study.

A substitution-by-analogy task was proposed as a manipulation that might prove taxing enough to prevent people from making the relevant changes by spelling out the words and which could be used to test all possible proposed sub-syllabic units without focussing attention on the phonological units under study (the focus, rather, became the identification and production of a pattern).[1] Subjects were presented with two pairs of words demonstrating the substitution of one sub-syllabic unit by another (e.g., substitution of any onset by /pl_/), and then were given a number of stimulus words in which to make the same substitution themselves. Pilot testing showed that this task was possible but not easy, could be set up to test all combinations of possible units, and was not transparent as to its purpose. The people tested in the pilot phase (mostly graduate students in linguistics) more often thought it was an intelligence test (similar to Miller Analogies, for instance) rather than a language task. On the assumption that the pre-vocalic consonant cluster, the vocalic nucleus (consisting of a vowel or diphthong), and the post-vocalic consonant cluster were reasonable units to manipulate, the six types of substitutions that were tested were onset, nucleus, coda, head (onset plus nucleus), rhyme (nucleus plus coda), and margins (onset plus coda; see Figure 4.3.1).

[1]The analogy task was suggested by Dr. B.L. Derwing as a way to assess the relative accessibility of phonological units. It was used here for substitutions and in Experiment 2 for deletions (which can be thought of as substitution by null).

Figure 4.3.1

Types of units included in the substitution-by-analogy task



The model of syllables as indivisible wholes would predict no differences in ease of manipulating various units, since none would be psychologically motivated (in fact, it would predict that the task was impossible). Alternatively, a model of the syllable as a string of equally weighted phonemes might predict that one segment ought to be just as easy to manipulate as another, but that groups of segments might present added difficulties in proportion to the number of segments involved. The models of syllable structure with sub-syllabic units would predict differences in the manipulability of the various units, depending on what the particular units are, where their boundaries occur, and how they are thought to be organized within the syllable. For example, the model of the syllable with a single intermediate level of equally weighted constituents would predict that there would be no differences in ease of manipulating the sub-syllabic units (although the question of interest would then become which units comprise the syllable; for instance, onset, nucleus, coda is one possibility).

The rhyme model proposed within metrical phonology (e.g., Liberman & Prince, 1977; McCarthy, 1979; Selkirk, 1980; 1984; Hayes, 1981) would predict that onset and rhyme would be most easily manipulted, since accessing the nucleus or coda would require going deeper into the structure. The head unit is not described by this model at all, since a major branch separates the onset and vowel; nor are the margins, which are not only discontiguous but separated by both a major and minor branch. Thus margins and head should be the least accessible units. In the case of a left-branching model, the major branch would be between the head and the coda, making them the most accessible. Within the head, the onset and the vowel nucleus would be the next most accessible, presumably both at the same level at the end of minor branches. This structure has no provision for rhyme, since a major branch separates the nucleus and coda, nor for margins, since there would be a major and minor branch between them, so both rhyme and margins would be predicted to be unnatural units.

The model proposed by Vennemann (in press) allows relevant predictions based on bonds and affinities. There are affinities among contiguous segments of the onsets and offsets (codas) and an onset bond that results in greater cohesion in the head (onset plus nucleus). It is not clear from Vennemann's description whether the onset bond would need to be broken to allow access to either the onset or the nucleus. If so, the head and coda (offset) would be most accessible, then onset, nucleus, and/ or margins (since there is nothing to predict that it would be harder to access both the onset and coda than just the onset once the onset bond is broken), and finally rhyme, since there is little to tie the nucleus and coda together.

A number of controls were built into the present study to ensure that any differences found among the unit substitutions were not caused by the number or relative saliency of the phonemes involved in the substitutions.

Method

Subjects. Forty native English speakers took part in the study (20 women and 20 men). All were university students or graduates and none had any specialized training in phonology beyond introductory courses.

Materials. Twenty sets of substitution analogies were devised. In each set there were two example pairs demonstrating the analogy, followed by four to six stimulus words. There were six different types of substitution: onset, nucleus (simple vowel or diphthong), coda, head (onset and nucleus), rhyme (nucleus and coda), margins (onset and coda). There were three sets each of onset, nucleus, head, and rhyme, and four of coda and margins.[2] Within each set, there were two spelling match targets, where the spelling of the response word (except, of course, for the substitution) matched the spelling of the stimulus word (e.g., *scum*-*plum* for a /pl_/ onset substitution), a mismatch target, where the spelling of the parts of the words not involved in the substitution did not match (e.g., *fought*-*plot*)[3], and a nonsense target, where the correct response was a nonsense word (e.g., stones-/plonz/).[4] The substitution types with an example of each, the example pairs (which were always spelling matches, for clarity), the stimulus words, and intended targets can be found in Table 4.3.1 (the complete list of stimuli used can be found in Appendix D).[5]

[2]The additional coda and margin sets were introduced to allow for direct comparison of phonemes in different positions.

[3]There is no /ɔ/ - /a/ contrast in the local dialect except before /r/, so these words rhyme.

[4]Due to an oversight one of the rhyme sets had an extra match target and no mismatch.

[5]All the stimulus words were real words except for one rhyme set, in which the first spelling match stimulus was inadvertantly the nonsense word pesk. See note 6 in the Results section.

Table 4.3.1

Substitution types with an example set of each (including example pairs, stimulus words, and targets)

Type	Example	Example Pairs	Stimulus:*target*	Type of Item
Onset	/str_/	cling-string	bride:*stride*	Match
		beam-stream	duck:*struck*	Match
			slate:*straight*	Mismatch
			coach:*/strotʃ/*	Nonsense
Nucleus	/_ɛ_/	braid-bread	steam:*stem*	Match
		dusk-desk	mint:*meant*	Match
			size:*says*	Mismatch
			look:*/lɛk/*	Nonsense
Coda	/_nts/	blimp-blintz	tempt:*tense/tents*	Match
		faith-faints	past:*pants*	Match
			hop:*haunts*	Mismatch
			love:*/lʌ nts/*	Nonsense
Head	/kræ_/	floss-crass	stink:*crank*	Match
		drift-craft	blush:*crash*	Match
			chic:*crack*	Mismatch
			snail:*/kræl/*	Nonsense
Rhyme	/_old/	baste-bold	monk:*mold*	Match
		strict-strolled	scant:*scold*	Match
			kept:*cold*	Mismatch
			chomp:*/tʃold/*	Nonsense
Margins	/st_k/	bathe-stake	glimpse:*stick*	Match
		plug-stuck	fast:*stack*	Match
			taught:*stock*	Mismatch
			dread:*/stɛk/*	Nonsense

The substitution targets were chosen so as to have at least one single and one cluster phoneme substitution per type. The stimulus words contained a variety of consonant clusters and vowels, with the constraint that the number of changes for onset and coda substitutions was equal (for each type of substitution, a total of 22 phonemes was substituted by a total of 24 phonemes), as was the number for head and rhyme substitutions (for each, a total of 34 phonemes substituted by 32 phonemes). Specific control items were also included to allow comparison of identical phonemes in different positions (/le_/:/_el/, /pl_/:/_lp/:/p_l/:/l_p/, /skæ_/:/_æks/, /sk_/ to /pl_/, /_sk/ to /_lp/, /k_s/ to /l_p/, /s_k/ to /p_l/; these items are marked with an asterisk in Appendix D). In addition to the 80 stimulus words of primary interest, there were extra items, bringing the total number of stimulus words (those requiring a response by the subject) to 92. The additional 12 items included seven cases of pre-vocalic glides, three cases of vowel nuclei which could be analysed either as syllabic /r/ or /ʌr/, and two other cases of post-vocalic /r/ (these items are marked with double asterisks in Appendix C).

Procedure. The subjects were tested individually in a quiet room. They were told they would playing a word game similar to Pig Latin and that they would be asked to make a new word or nonsense word by substituting sounds in the words given by the experimenter. The experimenter explained that she would say two pairs of example words that demonstrated the desired substitution analogy and then would give four to six stimulus words individually, pausing after each for the subject's answers. Subjects could have the example pairs repeated but were asked to keep such requests to a minimum. They were told that each substitution set was completely different from the others.

Sample sets were given before the session began, two to demonstrate what was meant by substitution (different sounds in cat can be substituted to get rat, coat or cap; while similar substitutions for dog give dog-fog, dogdig, dog-doll), and one to demonstrate analogy using deletion-by-analogy. In the latter case, two examples were given (trough-off, flow-owe) and then the subjects were asked to make the analogous deletions for smart, break, mold, and try). For the latter set, the instructions and example pairs were repeated if the subject did not get the first target (smart-art) correct. All subjects were able to do this sample set after one repetition of the instructions.

The 20 substitution sets were presented in a different random order for each subject (i.e., a subject might get an onset first, then a head, then a rhyme substitution, etc), but within each set the order was always two example pairs followed by two stimulus words that required spelling match responses, a word that required a spelling mismatch response and a word which required a nonsense word response. The rate of presentation was determined by the subject. There was a short pause after each set to make it clear to the subject that a new set was starting. The complete test generally took 20 to 30 minutes (range 10-45 minutes). The participants were asked for their comments afterwards and were provided with a brief explanation of the purpose of the study. The sessions were taped and subjects were told that this was in order to check the accuracy of the transcriptions. None were informed that reaction time was of interest.

Scoring. Responses were marked as correct (matching the target) or as one of seven types of error: incorrect substitution of onset, vowel, coda, head, rhyme, margins, or other change. For example, if the response after the example pairs might-plight; drank-plank (onset substitution of /pl_/) for scum was plum, it was considered correct; if the response was scuff, a coda substitution error was involved; if bum, an onset substitution error; if plank, an other change error. Reaction times were measured from the tape recordings with a stopwatch to the nearest tenth of a second, from the time of the completion of the utterance of the stimulus word by the experimenter to the beginning of the response by the subject.

Results

Analyses of correct responses

a) Analyses by type of substitution.

All of the subjects produced at least some correct responses (the range was 5 to 74 correct out of 80, mean=38.13). Because there was no significant difference in proportion correct between men (.50) and women (.46), all subsequent analyses were collapsed across sex. The proportion correct for each substitution type was calculated for each type of response (spelling match, spelling mismatch, nonsense word). In addition, since the spelling match response type included both the first and second stimulus word for each set (the order of presentation was always spelling match 1, spelling match 2, spelling mismatch, nonsense response), each of these was calculated separately. The mean proportion correct for each case can be found in Table 4.3.2 (the complete list of stimuli with proportion correct can be found in Appendix D).[6]

---

[6]The nonsense word pesk which was used inadvertantly in the first example of one of the rhyme sets probably caused the low scores in that set relative to the other rhyme sets. This did not effect the advantage for rhyme substitutions, however, since rhyme was the best overall.

Table 4.3.2

Mean proportion correct for substitution type by response type and position (n=40)

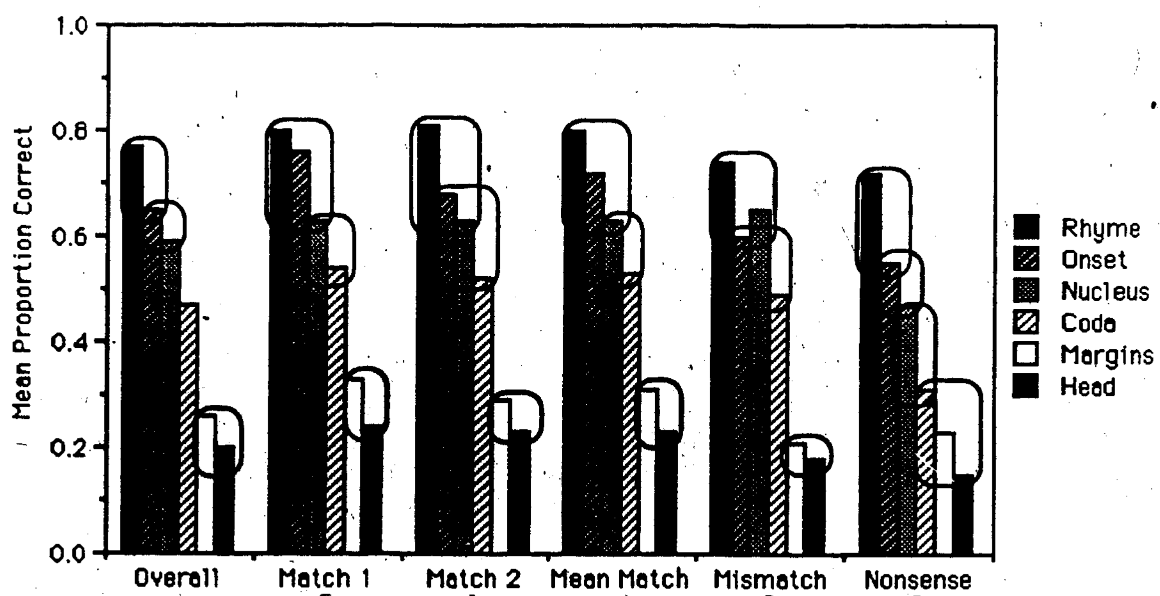
Pos	Substitution: Response:	Onset	Nucleus	Coda	Head	Rhyme	Margins
1	Spelling Match 1	.76	.63	.54	.24	.80	.33
2	Spelling Match 2	.68	.63	.52	.23	.81	.29
	Mean Match	.72	.63	.53	.23	.80	.31
3	Mismatch	.60	.65	.49	.18	.74	.21
4	Nonsense	.55	.47	.31	.15	.72	.23
	Overall	**.65**	**.59**	**.47**	**.20**	**.77**	**.26**

A repeated measures analysis of variance (ANOVA) with substitution type (6 types) and response type (spelling match, spelling mismatch, nonsense word response) as the within subject factors showed that both main effects were significant (substitution type $F(5,195)=48.40$, $p<.0001$; response type $F(2,78)=48.56$, $p<.0001$), as was the interaction ($F(10,390)=3.94$, $p<.0004$). Rhyme scores were significantly higher overall than all other types except onset substitution scores. Onset and nucleus substitution scores did not differ and both were significantly higher than coda, margins, and head substitution scores. Coda substitution scores were higher than the scores for margins and head substitutions, which did not differ significantly from each other (see Table 4.3.2; all pairwise comparisons between scores on different types of substitution were paired sample t-tests, using the Bonferroni criterion of $t(39)>3.18$ for k=15 tests and $p<.05$; Myers, 1979).

The same order of difficulty generally held for the individual response types, although there were more adjacent means which were not significantly different (because paired comparison tests were done both for the three response types and six substitution types; the statistic here was $t(39)>3.61$ for k=45 tests and $p<.05$). For the mean spelling match, rhyme, onset and nucleus substitutions, nucleus and coda substitutions, and head and margins substitutions did not differ from one another. For the mismatch response cases, rhyme, nucleus, and onset substitutions did not differ from one another, onset and coda substitutions did not differ, and again, margins and head substitutions did not. For the nonsense word response, rhyme and onset substitutions, onset and nucleus substitutions, nucleus and coda substitutions, and coda, margins, and head substitution score differences were not significant (see Figure 4.3.2).

Figure 4.3.2

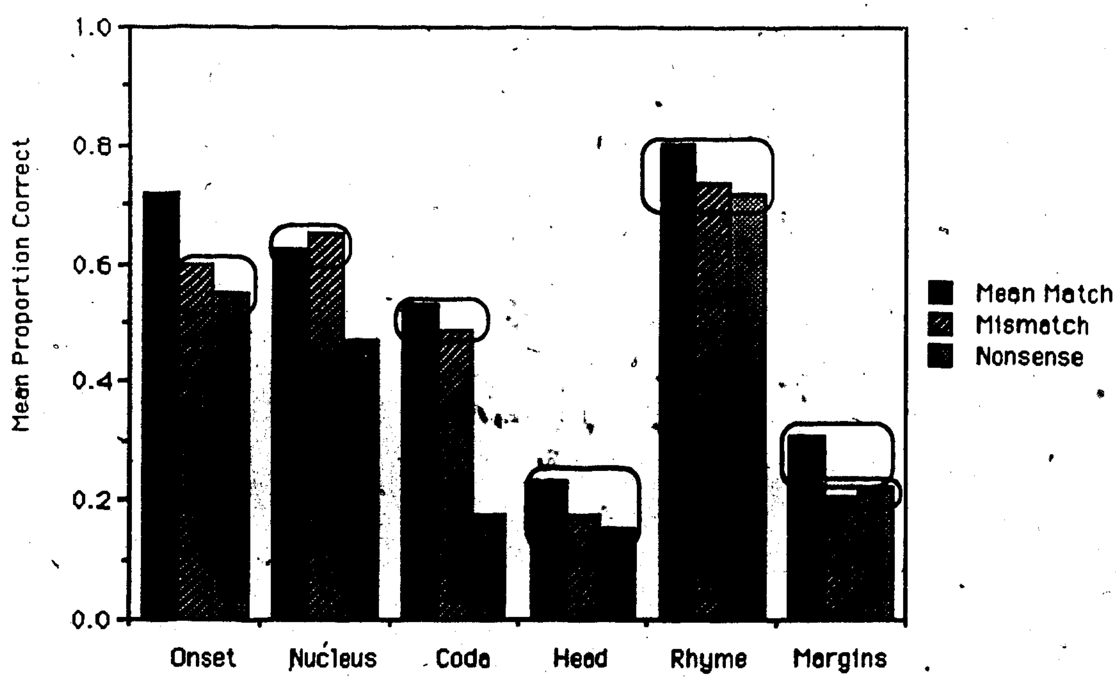
Mean proportion correct for each response type by substitution type (means within a circle are not significantly different $t(39)<3.61$)



The interaction arose because the advantage of mean spelling match over the other two response types did not hold for every substitution type (see Figure 4.3.3). Spelling match was significantly better than spelling mismatch only for the onset and margins substitutions, significantly better than nonsense word responses only for onset, nucleus and coda substitutions, and spelling mismatch was significantly better than nonsense word response only for nucleus and coda substitution types (all pairwise comparisons were Bonferroni t-tests for paired samples with $t(39)>3.70$, k= 63 tests and $p<.05$).

Figure 4.3.3

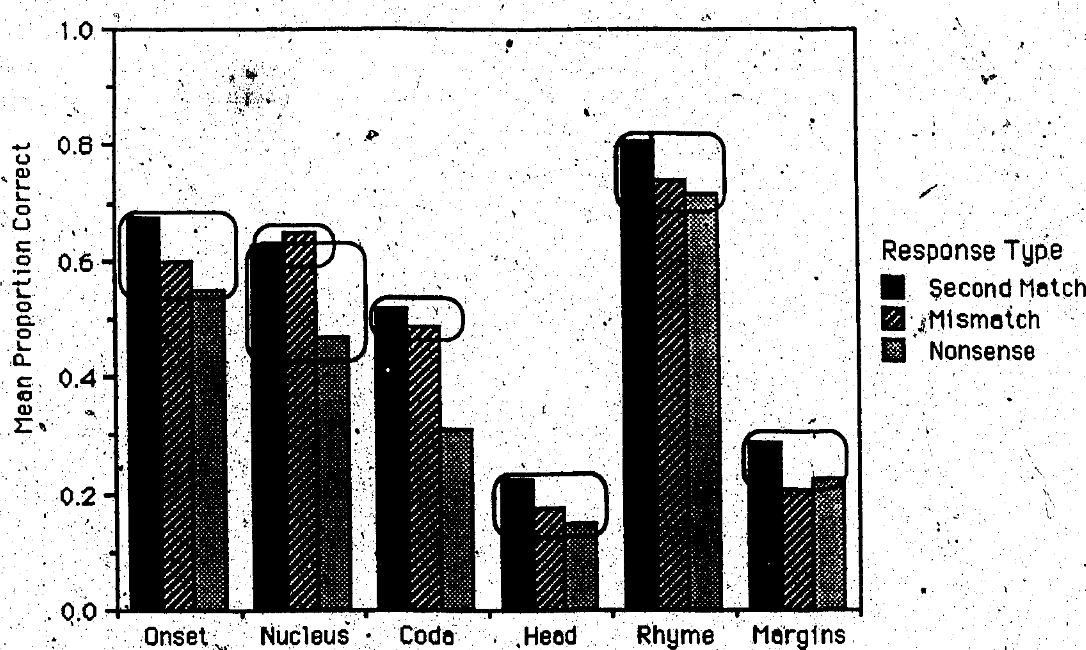
Proportion correct for mean spelling match, mismatch, and nonsense response types (means within a circle are not significantly different; $t(39)<3.70$).



To determine whether the slight advantage of mean spelling match over spelling mismatch might have been the result of a primacy effect (since the first and the second stimulus words within each set were both spelling matches) another repeated measures ANOVA (6x3) was carried out on proportion correct, with substitution type and response type as within subject factors. Response type in this case was with the second spelling match, spelling mismatch, and nonsense word, rather than with the mean spelling match. Again, there was a significant effect of substitution type ($F(5,195)=46.63$, $p<.0001$) and of response type ($F(2,78)=34.80$, $p<.0001$), and a significant interaction ($F(10,390)=3.43$, $p<.002$), but there was no advantage of spelling match over spelling mismatch. The only significant differences (with $t(39)>3.70$, k=63 tests and $p<.05$) were between spelling mismatch and nonsense word responses for nucleus substitutions and between both spelling match and mismatch and nonsense word responses for coda substitutions (see Figure 4.3.4). The only difference between the second spelling match and the mean spelling match analyses was that onset was not significantly different from coda in the second spelling match analysis but it was in the mean match analysis (see Figure 4.3.2).

Figure 4.3.4

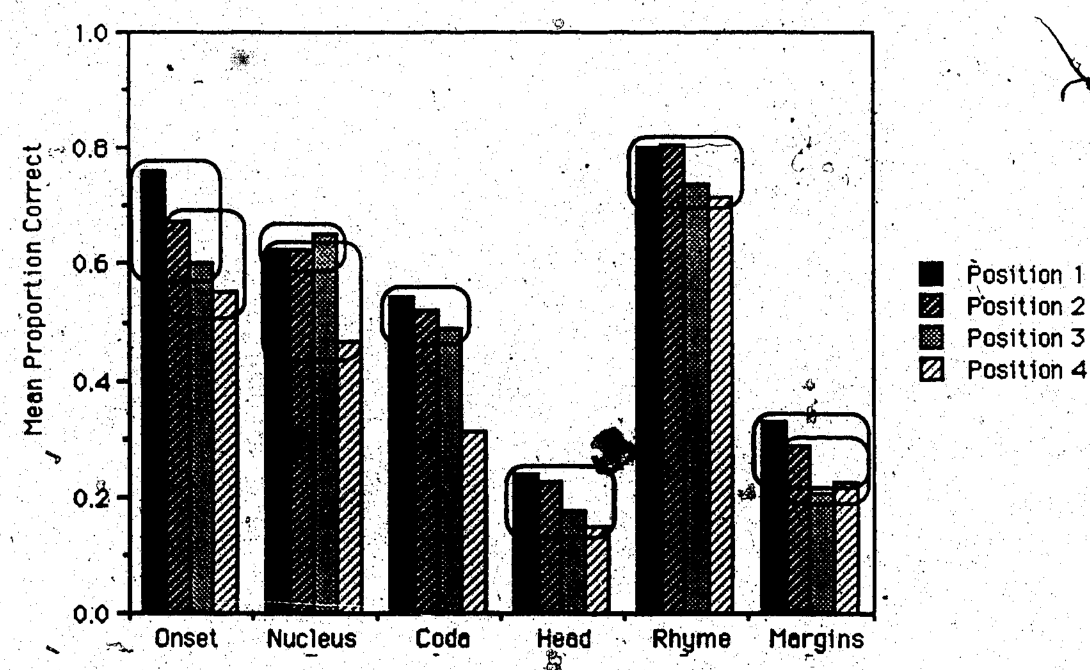
Proportion correct for second spelling match, mismatch, and nonsense response types (means within a circle are not significantly different; $t(39)<3.70$).



The presentation order for different response types was invariant, so just as the advantage of mean spelling match over spelling mismatch might be thought of as a primacy effect, so might the disadvantage for nonsense word responses be considered an ordering effect. Therefore, another repeated measures ANOVA was carried out with substitution type (6 types) and order of presentation (4 positions) as the within subject factors. Both main factors were significant (substitution type $F(5,195)=49.37, p<.0001$; order $F(3,117)=39.06, p<.0001$), as was the interaction ($F(15,385)=2.64, p<.007$). Despite the significant order effect and the appearance (see Figure 4.3.5) of a decrease in proportion correct with order, the only significant differences ($t(39)>3.90$ for k=96 tests, $p<.05$) were between position 4 and position 1 for onset substitutions, position 4 and position 3 for nucleus substitutions, position 4 and the other three positions for coda substitutions, and positions 3 and 1 for margins substitutions. The order of difficulty for the substitution types was generally the same for positions 1 and 2, except that onset substitution scores were significantly higher than coda substitution scores in position 1 but not in position 2 (see Figure 4.3.2; match1=pos1, match2=pos2, mismatch=pos3, nonsense=pos4).

Figure 4.3.5

Proportion correct for substitution type by response order (means within a circle are not significantly different; $t(39)<3.90$)



To determine whether the disadvantage of the nonsense/position 4 response was due to type or order, a *post hoc* comparison was made of the coda substitution types.[7] There were two coda cases where the mismatch response was in position 3 and the nonsense response in position 4, and two other cases where the mismatch response was in position 4 and the nonsense response in position 5. If there was a response effect, the proportion correct for nonsense should have been lower than mismatch, regardless of the position. Alternatively, if there was an order effect, the proportion correct for position 4 should have been equal, regardless of whether position 4 was a mismatch or a nonsense response. The means for the mismatch cases (pos3=.49, pos4=.49) were equal and the nonsense cases (pos4=.33, pos 5=.30) were not significantly different. However, the test of whether the two position cases were different proved to be significant ($t(1)=8.89$, $p<.05$), since the mean proportion correct for pos4/mismatch (.49) was greater than pos4/nonsense (.33), supporting the claim that the effect was due to response type rather than order of presentation.

b) Analyses by size of substitution.

It might be argued that one type of substitution is more difficult than another because proportionately more of the syllable has to be substituted for some sub-syllabic units than for others. To check for evidence for this size argument, the proportion correct was determined for each subject by size of substitution unit, where onset, nucleus, and coda substitutions were considered to be single unit substitutions and head, rhyme, and margins substitutions to be double unit substitutions (since each of these units can

[7]This comparison was possible because some of the 12 extra items included to study the question of glides in vowel nuclei occurred before the mismatch targets.

be described as containing two of the single units; i.e., head=onset+nucleus, rhyme=nucleus+coda, and margins=onset+coda).[8] A repeated measures ANOVA (3x2) for proportion correct with mean spelling match, spelling mismatch, and nonsense word responses *versus* unit size showed single unit substitutions were correct significantly more often than double unit substitutions ($F(1,39)=33.06, p<.0001$). Response type was significant ($F(2,78)=61.29, p<.0001$), as was the interaction ($F(2,78)=19.13, p<.0001$, see Table 4.3.3). The difference between single and double unit substitutions was significant for mean spelling match ($t(39)=5.55$) and spelling mismatch ($t(39)=7.02$), but not for nonsense word responses.[9] Although proportion correct was greatest overall for mean spelling match, the difference between mean spelling match and spelling mismatch was significant only for double unit substitutions ($t(39)=3.30$, $p<.05$). Nonsense word response scores were significantly lower than mean spelling match scores for both single ($t(39)=8.49$) and double unit substitutions ($t(39)=5.47$), as well as lower than spelling mismatch scores for single unit substitutions ($t(39)=6.05$).

---

[8]Note that size here does not refer to the number of phonemes involved in substitutions *per se* (that is, single segments *vs* clusters), since a head substitution of /le_/ involves two phonemes, just as the onset substitution /pl_/ does.

[9]Again, Bonferroni t-tests were used throughout these paired comparisons of means (for k=6, $t(39)>3,46$ and $p=.01$).

Table 4.3.3

Mean proportion correct for unit size and response type (n=40)

Substitution Size:		Single Unit	Double Unit
Response Type:			
Pos 1	Spelling Match 1	.63	.45
Pos 2	Spelling Match 2	.60	.46
	Mean Spelling Match	.62	.45
Pos 3	Spelling Mismatch	.57	.31
Pos 4	Nonsense	.43	.35
	Overall	**.56**	**.39**

c) Analyses of special cases

The phonemes /p/ and /l/ were included as substitution types in a variety of positions (onset, coda, and margins) to compare these substitution types directly. There was no difference between substitutions with these phonemes as onset (/pl_/ mean correct=72%) and coda (/_lp/ 63%), or between different orderings as margins (/p_l/ 33%; /l_p/ 29%), although all comparisons of either onset or coda with either of the margins cases were significant ($t(39)>4.31$, $p<.0001$). Comparisons were also made controlling for what was substituted. In one case of each of the onset, coda, and margins substitutions involving /p/ and /l/, the phonemes /s/ and /k/ were the stimulus phonemes. This pattern was a subcase of the overall proportion correct: the order was onset (/sk_/ to /pl_/, 83% correct), coda (/_sk/ to /_lp/, 60%), margins substitutions (/s_k/ to /p_l/, 48%; /k_s/ to /l_p/, 38% correct).

Another set of comparisons involved the phonemes /e/ and /l/, as head and rhyme. Substituting these phonemes as a rhyme was significantly more often correct than as a head (/_el/ 89%; /le_/ 20%; $t(39)=12.18$, $p<.0001$). A fairly direct comparison that was part of this set involved the phonemes /s/, /k/, and /æ/.[10] The rhyme substitution (/_æks/ to /_el/) had a much higher proportion correct (85%) than the head substitution (/skæ_/ to /le_/; 15%). One comparison which was less direct, but did involve the same nucleus, was the rhyme substitution /ruf/ to /rel/ (95% correct) *versus* the head substitution of /suð/ to /leð/ (25% correct).

There were 12 extra stimulus items containing pre-vocalic glides and post-vocalic /r/. The pre-vocalic glides were treated as part of the onset significantly more often (77%) than as part of the vowel nucleus (13%; $t(6)=4.43$, $p<.005$) irrespective of whether the glide was represented orthographically by a consonant or a vowel (e.g., swift *vs.* quest). The post-vocalic /r/ was treated as part of the vowel nucleus significantly more often (47%) than as part of the coda (16%, $t(4)=2.65$, $p<.05$). The difference for the /ʌr/ cases alone was even larger (nucleus 53% *versus* coda 3%), $t(2)=12.75$, $p<.005$).[11]

d). Reaction Time Analysis

Only reaction times (RT) for correct responses were considered because it is difficult to attach any meaningfulness to the length of time it takes to do something incorrectly. The pool of responses for this analysis was thereby limited and was further restricted by the decision to use only

[10]Unfortunately, there was no direct comparison pair included in the list (such as tail and late. However, the effect is so large for the close comparison that it probably does not matter.

[11]Some of the responses for these cases could not be classified as part of the margin or the nucleus, so the sum of the proportion of nucleus and coda responses did not equal one.

those subjects with at least two correct responses per substitution type. This resulted in a sample size of 16 (nine men and seven women). A repeated measures ANOVA (2x6) with sex as the between subjects and substitution type as the within subject factor showed a significant effect only of substitution type ($F(5,70)=4.93$, $p<.02$). Collapsing across sex increased the significance of the type effect ($F(5,75)=8.95$, $p<.002$). Paired comparisons of mean reaction time for each substitution type (for k=15 tests, using Bonferroni's $t(15)>3.48$ with $p<.05$) showed that rhyme substitutions (1.8 sec.) did not differ significantly from onset substitutions (2.1 sec.) but were faster than nucleus (2.6 sec.; $t(15)=3.85$), coda (3.1 sec.; $t(15)=4.24$), and head substitutions (4.6 sec.; $t(15)=5.23$). None of the substitution types differed significantly from margins substitutions (4.1 sec.), probably because of the relatively greater variance associated with the reaction times for margins substitutions (more than twice that of the next highest and 50 times that of the lowest RT). Onset substitutions were significantly faster than head substitutions ($t(15)=4.56$), as were nucleus ($t(15)=5.05$), and coda substitutions ($t(15)=4.55$).

A paired samples test of mean reaction time for different sized units showed that single unit substitutions were significantly faster than double unit substitutions (2.6 sec. *versus* 3.5 sec., respectively; $t(15)=3.79$, $p<.002$), which parallels the result of the accuracy measure.

Error analyses

Errors were coded in terms of seven possible error types (inappropriate substitution of onset, nucleus, coda, head, rhyme, margins, and other change) for each intended substitution. In three cases (for onset, coda, and rhyme substitutions) the entry on the diagonal was close to half of the total number of errors for that type (see Table 4.3.4). This is exactly what one might predict, if each substitution is considered to have two parts (*where* the substitution was and *what* it was). For example, onset errors for onset substitutions were the result of only the *what* being wrong, not the *where*, and the other errors were distributed across the rest of the substitution type possibilities (although it is by no means clear that these two types of errors are equally probable, and there are various other predictions which could have been made). In contrast, the entry on the diagonal for each of the head and margins substitutions was low relative to total errors for that type; most of the errors for these two targets were onset substitutions.

Table 4.3.4

Matrix of number of errors per substitution type (n=40 subjects; entries on the diagonal are underlined).

Error: Target:	Onset	Nucleus	Coda	Head	Rhyme	Margins	Other	Total
Onset	131	11	0	4	11	11	2	170
Nucleus	10	49	4	8	119	0	6	196
Coda	4	6	154	0	147	10	21	342
Head	190	43	13	29	50	29	31	385
Rhyme	10	29	9	0	49	0	13	110
Margins	199	25	48	11	59	81	49	472
TOTAL	**544**	**163**	**228**	**52**	**435**	**131**	**122**	**1675**
Mean	13.6	4.1	5.7	1.3	10.9	3.3	3.1	

The mean number of onset errors (13.6) was not significantly different from the number of rhyme errors (10.9), but was significantly higher than the coda (5.7; $t(39)=5.81$), nucleus (4.1, $t(39)=6.79$), margins (3.3,

$t(39)=7.08$), other change (3.1, $t(39)=7.03$), and head errors (1.3, $t(39)=8.54$). There were significantly more rhyme errors than nucleus, margins, other change, and head errors (all $t(39)>5.43$), but not coda errors; and both coda and other change errors were significantly more frequent than head errors ($t(39)=5.55$, $t(39)=4.09$, respectively).[12]

Overall pattern

The order of all three measures (proportion correct, reaction time, and error preference) did not differ much (see Table 4.3.5). For each, onset and rhyme substitutions were not significantly different from each other and were the highest (proportion correct), fastest (reaction time), and preferred (error), and margins and head substitutions were the lowest ranked and were also not significantly different from each other. Nucleus and coda substitutions fell in the middle, also not distinct from each other.

[12] All significant differences (i.e., $t(39)>3.98$) were set with Bonferroni's criterion for k=21 comparisons and significance at the .01 alpha level overall.

Table 4.3.5

Rank order for each measure by substitution type (n=40 for accuracy and errors, n=16 for reaction time).

Accuracy (Proportion Correct)		Reaction Time (Seconds)		Preferred Errors (Number)	
Rhyme	.77	Rhyme	1.8	Onset	13.6
Onset	.65	Onset	2.1	Rhyme	10.9
Nucleus	.59	Nucleus	2.6	Coda	5.7
Coda	.47	Coda	3.1	Nucleus	4.1
Margins	.26	Margins	4.1	Margins	3.3
Head	.20	Head	4.6	Other	3.1
				Head	1.3

Discussion

The results clearly support a model of syllable structure with a major break between the onset and rhyme and a minor one between the nucleus and coda. The performance on onset and rhyme substitutions was not significantly different either in proportion correct or reaction time. Although there was some overlap in performance on nucleus and coda substitutions, these two were generally lower than onset and rhyme substitutions, with nucleus better than coda substitutions. Since the vowel is generally thought of as the syllabic peak, carries stress, and can stand alone, it is not surprising that the nucleus was more salient than the coda. This result suggests that a simple tree structure for the syllable with equal end nodes would have to be adjusted to give more weight to the vowel. A right-branching structure for syllables predicts a disadvantage for head and margins substitutions, which was supported. The discontinuous margins were slightly easier to access than the head, perhaps because of the saliency of the vowel (i.e., substituting margins may also be thought of as substituting the stimulus nucleus into a target shell), whereas the head is not a predicted unit in this model and the leftover (i.e., the coda) is not as salient as other units (i.e., the onset and the rhyme).

Despite controls on the number of phonemes involved in substitutions, head substitutions were significantly more difficult than rhyme substitutions. Thus, no support was found for the head unit of the left-branching hierarchical model proposed by Iverson & Wheeler (1987), nor for that of Vennemann's (in press) description of bonds and affinities. Without the special body bond proposed to bind the onset and nucleus, Vennemann's model might be a useful alternative to a hierarchical model. It may be that bonds are language dependent. Based on the present results, a rhyme bond (between the nucleus and coda) would be more appropriate than a body bond to describe English.

A non-hierarchical model with a rhyme bond would presumably make the same predictions as a right-branching structure, except that differential affinities could account for why the nucleus appeared to be more accessible than the coda. If the coda is weakly bonded within itself (more weakly than the rhyme bond) and there is a segmental bond for the nucleus (to bind the vowel with off-glides), then one would predict that a CC coda would be less tightly bound than a vowel diphthong nucleus. There should be no difference between simple vowel nuclei and single segment codas, because there would be no affinities to play a role beyond the rhyme bond. An inspection of the data revealed that although the diphthongs were more often correct (53%) than coda clusters (45%), the test case of simple vowel and simple coda were not equal (71% and 50%, respectively). However, the vowel diphthongs were approximately equal to the simple coda. The preceding means were based, in some cases, on only one set of substitutions, however, and are therefore inconclusive. Further elaboration and testing of these affinities should be carried out (see Chapter 5 for a discussion).

The other models of syllable structure that can be rejected based on the results of this task are the 'flat' model of the syllable as a string of equally weighted phonemes and the model of the syllable as a single indivisible unit. Despite controls on the numbers of phonemes involved in substitutions (e.g., onset *vs.* coda substitutions, rhyme *vs.* head substitutions), there were significant differences between substitution scores with these sub-syllabic units. If syllables are treated as non-analysable wholes by native speakers, then one would predict no differences in the ability to access and to make substitutions with different sub-syllabic units (i.e., all would present major difficulties). Neither the core plus affix model nor the metrical model including appendices was tested here.

The question of whether people were relying on a spelling strategy to perform this task or, more generally, whether spelling plays a role in phonological manipulation, was not completely resolved in this study. There was no difference between spelling matches and mismatches with the substitution type analysis (i.e., with the six different types of substitutions analysed separately). However, a comparison of single unit substitutions (i.e., onset, nucleus, coda) with double unit substitutions (i.e., head, rhyme, margins) showed a significant advantage for spelling match targets for the double units over mismatch targets (although there was no difference for the single unit substitutions). There was little in common phonologically or orthographically between the stimulus and target for double unit substitutions with a spelling mismatch, since the only unit that did not change phonologically (i.e., which was not part of the substitution) was different orthographically. Under these optimal conditions the effect of orthography did emerge, with the spelling mismatch targets significantly more difficult than the match targets. This result serves as yet another caution that one cannot ignore the influence of awareness of spelling in phonological tasks, since spelling that matched obviously made the task easier.

Although at first glance the general pattern for the substitution type by order analysis appeared to indicate decreasing performance with distance from example pairs (i.e., the results for onset and head, Figure 4.3.5), the relationship was not strictly linear, because there was a large difference in performance between nonsense word responses and other types for the vowel and coda substitutions. The low performance for the position 4/nonsense word response may have been caused in part by forgetting the example analogy, but there was certainly a tendency for the subjects to pause before giving a nonsense word response, despite having been assured at the start of the experiment that some of the appropriate responses would be nonsense words. There seemed to be a sense of confirmation giving a real word response that was lacking with a nonsense word response. In fact, the majority of errors given for the position 4/nonsense items were real words that were phonologically similar to the target nonsense word response. Obviously, though, the effect of different substitution types was still stronger than this word bias, since the nonsense word response pattern paralleled that of the spelling match and spelling mismatch responses.

It was easier to substitute a single unit (e.g., onset, nucleus, coda) than a double unit (head, rhyme, margins), which would suggest that the factor determining difficulty was simply what proportion of the syllable changed. This explanation works for certain comparisons (e.g., when the two phonemes /p/ and /l/ comprised an onset or coda, performance was much better than when they comprised margins), but it does not account for the advantage of rhyme (a double unit) substitution over nucleus and coda (single units) substitutions. Nor does a simple number explanation account for the differences in results with identical phonemes in different positions. The phonemes (/l/ and /e/) were treated very differently depending on whether they comprised the rhyme or the head, with the clear advantage for rhyme substitutions supporting the reality of a right-branching syllable structure for English.

The results for the special cases involving pre-vocalic glides and post-vocalic /r/ supported the suggestion that post-vocalic but not pre-vocalic glides should be described as part of the vowel nucleus in English. The glide /y/ before /u/ was found to be treated as part of the nucleus in at least some speech errors (Shattuck-Hufnagel, 1986), although those results were fairly ambiguous between the onset and the nucleus. In the present study, post-vocalic, but not pre-vocalic, /r/'s and glides were treated as part of the nucleus.[13] The question of what should be included in the nucleus and

---

[13]Post-vocalic nasals and /l/ were treated as part of the coda, a decision which seemed justified by the present results, since there was no difference in the performance with stimuli involving post-vocalic /l/, nasals, and obstruents (see Appendix C for means). These segments were treated differently in Experiment 4. See Chapter 5 for further

indeed the syllable (i.e., appendices, as suggested by Vennemann, in press; Selkirk, 1984; and Giegerich, 1985; among others) remains unresolved.

It is encouraging that the pattern for preferred errors was very similar to those of proportion correct and reaction time. If onset and rhyme are somehow more natural units than the others tested, this is exactly what one would expect. It seems reasonable that the most natural strategy when dealing with an overly difficult task would be to resort to the substitution type that was most readily accessible (i.e., onset or rhyme). In contrast to rhyme (26% of the total number of errors) and onset errors (32% of the total), head errors were almost non-existent (3% of the total) and margins errors were also rare (8% of the total).

It is difficult to claim that the differential results for substitution types were simply a function of task demands, since examples were provided for each set, the sets were randomly ordered for each subject, and there was exposure to all the substitution types. If training is used as an argument, then margins and coda substitutions should have been preferred, since there were four sets of each of these substitution types in the test and only three of each of the others. There is another possible explanation for the advantage of onset and rhyme, however, that involves prior learning outside of the experimental setting. These two units may be more accessible because of real-life experience. Alliteration (onset reduplication) and rhyme are frequently used poetic devices in English and the ease of access found in this study may simply reflect familiarity with these types of manipulations.

---

discussion of the question of the make-up of the nucleus.

One way to counteract the prior learning effect might be to train people on each type of substitution until they could perform all types and then test to see which substitution resulted in more errors or slower reaction times. A problem with this approach is that it would be difficult (impossible?) to match all the prior experience of rhyme and onset and to equate many years of language experience with a training session of a few minutes, or even days or weeks. If people could not be trained on a particular substitution type, then no data would be obtained of the sort desired.[14] Trials to criterion would have to be measured in that case, but this would not eliminate the problem of prior exposure to onset and rhyme.

Although the criticism of prior experience with manipulating onsets and rhymes might be applied to the results of the present study and the previous two, it is difficult to do the same for a task which requires perceptual judgments rather than language production or manipulation. This was part of the motivation for the phonetic similarity task of Experiment 4.

[14] It is very likely that criterion would not always be reached for margins and head substitutions, since some people in the pilot test could not do these even after they had been given detailed explanations of what was required.

Experiment 4

Phonetic Similarity Judgment Task

The preceding experiments all involved language production tasks; none unequivocally involved language perception alone. The case that has been made so far for the psychological reality of sub-syllabic units is only valid for the mode that was used to establish it. It is also necessary to establish the reality of sub-syllabic units for perception, or else it could be argued that the 'reality' of the first experiments was partly an artifact of production task demands (i.e., language strategies used in artificial situations) and had little relationship to general language knowledge.

Judging the phonetic similarity of pairs of stimuli is a task which requires access to phonological perception. Features, phonemes, or groups of phonemes (e.g., those relating to onsets, codas, etc.) can be systematically varied to compare the effect of phonetic changes at various levels of phonological organization. Phonetic similarity judgments (PSJs) were used by Nelson and Nelson (1970) to compare the effects of 'formal' (orthographic) and 'acoustic' (sound) similarity in pairs of words. They found significant effects for both factors, as well as an interaction between them. Vitz and Winkler (1973) replicated the Nelson and Nelson study but argued that the visual presentation technique the latter had used increased the influence of orthographic similarity. Vitz and Winkler first asked subjects to make PSJs of aurally presented pairs of words and they then compared these judgments to scores based on the phonemic distance of comparison words to target words (predicted phonemic distance or PPD). They found very high correlations between their prediction measure and the PSJs (on the order of r=.9).

Derwing (1976) used the PSJ technique to assess the relative contributions of orthographic and phonological similarity and concluded that the phonetic dimension was more important; a result which Derwing and Nearey (1986) confirmed using aural, rather than written presentation. Nearey (1981) used the PSJ technique to investigate whether stops following syllable-initial /s/ were perceived as more similar to the voiced or to the voiceless phoneme (i.e., bill:spill *vs* pill:spill).[1] Derwing and Nearey (1986) also used the PSJ technique and predictions from phonemic distance metrics to address certain questions of phonemic theory (for example, the phonemic status of vowel nuclei, particularly diphthongs, stops after syllable-initial /s/, and the velar nasal) and found, in general, that a simple phonemic model seemed to be a fairly reasonable predictor of these phonetic similarity judgments.

However, there were some differences in the PSJs reported by Derwing and Nearey (1986) which the phonemic model could not account for. Both the degree and the locus of difference had an effect (see Figure 4.4.1). Multiple feature changes within the same segment had more of an effect on judgments than single feature changes (e.g., a change of voice, place, and manner in the final segment was much more serious than just one change; compare /hɪp-hɪz/ with /pɪt-pɪd/). There was also an enhanced effect of differences in word-final consonants as compared to word-initial consonants (e.g., the three feature change of /hɪp-hɪz/ lowered the similarity judgments much more than the equivalent change

[1]The study was also directed to other questions, such as the role of orthography in making these PSJs, but the point of interest for the present discussion is how the technique has been used to investigate strictly phonological issues.

in word-initial position /pɪp-zɪp/). A global feature solution did not work either, because two changes distributed across the syllable (e.g., a change of voice in both initial and final position) had more effect on similarity ratings than three feature changes in one position.[2]

Figure 4.4.1

Effect of different distribution of feature changes

C	V	C	:	C	V	C
αvce		αplace		βvce		βplace

is less similar than:

C	V	C	:	C	V	C
		αvce αplace				βvce βplace

The differential effect of initial and final position reported by Derwing and Nearey (1986) can also be viewed as an onset-coda difference, since three-phoneme (CVC) words were used. It is impossible to tell whether the 'position' effect was based on location within the word (i.e., initial or final positon) or changes in different sub-syllabic units (i.e., onset or coda). Just as there was an effect of degree of difference (i.e., number of features changed) and locus of difference (i.e., whether the features changed were within one segment or distributed among segments), there might be different effects based on which sub-syllabic unit is changed.

[2]Note that this is in line with Vitz and Winkler's (1973) phonemic model, since two changes in two positions would necessarily involve two phonemes, whereas three feature changes in one position affects only one phoneme.

The present set of experiments was designed to address the question of whether a model of syllable structure including sub-syllabic units such as the onset, nucleus (including post-vocalic liquids and nasals), and coda could account for phonetic similarity judgments in words with consonant clusters.[3] The effect of changes in various sub-syllabic units on similarity judgments gives an indication of relative weight. For example, Derwing and Nearey's (1986) results that changes in the onset (e.g., bat:fat) have less of an effect than changes in the coda (e.g., bat:bag) suggests that the coda or rhyme was weighted more heavily. If there is a relative effect of, for example, changes in the nucleus and the coda, the question of the membership of the nucleus can be assessed by comparing changes in post-vocalic glides, liquids, and nasals with post-vocalic obstruent changes.

Experiment 4A Number of Features and Units Changed

A special set of stimuli was designed to test the relative effect of changes in consonant clusters within a single unit (e.g., *sp*eak:*bl*eak for onsets and bu*st*:bu*gs* for codas) and parallel changes thought to cross the nucleus-coda boundary (e.g., pee*rs*:pee*ked*). Both consonants of the cluster were changed in four-phoneme words. Different combinations of small (one feature change per consonant) and large (more than one feature change per consonant) changes were included to assess the Derwing and Nearey (1986) prediction that the number of features changed affects similarity judgments. A phonemic metric based on the proportion of changed phonemes would predict no difference in the similarity

[3]Many thanks to Dr. T.M. Nearey for suggestions and help with the design of this set of experiments.

judgments among the stimuli, since two out of four phonemes were changed in each case. A metric which treats all consonant clusters as units (such as Vitz & Winkler's cluster index, 1973) would also predict that all of these changes would have equal effects on similarity judgments. Although there are no explicit predictions for similarity judgments from the models of syllable structure discussed in Chapter 3, the possibility exists that some sub-syllabic units may contribute more to sound similarity judgments than others.

4A Method

Subjects. Students enrolled in Introductory Linguistics at the University of Alberta took part in the study during the first week of classes. All reported that they had normal hearing and were native speakers of English. They were tested in two groups (n=29 and n=20). The listening conditions for the latter group were less than ideal, as they were (unintentionally) subjected to background noise perceptible through the sealed windows of the room during the second half of the test (there was a rock group performance in an outdoor area several hundred meters from the experimental room).[4]

Materials. Thirty pairs of real or phonotactically possible English words were designed to test the effect of near mismatches (coded here as x, a single feature change) and large mismatches (coded here as X, with at least a manner change and one other feature change) in various units. Three combinations of near and large mismatches for consonant clusters

[4] Even so, there was very little difference in performance between the two groups. See the Results section for details.

were included (i.e., xX, Xx, XX). Changes in cluster onsets (i.e., CCVC:XxVC; e.g., /swayp/:/grayp/) and cluster codas (i.e., CVCC:CVXx, e.g., /hIps/:/hIdz/) were compared, as were changes between cluster codas and those involving complex nuclei (CVRC:CVXx, e.g., /pirs/:/pivz/). Single near mismatches in the pre-vocalic and post-vocalic consonant of four two-segment words were also compared (e.g., CV:xV /bæ:dæ/ *vs* VC:Vx /æb:æd/, see Table 4.4.1; and Appendix E for the complete list).

Table 4.4.1

Comparisons for onset, coda, and complex nucleus changes (with example word pairs)

Onset		Coda		Complex Nucleus	
CV:xV	/bæ:dæ/	VC:Vx	/æb:æd/	VR:Vx	/æn:æm/
CCVC:XXVC	/spik:blik/	CVCC:CVXX	/dɛsk:dɛks/	CVRC:CVXX	/pirz:pikt/
CCVC:xXVC	/spik:ʃrik/	CVCC:CVxX	/wɪsp:wɪzd/	CVRC:CVxX	/sæŋk:sænd
CCVC:XxVC	/swayp:grayp/	CVCC:CVXx	/hɪps:hɪdz/	CVRC:CVXx	/pirs:pivz/

C=Consonant R=Liquid or nasal of complex nucleus V=Vowel
x=near mismatch X=large mismatch
The vowel nuclei are underlined.

The 30 word pairs were randomly ordered along with the 94 pairs from the following two experiments, with the only constraint being that no two adjacent pairs contained the same word. At the end of the list, there were 16 additional word pairs for a total of 140 word pairs.

Procedure. The words were recorded by a female native speaker of Western Canadian English using a Sony TC K55II stereo cassette deck on one channel and a Sennheiser microphone.[5] To control for the possible effect of recency on judgments, each pair of words was said once, then repeated in the reverse order with the number of the item following (e.g., "/klæsp/:/klæst/, /klæst/:/klæsp/, number 1") at a rate of approximately one word per second.[6]

The recording was played back for subjects in a classroom using the same kind of recorder and two Heco Soundmaster 5 speakers. The test portion was preceded by five example pairs, chosen to demonstrate the range of possible mismatches (from a single near match in the onset to an entire rhyme mismatch). The students were asked to listen to the example pairs and were then asked to note the perceived similarity of sound of the test pairs according to the scale given at the top of their test sheet (reproduced in Figure 4.4.2), where *0* indicated the pairs were completely different and *10* indicated they were exactly alike. The entire test session lasted approximately 30 minutes.

Figure 4.4.2

Scale used for phonetic similarity judgments

0	1	2	3	4	5	6	7	8	9	10
Completely different										Exactly alike

[5] Many thanks to Tracey M. Derwing for recording the words and assisting with the collection of the data.

[6] Thanks to Dr. T.M. Nearey for this suggestion. See also the Discussion section for further explanation.

4A Results & Discussion

Because of the background noise during the second half of the test for one group of subjects, independent *t*-tests were carried out between the two subject groups on the mean similarity score for all 140 pairs of stimuli. Nine tests were significant at the .05 level, as opposed to the seven expected purely by chance (i.e., .05 x 140=7); since these were distributed fairly evenly throughout the session, the two groups were combined.

To determine the effects of type of unit changed and degree of change, the mean similarity judgments for each type of four-phoneme word pair were subjected to a repeated measures analysis of variance (ANOVA), with unit (3 levels; cluster onset, cluster coda, and cluster that crossed the nucleus-coda boundary) and type of change (3 levels; large mismatches for both consonants XX and both combinations of a large mismatch and a near mismatch Xx, xX). Both the effect of unit ($F(2,96)=45.29$, $p<.0001$) and the interaction ($F(4,192)=16.07$, $p<.0001$) were significant, although the effect of type of change was not. The means for each combination of unit and size of change can be found in Table 4.4.2, which also includes the means for the comparison pairs for the single near mismatch in the onset, coda and nucleus of two-phoneme words. The mean similarity judgments for the individual word pairs can be found in Appendix E.

Table 4.4.2

Mean similarity judgments for onset, coda, and complex nucleus cases (standard deviations in italics)

Onset			Coda			Complex Nucleus		
CV:xV	7.4	*1.6*	VC:Vx	6.4	*1.6*	VR:Vx	6.8	*2.4*
CCVC:XXVC	5.1	*1.9*	CVCC:CVXX	5.1	*1.2*	CVRC:CVXX	3.0	*1.7*
CCVC:xXVC	4.9	*1.7*	CVCC:CVxX	4.5	*1.5*	CVRC:CVxX	4.3	*1.4*
CCVC:XxVC	5.1	*1.9*	CVCC:CVXx	5.2	*1.5*	CVRC:CVXx	2.8	*1.7*

C=Consonant R=Liquid or nasal of complex nucleus V=Vowel
x=near mismatch X=large mismatch
Vowel nuclei are underlined.

Independently of the ANOVA analysis, paired comparison *t*-tests were carried out between the onset and coda cases and the coda and complex nucleus cases in order to examine the effect of type of unit, and comparisons within each type were done to determine the nature of the interaction between unit type and size of change. There was little difference between changes in onsets and codas; the only significant paired comparison *t*-test for the onset-coda cases was the two-phoneme word comparison (using Bonferroni's $t(48)>3.23$ for k=20 tests and p<.05; Myers, 1979) and none of the comparisons between different types of changes (i.e., XX, Xx, xX) were significant for onsets or codas.

In contrast to the onset-coda comparisons, there was a significantly greater effect of large changes in the complex nucleus compared with the coda (i.e., the CVXX-CVXX and the CVXx-CVXx cases; vowel nuclei underlined in each case, $t(48)>9.28$). In addition, a large mismatch in the post-vocalic position of the complex nucleus cases had a significantly greater effect than a near mismatch in that position (i.e., C_VXX_ and C_VX_x were both significantly different from C_VxX_ [$t(47)>5.94$], but not from each other).

The C_VxX_ cases, then, did not differ significantly from the simple nucleus cases but did differ from the other two complex-nucleus cases. The explanation for this apparent anomaly lies in the post-vocalic consonants involved in the different comparisons. The C_VXX_ and C_VX_x cases both involved the phonemes /r/ and /l/ exchanging with obstruents and nasals, whereas the C_VxX_ cases involved only nasal to nasal changes (i.e., for reasons of simplicity in designing the stimuli the place feature change of nasals was used for the single feature comparison). The difference between a near and large mismatch in the post-vocalic position of complex nucleus words thus may have been caused by the specific phonemes involved rather than the size of change *per se*. The post-vocalic nasal did not seem to be treated differently from post-vocalic obstruents, since there was also no difference between the single segment change case for the coda (i.e., _V_C:_V_x, /æb:æd/) and for the complex nucleus (i.e., _VR_:_Vx_, /æm:æn/). Nasal phonemes are probably not perceived as being part of the vowel nucleus. Nasals were not treated as part of the nucleus in the design of the substitution task of the previous experiment (Experiment 3) and probably should not have been here.

The results of the first experiment thus provide support for the notion that the type of sub-syllabic unit affected is more important than simply the number of features changed (i.e., degree of change) since there were significant differences between the changes in post-vocalic consonant clusters depending on whether or not these involved the complex nucleus. The first experiment was designed to assess the relative effect of degree and location of change, but did not vary the number of phonemes affected within a word. The second experiment was designed to vary both the number of phonemes affected and the number of sub-syllabic units involved in the change.

Experiment 4B <u>Number of Phonemes and Units Changed</u>

The second experiment was designed to examine the effect of single feature changes in one or more phonemes and one or more sub-syllabic units. Of special interest was the effect of changes in post-vocalic consonants such as 'resonants' (coded as R; specifically the liquids /l/ and /r/ and the nasals), which have been treated as part of complex nuclei, and other post-vocalic consonants. Single phoneme changes that affected part or all of a syllabic unit (i.e., changing one phoneme in the coda cluster of a CC<u>V</u>CC word, or changing the coda of a CC<u>VR</u>C word; with the vowel nucleus underlined in each case) and two phoneme changes which affected either one or two syllabic units were included to assess the relative effects of changes in different syllabic units (e.g., the onset *vs* the coda), the effect of the number of phonemes changed, and the effect of the type of vowel nucleus.

Also included in this set of stimuli were specific pairs to test the predictions of two of the models discussed above (Chapter 3), which treat certain word-final consonants in special ways. The metrical model proposed by Selkirk (1984) considers all word-final coronal obstruents as appendices separate from the rest of the syllable. The core plus affix model of syllable structure proposed by Fujimura (1975; 1976) and Fujimura and Lovins (1978) counts word-final apical consonants as affixes if they agree in voicing with the preceding segment (such as the final /d/ of tend but not the final /t/ of tent). Comparisons were made to determine whether there was a difference between changing a consonant that could be described as an appendix or affix and other consonants.

4B Method

Subjects & Procedure. See Experiment A.

Materials. The stimuli for this second experiment were designed to test the effect of one and two near-segment mismatches (i.e., coded here as x, a single feature change) in various positions of five-segment monosyllabic words, half of which could be analysed as cluster onset+simple nucleus+cluster coda (i.e., CCVCC:CCVCx, e.g., /klæsp/:/klæst/) and half as cluster onset+complex nucleus+simple coda (i.e., CCVRC:CCVRx, e.g., /sport/:/spord/, with nuclei underlined). Four English words (i.e., clasp, flux, blast, drift) were used for the simple nucleus comparisons and four for the complex nucleus comparisons (i.e., twelve, crumbed, blared, sport). There were four one-segment mismatch cases and four two-segment mismatch cases (the pattern of changes were those shown in Table 4.4.3). The eight words, two levels of mismatch, and four positions of the mismatch resulted in a total of 64 word pairs for this stimulus set (see Appendix F for the complete list).[7]

[7]There were four word pairs which were included in the study but not the analysis because more than the appropriate number of changes were made (see Appendix F for details).

Table 4.4.3

Near phoneme mismatches in various positions (w[illegible]ample pairs)

Simple Nucleus		Complex Nucleus	
One segment changes:			
CCVCC:CCVCx	/klæsp/:/klæst/	CCVRC:CCVRx	/sport/:/spord/
CCVCC:CCVxC	/klæsp/:/klæʃp/	CCVRC:CCVxC	/sport/:/spolt/
CCVCC:CxVCC	/klæsp/:/kræsp/	CCVRC:CxVRC	/sport/:/skort/
CCVCC:xCVCC	/klæsp/:/glæsp/	CCVRC:xCVRC	/sport/:/ʃport/
Two segment changes:			
CCVCC:CCVxx	/klæsp/:/klæʃt/	CCVRC:CCVxx	/sport/:/spold/
CCVCC:CxVxC	/klæsp/:/kræʃp/	CCVRC:CxVxC	/sport/:/stolt/
CCVCC:xCVCx	/klæsp/:/glæst/	CCVRC:xCVRx	/sport/:/ʃpord/
CCVCC:xxVCC	/klæsp/:/træsp/	CCVRC:xxVRC	/sport/:/ʃtort/

C=Consonant R=Liquid or nasal of complex nucleus V=Vowel
x=near mismatch The vowel nuclei are underlined.

4B Results & Discussion

A repeated measures ANOVA (2x2) was carried out on the similarity judgments on the 64 pairs designed to test the effect of near phoneme mismatches in different sub-syllabic units, with type of nucleus (simple or complex) and the number of changes (one or two) as the dependent, repeated factors (see Appendix F for the mean similarity judgments for each word pair). There was a significant effect for type of nucleus ($F(1,47)=100.60$, $p<.0001$), with changes in words with simple nuclei having less effect on similarity scores than changes in words with complex nuclei (6.3 *vs* 5.4, respectively). The effect of number of changes was also significant overall ($F(1,47)=430.88$, $p<.0001$), with one change having less effect than two (6.8 *vs* 4.9); the interaction of nucleus by number was also significant ($F(1,47)=29.36$, $p<.0001$). Two changes in complex nucleus words had the greatest effect on similarity judgments, although all of the means were significantly different from one another (see Figure 4.4.3). The means for each type of comparison are displayed in Table 4.4.4 to allow for direct comparisons of the effect of changes between syllables with different types of nuclei.

Figure 4.4.3

Mean similarity judgments for number of changes by nucleus type

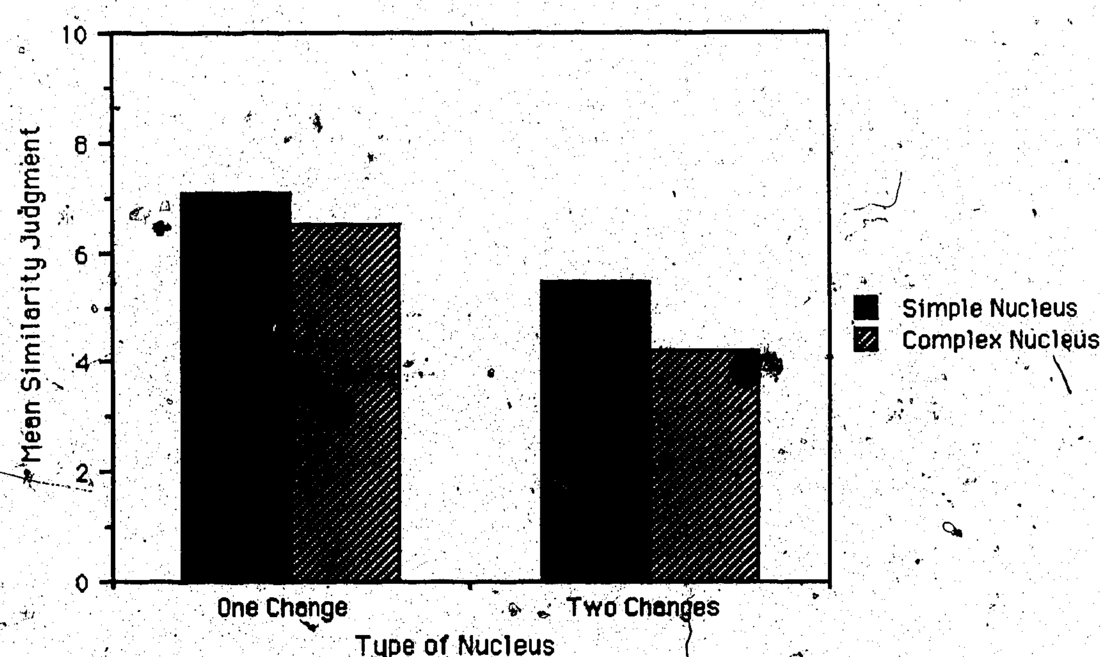


Table 4.4.4

Mean similarity judgments for simple and complex nucleus cases

Simple Nucleus		Complex Nucleus		*t*-tests
1 change:				
1. CCVCC:CCVCx	7.2	1. CCVRC:CCVRx	6.8	
2. CCVCC:CCVxC	7.0	2. CCVRC:CCVxC	5.0	*
3. CCVCC:CxVCC	6.9	3. CCVRC:CxVRC	7.0	
4. CCVCC:xCVCC	7.3	4. CCVRC:xCVRC	7.4	
	Mean=7.1		Mean=6.5	
2 changes:				
5. CCVCC:CCVxx	5.6	5. CCVRC:CCVxx	3.6	*
6. CCVCC:CxVxC	5.2	6. CCVRC:CxVxC	3.3	*
7. CCVCC:xCVCx	5.1	7. CCVRC:xCVRx	4.2	*
8. CCVCC:xxVCC	6.0	8. CCVRC:xxVRC	5.8	
	Mean=5.5		Mean=4.1	

* indicates a significant *t*-test comparison ($t(47)>4.63$, $p<.0001$)
C=Consonant R=Liquid or nasal of complex nucleus V=Vowel
x=near mismatch Vowel nuclei are underlined.

To examine the main effect of type of nucleus more closely, paired comparisons were made between changes in phonemes in equivalent positions relative to the vowel in each of the simple and complex nuclei cases. The effect seemed attributable primarily to changes directly affecting the nucleus. All of the pairwise comparisons between the simple and complex nucleus pairs involving changes in the post-vocalic position were significant (i.e., pair numbers 2, 5, and 6 and marked with an asterisk in Table 4.4.4; all $t(47)>4.63$, $p<.0001$; Bonferroni's criterion for the total number of *t*-tests carried out for this set of data was $t(47)>4.32$ for $k=120$ tests and $p=.01$). There was one case in which the nucleus was not involved in the change but the pairs were significantly different from one another (the number 7 pairs, which involved a change in the onset and the coda). To test whether this could be because a change in the single segment coda of a complex nucleus syllable word counted more than one change of a cluster coda, further comparisons were made.

Two changes in the simple nucleus cases had a roughly equivalent effect on similarity judgments, regardless of whether those changes involved two consonants of the onset or the coda (i.e., CCVCC:CCVxx or CCVCC:xxVCC), or one consonant in each of two units (i.e., CCVCC:CxVxC and CCVCC:xCVCx).[8] The complex nucleus case with two changes in the onset (i.e., CCVRC:xxVRC) was not different from any of the two phoneme change cases for the simple nucleus. Thus, two partial unit changes seemed equivalent to a whole unit change (at least comparing cluster onsets and codas). However, two changes that resulted in a partial onset and a complete coda change (i.e., the complex nucleus case CCVRC:xCVRx) received significantly lower similarity ratings than the two change cases with the simple nucleus (using Bonferroni's criterion of $t(47)>4.32$ for $k=120$ tests and $p=.01$).

Changes in the onset of simple nucleus words was equivalent to coda changes in those words, and onset changes in simple and complex nucleus words were equivalent. Changing the sonorant of a complex nucleus (e.g., CCVRC:CCVxC, /sport:spolt/) had a greater effect on similarity judgments than changing either the entire onset or coda of simple or complex nucleus syllables (e.g., CCVRC:xxVRC, /sport:ʃtort/; $t(47)>3.57$, $p<.0008$). This effect could not have been due only to the

[8]Changing the complete cluster coda of a simple nucleus syllable had a slightly greater effect than changing the whole onset, but not significantly ($t(47)=-1.43$). Changing the coda of a VC syllable also had a greater effect than changing the onset of a CV syllable (means of 6.4 and 7.4 respectively; see previous section for a discussion.)

particular consonants involved in the change (i.e., liquids and nasals), since changing pre-vocalic /r/'s was significantly less important for similarity judgments than changing post-vocalic /r/'s (6.9 and 4.6, respectively, $t(48)=11.1$, $p<.0001$). There was a difference between the effect of changes on nasals (i.e., CCVRC:CCVxC, /krʌmd:krʌnd/, mean=6.1) and liquids (e.g., /sport:spolt/, mean=4.6; $t(48)>5.44$, $p<.0001$)) in the complex nucleus, just as there was a difference between nasals and liquids in the first set of stimuli (see previous section).[9]

To examine the predictions suggested by Selkirk's (1984) metrical model, pairs of stimuli with changes in the syllable-final consonant which did and did not involve appendices were compared. Of the CCVRC:CCVRx set, there were three cases for which both members of the pairs had a consonant that would be defined as an appendix (i.e., a coronal obstruent, e.g., /spor*t*:spor*d*/, mean=7.0) which did not differ significantly from the one case where neither member of the pair had an appendix (i.e., /twɛlv:twɛlə/, mean=7.0). For the CCVRC:xCVRx set, there were two cases where the final consonant of both pairs could be classed as an appendix (e.g., /spor*t*:ʃpor*d*/, mean=3.8) and one case where neither member of the pair had an appendix (e.g., /twɛlv:dwɛlf/, mean=5.0). The match pairs (i.e., both with appendices or neither with appendices) did not differ significantly from one another. Clearly, a larger and more systematically designed list would be required to make any strong claims about the reality of the appendix as an independent unit.

Relevant pairs were also chosen to test the predictions of the

[9]The three PSJ experiments were designed together before any was analysed, so the nasals were included in the nucleus in this experiment. Experiment C used nuclei with only post-vocalic /r/'s, no /l/'s or nasals.

core+affix model (Fujimura, 1975; 1976). There were three cases from the CCVCC:CCVCx and CCVRC:CCVRx sets which involved a change in the status of the final consonant (from non-affix to affix; an apical consonant which agrees in voicing with the core-final segment, e.g., /klæs*p*:klæs*t*/, 7.6) and three cases for which both members of the pair had affix consonants (e.g., /flʌk*s*:flʌk*t*/, 6.1). This difference was significant ($t(47)=4.73$, $p<.0001$). For the CCVRC:xCVRx set, there were three cases with a change in affix status of the final consonant (mean=3.3) and one case for which neither of the pair had a consonant that qualified as an affix (mean=5.0). This difference was also significant ($t(48)=5.70$, $p<.0001$).

~~One would expect~~ that a change in the status of a syllabic unit (i.e., an affix) would affect similarity judgments more than changes which did not alter syllable structure. However, the single change cases which involved a change in the final consonant from an affix to a non-affix were treated as significantly more similar than those for which both members of the pair had affixes. For the two change cases the results were exactly the opposite (although more in line with expectations). Cases involving a change in the affix status of the word-final consonant were less similar than those without affixes word-finally. These comparisons give some indication of the reality of the distinctions these two models make, but a strong case cannot be made for or against either of these models without more evidence.

The results of this set of stimuli thus support the importance of the type of vowel nucleus, since changes in words with complex nuclei had a greater effect overall than changes in simple nucleus words. Two changes had more effect than single changes, independent of what units were affected. However, the unit was important, since a single change was significantly more important if it affected the nucleus or the entire coda than if it only affected part of the onset of coda. Combined with the results of the previous set of stimuli, a hierarchy has emerged with sub-syllabic units as most important, phonemes intermediate (in that a two-phoneme coda was affected differently from a single phoneme coda), and features (degree of change) as least important but still influential. The last experiment of this set was designed to test the effect of large feature changes in different number of phonemes and different proportion of sub-syllabic units.

Experiment 4C <u>Changes in the Rhyme</u>

A simplistic prediction of an onset-rhyme syllable model would be that a change in the rhyme would have a roughly equivalent effect, regardless of where it occurred (i.e., whether in the vowel, the post-vocalic consonant of the nucleus or the coda). The onset-nucleus-coda model, alternatively, would predict that the location of the change would have an effect, based on the relative weight of the rhyme component (i.e., nucleus or coda), and the phoneme model would predict that solely the number of phonemes affected would determine the similarity judgments. A third set of stimuli was therefore included to assess whether a phoneme model, a simple onset-rhyme model or a model that separated the rhyme into nucleus and coda best predicted similarity judgments. Words with rhymes ranging from one to four phonemes in length, with and without final consonant clusters (codas) were included.

4C Method

Subjects & Procedure. See Experiment A.

Materials. A set of thirty word pairs was designed to test the predictions of phonemic weight (as per Vitz & Winkler's PPD, 1973) *versus* the weight of sub-syllabic units (specifically nucleus, coda, and rhyme). These pairs had large mismatches in a variety of different combinations in the rhyme portion of monosyllabic words. There were vowel changes in words without codas (i.e., CV:CX, e.g., /bi:ba/), without onsets (i.e, VC:XC, e.g., /it:ɔt/), and with both onsets and codas (i.e., CCVCC:CCXCC, e.g., /bist:bɔst/) and a variety of changes involving the post-vocalic sonorant and coda. The sonorant was removed in some cases with no change in the coda (i.e., CVRCC:CVCC, e.g., /ʃarps:ʃaps/) and with changes in the coda (i.e., CVRCC:CVXX,e.g., /ʃarps:ʃagz/), and in one case the coda was changed with no change in the nucleus (i.e., CVRCC:CVRXX, e.g., /bardz:barft/). In two additional cases the complete rhyme was changed (i.e., CVRCC:CXXX, e.g., /bardz:bist/; and CVCC:CXXX, e.g., /bist:bavd/; see Appendix G for the complete list of stimuli).

4C Results & Discussion

There were significant differences in the similarity judgments for the word pairs set up to test the effect of changes in the rhyme, which a simple onset-rhyme syllable model could not account for. The results seemed best accounted for by an analysis based on the sub-syllabic units of onset, nucleus, and coda (the means can be found in Table 4.4.5). Changing one unit out of two (in this case actually changing one phoneme out of two;

numbers 1 and 2) was significantly different from making only one change out of three (numbers 3, 4, and 5), or two out of three (numbers 7 and 8) although pair number 6 did not differ significantly from pair numbers 1 and 2. If only one unit out of three was changed, then the effect was much less than if two units out of three were changed (compare the nucleus plus coda changes (6, 7, 8) with the nucleus (3, 4) or simple coda (5) changes), with all pairwise comparisons significant (second and third columns of Table 4.4.5; all pairwise comparisons for this set of data were based on Bonferroni's criterion of $t(47)>3.93$ for $k=28$ tests and $p<.01$).

Table 4.4.5

Mean similarity judgments for different unit changes in the rhyme

1 out of 2 units		1 out of 3 units		2 out of 3 units	
1. VC:XC	3.4	3. CVCC:CXCC	4.2	6. CVRCC:CVXX	2.8
2. CV:CX	3.4	4. CVRCC:CVCC	4.6	7. CVRCC:CXXX	1.7
		5. CVRCC:CVRXX	4.4	8. CVCC:CXXX	1.6

C=Consonant R=/r/ of complex nucleus V=Vowel
X=large mismatch

None of the one change out of three units cases (i.e., 3, 4, 5) were significantly different, nor were the one change out of two units cases (i.e., 1 and 2) nor the complete rhyme changes (i.e., 7 and 8). Changing a complex nucleus by removing the sonorant was a less serious change than completely changing the nucleus (number 6 was significantly different from numbers 7 and 8, Table 4.4.5).[10]

The relative similarity of the pairs in decreasing order for the predictions from Vitz and Winkler's (1973) PPD and the order obtained in

[10] All complex nuclei for this set of word pairs had post-vocalic /r/'s, not /l/'s or nasals.

the present study can be found in Table 4.4.6. The two correspond closely, although some different predictions can be made. Since pair numbers 3, 4, and 5 have different numbers of phoneme matches, the Vitz and Winkler metric predicted differences between means that the syllabic unit model did not; the lack of a significant difference in the results thus supports the latter model. Both models predicted no differences between pair numbers 1 and 2 and indeed, there was none; both models fell short, however, in the comparison between pair numbers 6 and 7, since neither could account for the observed difference. The phoneme-based model predicted a significant difference between numbers 6 and 8, which was realized, and between 7 and 8, which was not; whereas the unit-based model predicted no differences. It is difficult to choose one model over the other, solely on the basis of the results for this set of stimuli. The degree of similarity within the unit also seemed to matter.

Table 4.4.6

Model predictions for rhyme changes based on phonemes *vs.* sub-syllabic units (mean similarity judgments are also provided)

Vitz & Winkler Predictions

Pair Shape	Difference	Mean
4. CVRCC:CVCC	1/5	4.6
3. CVCC:CXCC	1/4	4.2
5. CVRCC:CVRXX	2/5	4.4
1. VC:XC	1/2	3.4
2. CV:CX		3.4
6. CVRCC:CVXX	3/5	2.8
7. CVRCC:CXXX		1.7
8. CVCC:CXXX	3/4	1.6

Sub-syllabic Unit Predictions

Pair Shape	Difference	Mean
4. CVRCC:CVCC		4.6
5. CVRCC:CVRXX	1/3	4.4
3. CVCC:CXCC		4.2
1. VC:XC	1/2	3.4
2. CV:CX		3.4
6. CVRCC:CVXX		2.8
7. CVRCC:CXXX	2/3	1.7
8. CVCC:CXXX		1.6

C=Consonant R=/r/ of complex nucleus V=Vowel
X=large mismatch

General Discussion

The three sub-syllabic units of onset, nucleus, and coda emerged unambiguously as units of perception in this experiment. A phoneme-based model failed to account for the pattern of results, as did a model based on consonant clusters, and the simple onset-rhyme syllable model was inadequate. In the first set of comparisons (Experiment A), there was no difference between changes in cluster onsets and codas, but changes in consonant clusters which crossed the nucleus-coda boundary had a much greater effect than changes in simple codas. The second set of comparisons (Experiment B) showed consistent differences between changes in syllables with simple and complex nuclei which could not be explained simply in terms of the number of changes involved (a phonemic metric), or between changes in consonant clusters (a cluster metric). The third set of comparisons (Experiment C) tested a simple onset-rhyme model, which predicted that any changes that affected the rhyme would be equal. The changes had very different effects based on whether they involved the vowel, the post-vocalic liquid, or the coda, suggesting that the nucleus and coda are weighted very differently.

The Complex Nucleus

The results of the present set of experiments provides strong support for a syllabic model including a complex nucleus of vowel plus liquid. Despite some suggestions (discussed in Chapter 3, above) that the nasal belongs with the vowel, there was little evidence of this here. Vowel changes and changes in the post-vocalic liquid affected similarity judgments significantly more than changes in post-vocalic nasals and obstruents or pre-vocalic consonants, suggesting that the nucleus is weighted more heavily in the syllable than the onset and the coda. Two changes across the nucleus-coda boundary had a much greater effect on judgments than two changes totally within either the cluster onset or the cluster coda of a simple nucleus word pair. A simple phonemic metric could not account for this result. Vitz and Winkler (1973) developed a cluster count to make an alternative set of predictions for words involving consonant clusters; the main effect of this alteration of their phonemic metric was to give proportionately more weight to the vowel. Even using this refined version, however, the crucial predictions for the word pairs used here (both the 30 word pairs of Experiment A and the 64 of Experiment B) were the same and did not account for the complex nucleus difference.

The Relative Weights of the Onset and the Coda

Despite some earlier suggestions that onsets and codas are weighted differently, the only significant difference between changes in the onset and the coda of simple nucleus words in the present study was between CV:xV̲ and VC:V̲x pairs of Experiment A, with the onset change treated as less important than the coda change. In all of the other comparisons, a single change in the onset was roughly equivalent to a single change in the coda, as was the comparison of two changes in the onset or coda. The results then suggest that there is no motivation to claim that the onsets and codas of CVC and CCVCC type syllables have different weights.

This is not consistent with the finding in other PSJ studies that word-final consonants functioned differently from consonants in other syllable positions. Derwing and Nearey (1986) reported, for example, that manner changes and two or three feature changes in word-final consonants affected similarity judgments more than changes in word-initial consonants, but they proposed no explanation for this effect. However, Nearey (personal communication) has suggested that this could have been caused by a recency effect (i.e., more recent changes may affect judgments more than earlier ones).

To try to eliminate any effect of recency in the present experiment, a control was built into the presentation of the word pairs.[11] The word pairs were said first in one order, repeated in the opposite order, and then the number of the word pair was given (corresponding to the number on the answer sheet). The item number was included to function as a 'suffix' (an intrusive element to recall), which has been shown in verbal recall studies (i.e., Crowder & Morton,1969; Darwin & Baddeley, 1974) to eliminate the recency effect. The results here thus support Nearey's suggestion that the initial-final difference reported in Derwing and Nearey (1986) was a recency effect.

Tests of other Models

The simple phonemic metric proposed by Vitz and Winkler (1973) worked fairly well for the stimuli of Experiment C, but a model that included sub-syllabic units was necessary to account for the pattern of results in Experiments A and B. Neither the appendix model suggested by proponents of metrical theory (e.g., Selkirk, 1984) nor the core plus affix model put forward by Fujimura (1975; 1976) were strongly supported by the present results. Further work would be necessary to test these adequately. There was no evidence that any refinement of the notion of nucleus consisting of vowel plus liquid (i.e., such as the qualification for long vowels suggested by Clements & Keyser, 1983) is necessary, although the claim that nasals can be part of the nucleus must be re-evaluated. One difficulty with the present technique, however, is that it is difficult to separate the effects of changes in the post-vocalic liquids /r/ and /l/ because the usual near feature change was from one to another.

Summary

In general, the results of the phonetic similarity task support the psychological reality of the sub-syllabic units of onset, complex nucleus, and coda, and provide evidence for the relative weights of these units within the syllable. The onset and coda are roughly equivalent in weight and the nucleus is weighted more heavily, whether it is simple or complex. Both a simple structure of sub-syllabic units directly subordinate to the syllable and a right-branching structure including a rhyme unit could predict this pattern. In fact, it is impossible to separate the predictions of a hierarchical model from the predictions based on relative weights of units.

It is clear that the simplistic notion that sub-syllabic units can replace phonemes and features in phonological description is premature, since the degree of change (features), the proportion of phonemes changed within a unit, and the type of sub-syllabic unit changed all had important effects on the similarity judgments.

CHAPTER 5: CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

The previous chapter of experimental work included four experiments which tested the reality of sound representation intermediate between syllables and phonemes (sub-syllabic units). The results clearly support a model of syllable structure with a major branch or break between the onset and rhyme and provide considerable support for a secondary division or branch between the vowel nucleus and the coda.

5.1 Review of the experiments

The counting task (Experiment 1) tested the difference in ability to count syllables, sub-syllabic units, and phonemes. Both children and high school students could count syllables well, a finding which fits with previous work (e.g., Liberman *et al.* 1974; Fox & Routh, 1975, etc.). A new finding was that all groups of subjects could count sub-syllabic units as well as or better than phonemes. Of the three analyses of sub-syllabic counting used in this study, the 'rhyme' analysis (i.e., that syllables are comprised of onsets and rhymes) accounted for responses best overall (38%), with no differences among subject groups. The 'vowel' analysis (treating syllables as divided into onset-vowel-coda) resulted in scores which were lower than the rhyme analysis (28%) but showed a pattern of differences among the groups. The 'nucleus' analysis (onset-nucleus-coda, treating post-vocalic liquids and nasals as part of the nucleus, separate from the coda), on the other hand, was almost as high (35%) as the rhyme analysis and still showed a group effect, with increasing scores with age.

Kindergarten children's responses for sub-syllabic counting were best accounted for with an onset-rhyme analysis, but the Grade One children and the High School students responses were best accounted for with a nucleus analysis. In all groups the nucleus analysis seemed to be better than the vowel analysis.

The children were not able to count phonemes very well and there was no significant difference between the scores of the Kindergarten and Grade One students, suggesting that developmental claims for phonemic analysis ability should be re-evaluated. The low scores (51%) for phoneme counting for the High School students were somewhat unexpected, since phonemes have been assumed to be accessible to literate adults. The high correlation of mean number provided for each word with phoneme count (r=.95) and a suggestion by Perin (1983) provided an explanation for this apparent anomaly. Short words (e.g., CVC) were counted fairly accurately but the mean number for longer words (e.g., CCVCC) was consistently lower than the phoneme count, which resulted in a high correlation but low accuracy. This reinforces the suggestion made earlier that experiments using only stimuli with simple canonical forms can provide, at best, limited information about the representation of sound, and may, in fact, be misleading.[1]

One of the biggest problems with the counting task was the difficulty of determining what units people were actually counting, except by inference from the consistency of the response patterns exhibited. The deletion-by-analogy task of the second experiment made it more obvious what units

[1]The same caution should be applied to the results of Experiment 1 here, which used few multisyllabic words for the sub-syllabic and phoneme counting tasks (to keep the number of phonemes per word within manageable limits for the children).

were being accessed. Not only were onset deletions performed best (54% correct), the preferred type of error was to give a rhyming word. Coda deletions proved to be difficult for the children (33% correct overall, but only 12% correct for the children) but not for the High School students (94% *vs* 89% correct for onset deletions). Incomplete unit deletions (i.e., part of an onset or coda) were performed relatively poorly by all of the subjects (21% correct overall). Although one limitation of the design of this experiment was that the three deletion lists were always presented in the same order, it does not seem likely that this was the cause of the low scores for the incomplete unit deletions, since a replication using random order of presentation with a different group of High School students showed the same pattern of difficulty.

For a given type of deletion there was no difference in performance between single phoneme and cluster deletions for the children, although High School students performed better with single segment deletions (92% *vs* 72%). Cluster onset and incomplete unit deletions seemed to cause problems for these students, perhaps because of a word strategy, since in most of these cases a word resulted either when the correct, cluster deletion was made or when an incorrect, single segment deletion was made. This result provides evidence that phonemes are readily accessible units to these older, literate students, in contrast to the primacy of sub-syllabic units (that is, onset and rhyme) for children.

To investigate the complete range of possible sub-syllabic units (i.e., onset, nucleus, coda, head, rhyme, margins) for adults, a further variant of the analogy technique was employed, called substitution-by-analogy. The results were scored in terms of accuracy, response latencies, and types of error. Rhymes and onsets were the easiest/fastest to manipulate (77%, 1.8 sec. and 65%, 2.1 sec., respectively), nuclei and codas next (59%, 2.6 sec. and 47%, 3.1 sec., respectively), and margins and heads by far the most difficult/slowest (26%, 4.1 sec. and 20%, 4.6 sec., respectively). There was a strong preference for errors involving changes in the onsets or rhymes with very few errors of the other types (coda, nucleus, margins, head, other changes). Despite controls on the number of phon[illegible]ed in substitutions (e.g., rhyme *vs* head substitutions), the[illegible]dvantage for the sub-syllabic units of rhyme, onset, nucleus, a[illegible]the head and margins units. This result [illegible] be explained by assuming that only phonemes are real to liter[illegible]ers.

Although the finding that o[illegible]nd-rhyme are accessible to speakers is not new (see Treiman, 1983; 1985a, 1986), the present study investigated six different hypothesized sub-syllabic units in the same task with the same speakers, which allowed for direct comparisons. The onset, rhyme, nucleus, and coda emerged unambiguously h[illegible]as sub-syllabic units (cf. Treiman, 1983; but see also Treiman, 1984) over head and margins units, which were relatively inaccessible to manipulation. The differential difficulty in manipulating the rhyme, onset, nucleus, and coda allows us to make some inferences about the structure of the syllable. Onsets and rhymes were best overall, which suggests a major branch or weak bond between the two. Performance on vowel nuclei was generally better than on codas. While this is perhaps not surprising (given the status of the vowel as the syllable peak and carrier of stress), it does suggest that a simple tree structure for the syllable with equally-weighted end nodes of nucleus and coda would have to be adjusted to give more weight to the former.

The first three experiments provide evidence that sub-syllabic units are accessible to various manipulations (i.e., counting, deletion, substitution), but they do not unequivocally provide evidence that these units are involved in language perception. The phonetic similarity judgment task (Experiment 4) confirmed the psychological reality of onset, nucleus, and coda for perceptual judgments. Predictions for similarity judgments using a model of the syllable with these units were consistently better than a model based on a string of independent segments or a simple onset-rhyme structure. Changes in the onsets and codas were roughly equivalent, whereas changes in the nucleus had much greater effects, suggesting a syllable structure with similarly weighted onsets and codas and more heavily weighted vowel nuclei. Changing post-vocalic liquids had a much greater effect on judgments than changing other post-vocalic consonants and a similar effect to changing the vowel, confirming that liquids are treated as part of the vowel nucleus. A refinement of the notion of nucleus emerged in that post-vocalic nasals did not seem to be treated differently from other non-liquid consonants.

5.2 Models of syllable structure

The models of syllable structure that can be definitively rejected based on the results of these tasks are the models of syllables as strings of equally weighted phonemes and as indivisible units. Despite controls on the number of phonemes involved in substitutions (Experiment 3) and changes (Experiment 4), there were significant differences in the treatment of sub-syllabic units. The results from the first and second experiments also support the primacy of sub-syllabic units over phonemes, as the children were more successful at counting and deleting at least some of these. By the same token, if syllables were treated as non-analysable wholes, then one would expect that speakers would be completely unable to count, delete, or substitute sub-syllabic units in the way that they did in these tasks, and that changes in various positions of the syllable would have similar effects on phonetic similarity judgments, yet large differences were found for different syllable position and for different units.

The results do not support the left branching hierarchical model (Iverson & Wheeler, 1987) nor Vennemann's (in press) description of a head-coda structure, since there was no evidence for a head unit. In a direct comparison of the manipulability of different units (the substitution-by-analogy task), the rhyme was significantly more accessible than the head, suggesting that the former has psychological reality to English speakers. Little evidence was found to support the core+affix model proposed by Fujimura (1975; 1976) and Fujimura and Lovins (1978).

The psychological reality of both onset and rhyme units received strong support from the results of the present experiments, since these two units seemed very accessible to children (counting, Experiment 1; deletion, Experiment 2) and were most readily manipulated by adults (Experiment 3). Caution should be exercised before accepting a simple onset-rhyme model of syllable structure, however, since the nucleus and coda also emerged as psychologically viable sub-syllabic units in the experiments. Claims for a psychological model of syllable structure based solely on the results of an advantage for onset and rhyme is subject to the criticism of cultural influence, since alliteration (onset reduplication) and rhyme are familiar in English verse and poetry.

Why alliteration and rhyme are used so frequently as poetic devices in

English in the first place, however, is unclear, unless the poetic tradition is seen as exploiting the units that the language provides. It appears that there are three possibilities:

1) Onset and rhyme are used in English quite by chance and their advantage in the present experiments was simply a reflection of this prior learning/experience. In this case, testing people completely naive with respect to English poetry, nursery rhymes, songs, advertising, etc. would be required. (It would be difficult to find such individuals.)
2) The use of onset and rhyme in English (for whatever reason) imposes a structure on English syllables. It may be that accessing this imposed structure resulted in the advantage for onset and rhyme, but then one would expect this advantage in production-oriented tasks but not in less direct tasks.[2]
3) Onset and rhyme are part of the more complex syllable structure which forms an integral part of phonological knowledge and its influence becomes apparent in tasks which require sub-syllabic analysis, such as those used in the present study.

The syllabic model of onset-vowel/nucleus-coda as equally weighted sub-syllabic units is also too simplistic to explain the results found in the present tasks. This model ignores the importance of rhyme, especially for the young children, who demonstrated little awareness of codas (i.e., sub-syllabic counting scores were best accounted for with an onset-rhyme analysis, Experiment 1; deletion of codas was very poor, Experiment 2). Further, although onset and coda were sometimes treated as though they

[2]An example of a less direct task is the phonetic similarity judgment (PSJ) task which was used in Experiment 4.

had equal importance or weight in the syllable (i.e., the high school students performed equally well on both types of deletions, Experiment 2; phonetic similarity judgments were affected equally by changes in these units, Experiment 4), there were also differences (i.e., onset substitutions were much easier than coda substitutions, Experiment 3).

To be consistent with the present results, a model of syllable structure would have to include provision for onset and rhyme, as well as nucleus and coda. A simple vowel structure for the nucleus did not account for the treatment of post-vocalic glides (/y/, /w/) and /r/ in the substitution task, nor for the effect that changes in these elements had on phonetic similarity judgments. Although the question of what the nucleus consists of remains unresolved (see next section for a discussion), it is clear that it can be more complex than a simple vowel.

The right-branching model most often discussed for English syllables consists of onset and rhyme on one level and vowel/nucleus and coda as sub-constituents of the rhyme. This model would have to be refined somewhat to be consistent with the results reported here; the onset and coda would be similar in weight, with a heavier nucleus (as indicated by the results of the phonetic similarity judgment task). Nothing in the results compels a tree structure, such as that discussed by Selkirk (1984), and the inconclusiveness of the tests for the reality of a separate appendix might be taken as an indication that there are some problems with the metrical model of the syllable.

The model of syllabic structure consisting of bonds and affinities proposed by Vennemann may be a viable alternative to a hierarchical model. Differential affinities would account for preference of units, as discussed in Experiment 3, above. It would be necessary to formalize the predictions from this model in order to adequately assess it. The substitution-by-analogy technique could be used with a variety of post-vocalic consonants to determine if different affinities could explain the differences already noted among post-vocalic /r/, /l/, and nasals. The present results most strongly support the notion of a major branch or weak (easily breakable) bond between the onset and rhyme, with a minor branch or bond between the nucleus and coda, but cannot discriminate between a hierarchical syllable structure and a flat structure based on bonds.

5.3 Specific questions

The results from the four experiments support the no[illegible]a heavily weighted syllable nucleus and provides some evidence as to the membership of the nucleus. Post-vocalic liquids were generally treated as part of the nucleus and pre-vocalic liquids were always part of the onset. The vowel-/r/ sequences were clearly units although the vowel-/l/ sequences did not seem so strongly bound. The evidence from the phonetic similarity task suggests that nasals are not treated as part of the vowel (note that Treiman, 1984, also found differential effects for post-vocalic liquids and nasals). Post-vocalic glides were clearly part of the vowel nucleus (i.e., vowel diphthongs) and pre-vocalic glides were generally treated as part of the onset. One way to explain these results is in terms of differential affinities such as those discussed by Vennemann. Another is by suggesting that post-vocalic glides and /r/ are actually part of the core nucleus and that /l/ and the nasals are treated as part of the complex nucleus if the core is simple (i.e., just a vowel) and as part of the coda if the core is complex.

The status of initial /s/+stop clusters was not resolved because there was some tendency to treat the /s/+stop clusters as separate units in three-phoneme onsets (e.g., /str/, /spl/) in the deletion-by-analogy task, suggesting that these are units within the onset.

A third question which has not been resolved by the results of the present experiments is the generalizability of the results to multisyllabic words. It is possible that sub-syllabic units emerge in the analysis of monosyllabic words but that some other units (e.g., syllables, phonemes) would emerge with more complex words. The problem of ambisyllabicity also has not been addressed. The division of syllables into onsets, nuclei (which is also problematic; see the discussion above), and codas becomes much more complicated with ambisyllabic consonants (e.g., the division between the first syllable's coda and the second syllable's onset in a word like Boston; example from Kahn, 1976).

5.4 Sub-syllabic units in languages other than English

In order to truly evaluate whether the evidence for sub-syllabic units found in the present study confirms their psychological reality or whether they are artifacts of other knowledge (such as knowledge of spelling or poetry), the techniques used in this study would have to be extended to a variety of languages with different orthographic and poetic traditions. It would not be surprising if different languages had different syllable-internal structure, and one would predict that different units might be preferentially accessible based on the particular phonotactic properties of the language. Because of the suggestion that orthography may be related to sound representation, Japanese (which has a syllabary) and Mandarin and Cantonese (with character systems) would be interesting languages to study. There is another language (more accessible to the present author) which has no standard orthography at all.

Swiss German is the cover term for German dialects in Switzerland. The Swiss speak their dialect (be it Bärndütsch, Züritüütsch, Baseldytsch, etc.) at home, with friends, at work, socializing, and so on, but use a standard form of German (Schriftdeutsch) for lectures, formal meetings, church services, and all written communication. Although the dialect is sometimes seen written (e.g., in ads, letters to friends, quotes in the newspaper), there is no standard written form. Swiss German differs from standard German in the conjugation of the verbs, the form of verb tenses, some pronouns, cases, morphological affixes, lexical items, and most noticeably, pronunciation. Although Keller (1979) included Swiss German in his book of German dialects, he consistently referred to it as a language, not a dialect. Although some southern German dialects are similar to Swiss German and may even be mutually intelligible, the Swiss tend to consider their language to be quite distinct from German. Germans from other than the southern area do not seem to be able to understand Swiss German without considerable exposure to it and Swiss German people automatically switch to standard German when speaking to Germans (or, indeed, to any foreigner).

It is therefore interesting to study phonological questions in Swiss German, with less fear of orthographic influence. There may be influence of German spelling-sound correspondences, but at least there is no standardized form to guide judgments. Not only is there no standard Swiss German orthography, a significant proportion of the Swiss German vocabulary is unique, with no apparent relationship to German roots or lexical items. If orthography was influencing phonemic perceptions, presumably it could only do so if there was some representation available, such as German spelling for words related to German. For uniquely Swiss words, with no standard spelling, it is unlikely an orthographic representation is already stored.[3] It would therefore seem a good language with which to assess the reasonableness of the model of sound representation. Can Swiss German speakers give reliable counts and make similarity judgments for words of their language that relate to the predictions of sub-syllabic and phoneme models in the same way English speakers can?

A group of university students (n=16) were asked to note the number of speech sounds ('Luutelemente') in 25 uniquely Swiss and 25 German-related words and then (in a second presentation of the 50 words), write the words as if they were writing them in a letter to a friend. The correlation between the mean number provided by the students and the number with a segmental phoneme representation[4] was extremely high (r=.97), as was the number correct scored with the segmental representation counts (82%). Despite the high scores and very consistent spellings, there were some problem areas. Affricates (e.g., /kχ/, /tʃ/, /pf/, /ts/) and aspirated stops (which are unusual in the language, but do occur word-initially) were treated as single segments in the segmental representation count, but the students were split in terms of whether these were one segment or two.

---

[3] Swiss German speakers seem very aware of unique Swiss words. In learning the language and requesting repetitions or explanations of certain items, I was invariably met with laughter and comments to the effect that the item was very unusual or strange because there was no German cognate. Some of these items are very frequently used (e.g., [bɔʃtʌ] to shop or [tœf] motorcycle).

[4] The transcriptions were done by the experimenter and confirmed by Dr. Spörri of the Phonetisches Labor, Universität Zürich.

The influence of standard German letter-sound correspondences could not be blamed directly, since two of the affricates were generally represented by single letters (/kχ/=k, /ts/=z), one with two letters (/pf/=pf), and one with four letters (/tʃ/=tsch), and the velar nasal was unambiguously treated as a single segment, although it was represented with two letters in the spellings (i.e., ng).

To assess the accessibility of units intermediate between phonemes and syllables, a pilot study was carried out, in which Zürich high school students were asked to note how many chunks of sound ('Luutgruppe') monosyllabic Swiss German words contained. They were first provided with an example word split in a variety of ways (e.g., /sprɪnts/ [a type of cheese] could be divided into two, /spr-ɪnts/, /sp-rɪnts/, /sprɪn-ts/; three, /sp-rɪn-ts/, /spr-ɪnt-s/, /s-prɪ-nts/; or into four /s-pr-ɪn-ts/, /sp-rɪ-nt-s/, /sp-rɪn-t-s/). Generally, words without onsets (e.g., /arm/) or without codas (e.g., /χli/) received counts of two, and words with both onsets and codas (e.g., /kʃaft/) received counts of three. Interestingly, there was evidence that post-vocalic liquids, nasals, and single obstruents (i.e., not clusters) were treated as part of the nucleus. Although diphthongs were counted as two segments by the students in the phoneme counting study, they were treated as single units here.

The unique status of the nucleus also emerged in a phonetic similarity judgment task with the same students, as did a difference between the onset and the coda. Changes in the onset (e.g., /gœrps:dœrps/, mean=4.4 on a scale from 0 to 6) had less effect than changes in the coda (e.g., /gœrps:gœrks/, mean=3.9) which, in turn, had less effect on judgments than changes in the post-vocalic liquid (e.g., /gœrps:gœlps/, mean=2.8). This general pattern held for all 54 word pairs studied, but the findings cannot be taken as conclusive, as not all possible comparisons were made.

There are many types of phonological questions which could be addressed using the techniques of the present study. The apparent saliency of the nucleus and coda in Swiss German contrast somewhat to the importance of rhyme in English. Interestingly, semi-rhyme (where the vocalic element is manipulated without changing the coda or vice versa) is often used in Swiss German for children's expressions and songs (e.g., /bif baf buf und du bɪʃ e:r und redliχ dus/, similar to the English counting verse /ini mini mayni mo/), which may be the result of syllable structure and the cause of the differential results in the phonetic similarity judgments. The more direct technique of substitution-by-analogy could be used to assess the relative manipulability of sub-syllabic units in Swiss German to contrast them with, for example, those in English. From the importance of vowels in German and Swiss German morphology and evidence from pilot studies, the prediction can be made that the nucleus will be easily manipulated as will its complement, the margins (i.e., substituting margins can also be thought of as leaving the vowel and vice versa).

Vennemann's (in press) model of the syllable as a head and coda was initially applied to German and may indeed be more appropriate than a right-branching (onset-oriented) model for languages closely allied to German. Finding that units other than onset and rhyme are not real to speakers of languages other than English would not contradict the claims of the present study at all. The psychological reality of sub-syllabic units in English has been established and it remains to determine which units are real for various languages, rather than assuming that only phonemes and syllables are relevant to the description of language sounds.

References

Alegria, J., Pignot, E., & Morais, J. Phonetic analysis of speech and memory codes in beginning readers. *Memory & Cognition*, 1982, *10*, 451-456.

Anderson, J. Syllabic or non-syllabic phonology? *Journal of Linguistics*, 1969, *5*, 136-142.

Anderson, J., & Jones, C. Three theses concerning phonological representations. *Journal of Linguistics*, 1974, *10*, 1-26.

Anderson, J.M. & Jones, C. *Phonological Structure and the History of English*. Amsterdam: North-Holland Pub. Co., 1977.

Bailey, C.-J.N. *Gradience in English syllabization and a revised concept of unmarked syllabization*. Bloomington: Indiana University Linguistics Club, 1978.

Baron, J., Treiman, R., Wilf, J.F., & Kellman, P. Spelling and reading by rules. In: U. Frith (Ed.), 1980, 159-194.

Barton, D., Miller, R., & Macken, M.A. Do children treat clusters as one unit or two? *Papers and Reports on Child Language Development*, 1980, *18*, 105-137.

Basbøll, H. Some conditioning phonological factors for the pronunciation of short vowels in Danish with special reference to syllabification. *ARIPUC*, 1972, *6*, 185-210.

Basbøll, H. Notes on Danish consonant combinations. *ARIPUC*, 1973, *7*, 103-142.

Basbøll, H. The phonological syllable with special reference to Danish. *ARIPUC*, 1974, *8*, 39-131.

Bell, A. & Hooper, J.B. (Eds.) *Syllables and Segments*. Amsterdam: North-Holland Pub. Co., 1978.

Bell, A. & Hooper, J.B. Issues and evidence in syllabic phonology. In: A. Bell & J.B. Hooper (Eds.), 1978a, 3-22.

Blachman, B.A. Relationship of rapid naming ability and language analysis skills to kindergarten and first-grade reading achievement. *Journal of Educational Psychology*, 1984, *76*, 610-622.

Bondarko, L.V. The syllable structure of speech and distinctive features of phonemes. *Phonetica*, 1969, *20*, 1-40.

Bradley, L. & Bryant, P.E. Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, 1978, *271*, 746-747.

Bradley, L. & Bryant, P.E. Categorizing sounds and learning to read - a causal connection. *Nature*, 1983, *301*, 419-421.

Browman, C.P. Tips of the tongue and slips of the ear. Implications for language processing. *UCLA Working Papers in Phonetics*, 1978, *42*.

Browman, C.P. Perceptual processing: Evidence from slips of the ear. In: V.A. Fromkin (Ed.), 1980, 213-230.

Bruce, D.J. The analysis of word sounds by young children. *British Journal of Educational Psychology*, 1964, *34*, 158-170.

Buckingham, H.W. Jr. On correlating aphasic errors with slips-of-the-tongue. *Applied Psycholinguistics*, 1980, *1*, 199-220.

Cairns, C.E. & Feinstein, M.H. Markedness and the theory of syllable structure. *Linguistic Inquiry*, 1982, *13*, 193-225.

Carroll, J.B., Davies, P., & Richman, B. *Word Frequency Book*. New York: American Heritage Pub. Co. Ltd., 1971.

Chomksy, C. Invented spelling in the open classroom. *Word*, 1975, *27*, 499-518.

Chomsky, N. & Halle, M. *The Sound Pattern of English*. New York: Harper & Row, 1968.

Clements, G.N. & Keyser, S.J. *CV Phonology. A Generative Theory of the Syllable*. Cambridge, Mass.: MIT Press, 1983.

Conrad, R. Acoustic confusions in immediate memory. *British Journal of Psychology*, 1964, *55*, 75-84.

Conrad, R. Speech and reading. In: J.F. Kavanagh & I.G. Mattingly (Eds.), 1972, 205-240.

Content, A., Kolinsky, R., Morais, J., & Bertelson, P. Phonetic segmentation in prereaders: effect of corrective information. *Journal of Experimental Child Psychology*, 1986, *42*, 49-72.

Cooper, L.A. Strategies for visual comparison and representation: Individual differences. In: F.J. Sternberg (Ed.). *Advances in the Psychology of Human Intelligence*. Hillsdale, N.J.: Laurence Erlbaum Assoc., Pub., 1982, 77-124.

Cowan, N. & Leavitt, L.A. Talking backward: exceptional speech play in late childhood. *Journal of Child Language*, 1982, *9*, 481-495.

Crowder, R.G. & Morton, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, *5*, 365-373.

Darwin, C.J. & Baddeley, A.D. Acoustic memory and the perception of speech. *Cognitive Psychology*, 1974, *6*, 41-60.

Davidsen-Nielsen, N. Syllabification in English words with medial sp, st, sk. *Journal of Phonetics*, 1974, *2*, 15-45.

Derwing, B.L. *Transformational Grammar as a Theory of Language Acquisition*. Cambridge: Cambridge University Press, 1973.

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

Derwing, B.L. Methodological issues in experimental phonology. Talk presented to the BYU Student Linguistic Society, Brigham Young University, Provo, Utah, February, 1985.

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

Derwing, B.L. & Nearey, T.M. Experimental phonology at the University of Alberta. In: J.J. Ohala & J.J. Jaeger (Eds.), 1986, 187-209.

Derwing, B.L., Nearey, T.M., & Dow, M.L. On the phoneme as the unit of the 'second articulation'. *Phonology Yearbook*, 1986, *3*, 45-69.

Dinnsen, D.A. A re-examination of phonological neutralization. *Journal of Linguistics*, 1985, *21*, 265-279.

Dow, M.L. On the role of orthography in experimental phonology. Msc thesis, University of Alberta, 1981.

Downing, J. (Ed.). *Comparative Reading. Cross-National Studies of Behavior and Processes in Reading and Writing.* New York: The Macmillan Co., 1973.

Ehri, L.C. Learning to read and spell. Invited address presented at the American Educational Research Association, Chicago, April 1, 1985.

Ehri, L.C. & Wilce, L.S. The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology*, 1979, *71*, 26-40.

Ehri, L.C. & Wilce, L.S. The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, 1980, *1*, 371-385.

Ehri, L.C. & Wilce, L.S. Cipher versus cue reading: An experiment in decoding acquisition. *Journal of Educational Psychology*, 1987a, *79*, 3-13.

Ehri, L.C. & Wilce, L.S. Does learning to spell help beginners learn to read words? *Reading Research Quarterly*, 1987b, *22*, 47-65.

Ehri, L.C. & Wilce, L.S. The influence of spellings on speech: Are alveolar flaps /d/ or/t/? In: D. Yaden & S. Templeton (Eds.) *Metalinguistic Awareness and Beginning Literacy*. Exeter, N.H.: Heinemann Educational Books, Inc., in press.

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

Fink, R. Orthography and the perception of stops after s. *Language and Speech*, 1974, *17*, 152-159.

Foss, D.J., Harwood, D.A. & Blank, M.A. Deciphering decoding decisions: Data and devices. In: R.A. Cole (Ed.), *Perception and Production of Fluent Speech.* Hillsdale, N.J.: Laurence Erlbaum Assoc., Pub., 1980, 165-199.

Foss, D.J. & Swinney, D.A. On the psychological reality of the phoneme: Perception, identification, and consciousness. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 246-257.

Fourakis, M. Should neutralization be redefined? *Journal of Phonetics*, 1984, *12*, 291-296.

Fourakis, M. & Iverson, G.K. On the 'incomplete neutralization' of German final obstruents. *Phonetica*, 1984, *41*, 140-149.

Fox, B. & Routh, D.K. Analysing spoken language into words, syllables, and phonemes: A developmental study. *Journal of Psycholinguistic Research*, 1975, *4*, 331-342.

Fox, B. & Routh, D.K. Phonemic analysis and severe reading disability in children. *Journal of Psycholinguistic Research*, 1980, *9*, 115-119.

Fox, B. & Routh, D.K. Phonemic analysis and synthesis as word attack skills: revisited. *Journal of Educational Psychology*, 1984, *76*, 1059-1064.

Frith, U. *Cognitive Processes in Spelling*. New York: Academic Press, 1980.

Fromkin, V.A. Neuro-muscular specification of linguistic units. *Language and Speech*, 1966, *9*, 170-199.

Fromkin, V.A. (Ed.) *Errors in Linguistic Performance. Slips of the tongue, ear, pen, and hand.* New York: Academic Press, 1980.

Fry, D.B. The function of the syllable. *Zeitschrift für Phonetik*, 1964, *17*, 215-221.

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

Fujimura, O. Syllables as a unit of speech recognition. *IEEE Transactions - Acoustic Speech and Signal Processing*, 1975, *23*, 82-87.

Fujimura, O. Syllables as concatenated demisyllables and affixes. *Journal of the Acoustical Society of America*, 1976, *59* (supp), S55.

Fujimura, O. & Lovins, J.B. Syllables as concatenative phonetic units. In: A. Bell & J.B. Hooper (Eds.), 1978, 107-120.

Garnes, S. & Bond, Z. A slip of the ear: A snip of the ear? A slip of the year? In: V.A. Fromkin (Ed.), 1980, 231-239.

Gelb, I.J. *A Study of Writing*. Chicago: The University of Chicago Press, 1952.

Giegerich, H.J. *Metrical Phonology and Phonological Structure. German and English*. Cambridge: Cambridge University Press, 1985.

Gleitman, L.R. & Rozin, P. Teaching reading by use of a syllabary. *Reading Research Quarterly*, 1973, *8*, 447-483.

Goldsmith, J. *Autosegmental Phonology*. Bloomington: Indiana University Linguistics Club, 1976.

Golinkoff, R.M. Critique: phonemic awareness skills and reading achievement. In: F.B. Murray & J.J. Pikulski (Eds.) *The Acquisition of Reading. Cognitive, Linguistic and Perceptual Prerequisites*, Baltimore: University Park Press, 1978, 23-41.

Gresser, J.Y. & Mercier, G. Automatic segmentation of speech into syllabic and phonemic units: Application to French words and utterances. In: G. Fant & M.A.A. Tatham (Eds.) *Auditory Analysis and Perception of Speech*. New York: Academic Press, 1975, 359-382.

Halle, M. & Vergnaud, J.-R. Three dimensional phonology. *Journal of Linguistic Research*, 1980, *1*, 83-105.

Hansen, D. & Rodgers, T. An exploration of psycholinguistic units in initial reading. In: K.S. Goodman (Ed.) *The Psycholinguistic Nature of the Reading Process*. Detroit: Wayne State University Press, 1968, 61-101.

Harrigan, J.E. Initial reading instruction: Phonemes, syllables or ideographs? *Journal of Learning Disabilities*, 1976, *9*, 74-80.

Hayes, B. *A Metrical Theory of Stress Rules*. Bloomington: Indiana University Linguistics Club, 1981.

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

Hoard, J.E. Juncture and syllable structure in English. *Phonetica*, 1966, *15*, 96-109.

Hoard, J.E. Aspiration, tenseness, and syllabication in English. *Language*, 1971, *47*, 133-140.

Hockett, C.F. A manual of phonology. *International Journal of American Linguists*, 1955, *21* (4).

Hockett, C.F. *A Course in Modern Linguistics*. New York: The MacMillan Co., 1958.

Hofmann, T.R. Initial clusters in English. *MIT-RLE-QPR*, 1966, *84*, 263-274.

Hohn, W.E. & Ehri, L.C. Do alphabet letters help prereaders acquire phonemic segmentation skill? *Journal of Education Psychology*, 1983, *75*, 752-762.

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

Householder, F.W. *Linguistic Speculations*. London: Cambridge University Press, 1971.

Hultzén, L.S. Consonant clusters in English. *American Speech*, 1965, *40*, 5-19.

Iverson, G.K. & Wheeler, D.W. Phonological categories and constituents. Paper presented at the Conference on Linguistic Categorization. 16th Annual UW-Milwaukee Linguistics Symposium, April 10-11, 1987.

Jaeger, J.J. Categorization in phonology: An experimental approach. PhD dissertation, University of California at Berkely, 1980a.

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

Jaeger, J.J. On the acquisiton of abstract representations for English vowels. *Phonology Yearbook*, 1986, *3*, 71-97.

Jastak, J.F., Bijou, S.W. & Jastak, S. *Wide Range Achievement Test*. Wilmington: Jastak Assoc., 1978.

Jones, C. Some constraints on medial consonant clusters. *Language*, 1976, *52*, 121-130.

Kahn, D. *Syllable-based Generalizations in English Phonology*. Bloomington: Indiana University Linguistics Club, 1976.

Kavanagh, J.F. & Mattingly, I.G. (Eds.) *Language by Ear and by Eye*. Cambridge, Mass.: The MIT Press, 1972.

Kavanagh, J.F. & R.L. Venezky (Eds.) *Orthography, Reading and Dyslexia*. Baltimore: University Park Press, 1980.

Keller, R.E. *German Dialects. Phonology and morphology and selected readings*. Manchester: University of Manchester Press, 1979.

Kerek, A. The phonological relevance of spelling pronunciation. *Visible Language*, 1976, *10*, 323-338.

Kiparsky, P. Remarks on the metrical structure of the syllable. In: W. Dressler, O. Pfeiffer, & J. Rennison (Eds.) *Phonologica 1980 Akten der Vierten Internationalen Phonologie-Tagung*. Innsbruck: Innsbruck Beiträge zur Sprachwissenschaft, 1981, 245-256.

Kloster Jensen, M. von Die Silbe in der Phonetik und Phonemik. *Phonetica*, 1963, *9*, 17-30.

Kohler, K.J. Is the syllable a phonological universal? *Journal of Linguistics*, 1966, *2*, 207-208.

Kramer, A.F. & Donchin, E. Brain potentials as indices of orthographic and phonological interaction during word matching. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 1987, *13*, 76-86.

Krámský, J. *The Phoneme. Introduction to the History and Theories of the Concept*. München: Wilhelm Fink, 1974.

Levitt, J. The influence of orthography on phonology: a comparative study (English, French, Spanish, Italian, German). *Linguistics*, 1978, *208*, 43-67.

Liberman, I.Y., Shankweiler, D., Fischer, F.W., & Carter, B. Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 1974, *18*, 201-212.

Liberman, M. The intonational system of English. PhD dissertation, MIT, Cambridge, Mass., 1975.

Liberman, M. & Prince, A. On stress and linguistic rhythm. *Linguistic Inquiry*, 1977, *8*, 249-336.

Linell, P. *The Written Language Bias in Linguistics*. Linköping: University of Linköping, 1982.

Lowenstamm, J. On the maximal cluster approach to syllable structure. *Linguistic Inquiry*, 1981, *12*, 575-604.

Lüdtke, H. Die Alphabetschrift und das Problem der Lautsegmentierung. *Phonetica*, 1969, *20*, 147-176.

MacKay, D.G. The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 1972, *3*, 210-227.

MacKay, D.G. Speech errors inside the syllable. In: A. Bell & J.B. Hooper (Eds.), 1978, 201-212.

Malmberg, B. The phonetic basis for syllable division. *Studia Linguistica*, 1955, *9*, 80-87.

Mann, V.A. Phonological awareness: the role of reading experience. *Cognition*, 1986, *24*, 65-91.

Marcel, T. Phonological awareness and phonological representation. In: U. Frith (Ed.), 1980, 373-403.

McCarthy, J.J. On stress and syllabification. *Linguistic Inquiry*, 1979, *10*, 443-465.

McCawley, J.D. *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton, 1968.

McNeil, J.D. & Stone, J. Note on teaching children to hear separate sounds in spoken words. *Journal of Educational Psychology*, 1965, *56*, 13-15.

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

Mehler, J., Sequi, J., & Frauenfelder, U. The role of the syllable in language acquisition and perception. In: T. Myers, J. Laver, & J. Anderson (Eds.), 1981, 295-305.

Menn, L. Phonological units in beginning speech. In: A. Bell & J.B. Hooper (Eds.), 1978, 157-171.

Mitleb, F. Segmental and non-segmental structure in phonetics: evidence from foreign accent. PhD dissertation, Indiana University, Bloomington, 1981.

Morais, J. Literacy and awareness of the units of speech: implications for research on the units of perception. *Linguistics*, 1985, *23*, 707-721.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. Literacy training and speech segmentation. *Cognition*, 1986, *24*, 45-64.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 1979, *7*, 323-331.

Moskowitz, B.A. On the status of vowel shift in English. In: T.E. Moore (Ed.) *Cognitive Development and the Acquisition of Language*. New York: Academic Press, 1973, 223-260.

Myers, J.L. *Fundamentals of Experimental Design*. Toronto: Allyn & Bacon, Inc., 1979.

Myers, T., Laver, J., & Anderson, J. *The Cognitive Representation of Speech*. Amsterdam: North-Holland Pub., Co., 1981.

Nearey, T.M. The psychological reality of phonological representations. In: T. Myers, J. Laver, & J. Anderson (Eds.), 1981, 359-377.

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

O'Connor, J.D. & Trim. J.L.M. Vowel, consonant, and syllable - a phonological definition. *Word*, 1953, *9*, 103-122.

O'Dell, M. & Port, R. Discrimination of word final voicing in German. *Journal of the Acoustical Society of America*, 1983, *73* (supp), S31.

Ohala, J.J. & Jaeger, J.J. *Experimental Phonology*. New York: Academic Press, 1986.

Paul, R. Invented spelling in kindergarten. *Young Children*, 1976, *31*, 195-200.

Perin, D. Phonemic segmentation and spelling. *British Journal of Psychology*, 1983, *74*, 129-144.

Pike, K.L. *Phonemics: A Technique for Reducing Languages to Writing.* Ann Arbor: University of Michigan Press, 1947.

Pike, K.L. *Language in Relation to a Unified Theory of the Structure of Human Behavior.* The Hague: Mouton & Co., 1967.

Pilch, H. Phonemic constituent analysis. *Phonetica*, 1966, *14*, 237-252.

Port, R., Mitleb, F., & O'Dell, M. Neutralization of obstruent voicing in German is incomplete. *Journal of the Acoustical Society of America*, 1981, *70* (supp 1), S13.

Pulgram, E. Syllable, word, nexus, cursus. The Hague: Mouton, *Janua Linguarum* 1970, *81*.

Rack, J.P. Orthographic and phonetic coding in developmental dyslexia. *British Journal of Psychology*, 1985, *76*, 325-340.

Read, C. Pre-school children's knowledge of English phonology. *Harvard Educational Review*, 1971, *41*, 1-34.

Read, C. Children's judgments of phonetic similarities in relation to English spelling. *Language Learning*, 1973, *23*, 17-38.

Read, C., Yun-Fei, Z., Hong-Yin, N., & Bao-Qing, D. The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition*, 1986, *24*, 31-44.

Reber, A.S. & Scarborough, D.L. (Eds.) *Toward a Psychology of Reading.* Hillsdale, N.J.: Laurence Erlbaum Assoc., Pub., 1977.

Rosner, J. & Simon, D.P. The auditory analysis test: An initial report. *Journal of Learning Disabilities*, 1971, *4*, 384-392.

Rozin, P., Bressman, B., & Taft, M. Do children understand the basic relationship between speech and writing? The mow-motorcycle test. *Journal of Reading Behavior*, 1974, *6*, 327-334.

Rozin, P. & Gleitman, L.R. The structure and acquisition of reading II: The reading process and the acquisiton of the alphabetic principle. In: A.S. Reber & D.L. Scarborough (Eds.), 1977, 55-141.

Rozin, P., Poritsky, S., & Sotsky, R. American children with reading problems can easily learn to read English represented by Chinese characters. *Science*, 1971, *171*, 1264-1267.

Ruske, G. & Schotola, T. An approach to speech recognition using syllabic decision units. *IEEE Transactions - Acoustic Speech and Signal Processing*, 1978, *26*, 722-725.

Sakamoto, T. Reading of Hiragana. In: J.F. Kavanagh & R.L. Venezky (Eds.), 1980, 15-24.

Sakamoto, T. & Makita, K. Japan. In: J. Downing (Ed.), 1973, 440-465.
Santa, J.L., Santa, C., & Smith, E.E. Units of word recognition: Evidence for the use of multiple units. *Perception & Psychophysics*, 1977, *22*, 585-591.
Sapir, E. The psychological reality of phonemes (1933). Reprinted in: D.G. Mandelbaum (Ed.) *Selected Writings of Edward Sapir*. Los Angeles: University of California Press, 1951, 46-60.
Savin, H.B. & Bever, T.G. The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 295-302.
Scholes, R.J. Syllable segmentation and identification in American English. *Linguistics*, 1967, *36*, 55-77.
Seidenberg, M.S. & Tanenhaus, M.K. Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, *5*, 546-554.
Selkirk, E.O. The role of prosodic categories in English word stress. *Linguistic Inquiry*, 1980, *11*, 563-605.
Selkirk, E.O. On the major class features and syllable theory. In: M. Aronoff & R.T. Oehrle (Eds.) *Language Sound Structure*. Cambridge, Mass.: The MIT Press, 1984, 107-136.
Shattuck-Hufnagel, S. Sublexical units and suprasegmental structure in speech production planning. In: P.F. MacNeilage (Ed.) *The Production of Speech*. New York: Springer-Verlag, 1983, 109-136.
Shattuck-Hufnagel, S. The representation of phonological information during speech production planning: evidence from vowel errors in spontaneous speech. *Phonology Yearbook*, 1986, *3*, 117-149.
Singh, S., Woods, D.R., & Becker, G.M. Perceptual structure of 22 prevocalic English consonants. *Journal of the Acoustical Society of America*, 1972, *52*(6), 1698-1713.
Skjelfjord, V.J. Teaching children to segment spoken words as an aid in learning to read. *Journal of Learning Disabilities*, 1976, *9*, 297-306.
Skousen, R. English spelling and phonemic representation. *Visible Language*, 1982, *17*, 28-38.
Smith, I.R. Abstractness in phonology: The English velar nasal. *Linguistics*, 1982a, *20*, 391-409.
Smith, I.R. Morphological operations, historical evidence and English [ŋ]. *Proceedings of the XIIIth International Congress of Linguists*, 1982b, 572-575.
Spoehr, K.T. & Smith, E.E. The role of syllables in perceptual processing. *Cognitive Psychology*, 1973, *5*, 71-89.

Stemberger, J.P. The nature of /r/ and /l/ in English: evidence from speech errors. *Journal of Phonetics*, 1983, *11*, 139-147.

Stemberger, J.P. & Treiman, R. The internal structure of word-initial consonant clusters. *Journal of Memory and Language*, 1986, *25*, 163-180.

Stroop, J.R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 1935, *18*, 643-661.

Taft, M. & Hambly, G. The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language*, 1985, *24*, 320-335.

Tanenhaus, M.K., Flanigan, H.P., & Seidenberg, M.S. Orthographic and phonological activation in auditory and visual word recognition. *Memory & Cognition*, 1980, *8*, 513-520.

Taylor, I. Writing systems and reading. In: G.E. Mackinnon & T. G. Waller (Eds.) *Reading Research: Advances in Theory and Practice, Vol. 2*. New York: Academic Press, 1981, 1-51.

Templeton, S. & Scarborough-Franks, L. The spelling's the thing: knowledge of derivational morphology in orthography and phonology among older students. *Applied Psycholinguistics*, 1985, *6*, 371-390.

Terman, L. & Merrill, M. *Stanford-Binet Intelligence Scale*. Boston: Houghton Mifflin Company, 1973.

Treiman, R. The structure of spoken syllables: Evidence from novel word games. *Cognition*, 1983, *15*, 49-74.

Treiman, R. On the status of final consonant clusters in English syllables. *Journal of Verbal Learning and Verbal Behavior*, 1984, *23*, 343-356.

Treiman, R. Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology*, 1985a, *39*, 161-181.

Treiman, R. Phonemic awareness and spelling: Children's judgments do not always agree with adults. *Journal of Experimental Child Psychology*, 1985b, *39*, 182-201.

Treiman, R. Spelling of stop consonants after /s/ by children and adults. *Applied Psycholinguistics*, 1985c, *6*, 201-282.

Treiman, R. The division between onsets and rimes in English syllables. *Journal of Memory and Language*, 1986, *25*, 476-491.

Treiman, R. & Baron, J. Segmental analysis ability: development and relation to reading ability. In: G.E. Mackinnon & T.G. Waller (Eds.) *Reading Research: Advances in Theory and Practice (Vol. 3)*. New York: Academic Press, 1981, 159-198.

Treiman, R. & Breaux, A.M. Common phoneme and overall similarity relations among spoken syllables: their use by children and adults. *Journal of Psycholinguistic Research*, 1982, *11*, 569-598.

Treiman, R. & Chafetz, J. Are there onset- and rime-like units in printed words? In: U. Coltheart (Ed.), *Attention and Performance XII.* London: Erlbaum, to appear.

Treiman, R. & Danis, C. Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, in press.

Vennemann, T. Theory of syllabic phonology. *Linguistische Berichte*, 1972, *18*, 1-18.

Vennemann, T. The rule dependence of syllable structure. In: C. Duncan-Rose *et al.* (Eds.). *A Festschrift for Robert P. Stockwell* (in press).

Vihman, M.M. Phonology and the development of the lexicon: evidence from children's errors. *Journal of Child Language*, 1981, *8*, 239-264.

Vihman, M.M. A note on children's lexical representations. *Journal of Child Language*, 1982, *9*, 249-253.

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

Von Essen, O. Die Silbe - ein phonologischer Begriff. *Zeitschrift für Phonetik*, 1951, *5*, 199-203.

Walker, W. Notes on native writing systems and the design of native literacy programs. *Anthropological Linguistics*, 1969, *11*, 148-166.

Wang, H.S. On the productivity of vowel shift alternations in English: an experimental study. PhD dissertation, University of Alberta, 1985.

Wang, H.S. & Derwing, B.L. More on English vowel shift: the back vowel question. In: J.J. Ohala & J.J. Jaeger (Eds.), 1986, 99-116.

Warren, R.M. *Auditory Perception. A New Synthesis*. New York: Pergamon Press, 1982.

Waterson, N. Perception and production in the acquisition of phonology. In: W. von Raffler-Engel & Y. Lebrun (Eds.) *Baby Talk and Infant Speech*. Amsterdam: Swets & Zeitlinger B.V., 1976, 294-322.

Waterson, N. A tentative developmental model of phonological representation. In: T. Myers, J. Laver, & J. Anderson (Eds.), 1981, 323-333.

Wishart, D. *Clustan User Manual*. Edinburgh: Edinburgh University Program Unit, 1978.

Wood, C.C. & Day, R.S. Failure of selective attention to phonetic segments in consonant-vowel syllables. *Perception & Psychophysics*, 1975, *17*, 346-350.

Woods, A., Fletcher, P., & Hughes, A. *Statistics in Language Studies.* Cambridge: Cambridge University Press, 1986.

Zhurova, L.Ye. The development of analysis of words into their sounds by preschool children. In: C.A. Ferguson & D.I. Slobin (Eds.) *Studies of Child Language Development*. New York: Holt, Rinehart & Winston, Inc., 1973, 141-167.

Zwicker, E., Terhardt, E., & Paulus, E. Automatic speech recognition using psychoacoustic models. *Journal of the Acoustical Society of America*, 1979, *65*, 487-498.

List of Appendices

Appendix A

Experiment 1. List of words and counts for each of the analyses used (i) in all three tasks, (ii) in the sub-syllabic and phoneme counting tasks only, and (iii) in the syllable counting task only.

Word	Syllable Count	Sub-Syllabic Count: Rhyme	Sub-Syllabic Count: Vowel	Sub-Syllabic Count: Nucleus	Phoneme Count
i. All Counting tasks:					
ear	1	1	2	1	2
in	1	1	2	1	2
ape	1	1	2	2	2
ink	1	1	2	2	3
gas	1	2	3	3	3
plan	1	2	3	2	4
glare	1	2	3	2	4
break	1	2	3	3	4
toast	1	2	3	3	4
drift	1	2	3	3	5
penny	2	4	4	4	4
apron	2	3	4	3	5
item	2	3	4	3	4
vary	2	4	4	4	4
soda	2	4	4	4	4
ii. Sub-syllabic and Phoneme Counting only:					
art		1	2	2	3
ox		1	2	2	3
ask		1	2	2	3
tar		2	3	2	3
pin		2	3	2	3
rate		2	3	3	3
pray		2	2	2	3
spit		2	3	3	4
fence		2	3	3	4
soft		2	3	3	4
storm		2	3	3	5
drink		2	3	3	5
spend		2	3	3	5
grasp		2	3	3	5
flask		2	3	3	5

iii. Syllable Counting only:

Two syllables: flutter, private, stupid, pressure, climate

Three syllables: benefit, coconut, gravity, delicate, visitor, relative, definite, regular, typical, popular

Appendix B

Experiment 1. Training triads for the counting tasks

Syllabic	Sub-syllabic	Phonemic
ba-na-na	a-ppl-y	s-t-r-ay
he-ro	st-ay	f-l-y
my	eye	uh
ha-ppi-ly	fr-i-sk-y	c-l-ue
ca-noe	fl-y	m-y
stay	a	a
po-li-cy	fr-o-st-y	s-p-r-y
he-llo	fr-ee	b-ay
bay	owe	owe
ho-li-day	a-gr-ee	s-t-ay
co-ffee	gr-ay	l-ow
eye	uh	eye

Appendix C

Experiment 2. Percentage correct for individual pairs of words used in the deletion-by-analogy task (n=64 subjects).

Test Items (excluding analogy pairs):*

	Single Phoneme Deletions		Cluster Deletions	
Onset:				
	1. mold:(old)	58	3. flow:(owe)	44
	2. can:(an)	53	4. break:(ache)	56
	7. tar:(are)	53	5. split:(it)	44
	8. both:(oath)	56	6. spin:(in)	53
	10. meat:(eat)	64	9. trend:(end)	59
Coda:				
	1. tease:(tea)	39	2. loft:(law)	27
	3. base:(bay)	30	4. toast:(toe)	36
	5. grate:(gray)	31	7. traced:(tray)	33
	6. juice:(jew)	36	9. most:(mow)	30
	8. keep:(key)	36	10. grouped:(grew)	31
Incomplete Unit:				
	3. smack:(sack)	20	1. spring:(sing)	17
	6. shrug:(rug)	36	2. sixth:(sick)	8
	7. hand:(had)	14	4. strange:(range)	30
	8. trail:(rail)	31	5. splash:(lash)	17
	9. please:(peas)	28	10. strip:(sip)	9

* Numbers refer to order in which the pairs were presented.

Appendix D

Experiment 3. Substitution types, example pairs, stimulus words and targets, type of response, and percentage correct (n=40 subjects)

Substitutions		Examples	Stimulus:*target*	Type	%
Onset	/pl_/*	might-plight	scum:*plum*	Match	83
		drank-plank	treat:*pleat*	Match	68
			fought:*plot*	Mismatch	68
			stones:/*plonz*/	Nonsense	70
			choir	Glide**	
	/str_/	cling-string	bride:*stride*	Match	78
		beam-stream	duck:*struck*	Match	63
			slate:*straight*	Mismatch	48
			coach:/*strotʃ*/	Nonsense	53
	/g_/	frown-gown	splash:*gash*	Match	68
		tape-gape	chimp:*gimp*	Match	73
			trial:*guile*	Mismatch	65
			scrooge:/*guʒ*/	Nonsense	43
			swift	Glide**	
Vowel	/_ay_/	glade-glide	fraught:*fright*	Match	48
		dove-dive	hoist:*heist*	Match	55
			stool:*style*	Mismatch	63
			breath:/*brayθ*/		
			square	Glide, Nucleus**	
			cute	Glide**	
	/_ɛ_/	braid-bread	steam:*stem*	Match	75
		dusk-desk	mint:*meant*	Match	73
			size:*says*	Mismatch	75
			look:/*lɛk*/	Nonsense	60
	/_u_/	bath-booth	rust:*roost*	Match	65
		crown-croon	stowed:*stewed*	Match	60
			blaze:*blues*	Mismatch	58
			laughed:/*luft*/	Nonsense	48
			perch	Nucleus**	

* control items ** specific comparison pairs

Substitutions		Examples	Stimulus:*target*	Type	%
Coda	/_m/	tote-tome	shank:*sham*	Match	90
		clasp-clam	prince:*prim*	Match	45
			socked:*psalm*	Mismatch	48
			meld: /*mɛm*/	Nonsense	18
			jerk	Nucleus**	
			tears	Nucleus**	
	/_nts/	blimp-blintz	tempt:*tense*	Match	33
		faith-faints	past:*pants*	Match	30
			hop:*haunts*	Mismatch	28
			love:/*lʌnts*/	Nonsense	20
	/_ks/*	stripe-strikes	milk:*mix*	Match	35
		fond-fox	slapped:*slacks**	Match	55
			hedge:*hex*	Mismatch	50
			claim: /*kleks*/	Nonsense	43
			learn	Nucleus**	
			port	Nucleus**	
	/_lp/*	heck-help	ask:*alp**	Match	60
		gust-gulp	yen:*yelp*	Match	78
			wedge:*whelp*	Mismatch	70
			tint:/*tɪlp*/	Nonsense	45
Head	/kræ_/	floss-crass	stink:*crank*	Match	20
		drift-craft	blush:*crash*	Match	15
			chic:*crack*	Mismatch	15
			snail: /*kræl*/	Nonsense	10
	/le_/*	broke-lake	price:*lace*	Match	23
		freeze-laze	soothe:*lathe**	Match	25
			scat:*late**	Mismatch	15
			gruff: /*lef*/	Nonsense	18
	/slɪ_/	trap-slip	fret:*slit*	Match	30
		bonk-slink	strong:*sling*	Match	28
			tomb:*slim*	Mismatch	23
			brusque: /*slɪsk*/	Nonsense	18

* control items ** specific comparison pairs

Substitutions		Examples	Stimulus:*target*	Type	%
Rhyme	/_el/*	most-mail	roof:*rail**	Match	95
		stun-stale	pound:*pail*	Match	95
			flax:*flail*	Match	85
			spite:/*spel*/	Nonsense	83
			quest	Glide**	
	/_ʌnt/	freak-front	pesk:*punt*	Match	55
		stealth-stunt	groups:*grunt*	Match	55
			wreath:*runt*	Mismatch	55
			crimps:/*krʌnt*/	Nonsense	43
	/_old/	baste-bold	monk:*mold*	Match	90
		strict-strolled	scant:*scold*	Match	88
			kept:*cold*	Mismatch	93
			chomp:/*tʃold*/	Nonsense	90
Margins	/l_p/*	stoke-lope	kiss:*lip**	Match	38
		gash-lap	teach:*leap*	Match	30
			brought:*lop*	Mismatch	28
			chunk:/*lʌp*/	Nonsense	20
			mute	Glide**	
	/p_l/*	sheen-peel	sick:pill	Match	48
		dozed-pole	ride:pile	Match	43
			good:*pull*	Mismatch	8
			vest:/*pɛl*/	Nonsense	35
	/b_t/	strife-bite	gloom:*boot*	Match	28
		tenth-bet	mouse:*bout*	Match	23
			niece:*beet*	Mismatch	25
			crook:/*bʊt*/	Nonsense	18
			swap	Glide**	
	/st_k/	bathe-stake	glimpse:*stick*	Match	20
		plug-stuck	fast:*stack*	Match	20
			taught:*stock*	Mismatch	23
			dread:/*stɛk*/	Nonsense	18

* control items ** specific comparison pairs

Appendix E

Experiment 4. ·Words used in the comparisons of onset, coda, and nucleus changes with mean PSJ for each (30 word pairs).

		Mean
Onset Changes:		
CV:xV	næ:mæ	7.7
	bæ:dæ	7.1
CCVC:XXVC	spik:blik	5.4
	skut:brut	4.5
	grɪp:snɪp	5.4
CCVC:xXVC	kyud:glud	5.[illegible]
	kyud:prud	5.1
	spik:ʃrik	4.2
CCVC:XxVC	swayp:grayp	5.2
	swʌm:glʌm	5.5
	slɪm:grɪm	4.8
Coda Changes:		
VC:Vx	æb:æd	6.4
CVCC:CVXX	bʌst:bʌgz	3.8
	dɛsk:dɛks	5.5
	kʌps:kʌsp	6.1
CVCC:CVxX	wɪsp:wɪzd	4.9
	lɪft:lɪvz	4.2
CVCC:CVXx	hɪps:hɪdz	6.3
	tʌbz:tʌks	4.8
	kʌsp:kʌpt	4.5
Complex Nucleus:		
VR:Vx	æm:æn	6.8
CVRC:CVXX	kʌld:kʌps	2.8
	pirz:pikt	3.2
	rild:rimz	2.9
CVRC:CVxX	sæŋk:sænd	4.5
	læmp:lænd	4.7
	læmp:læbz	3.8
CVRC:CVXx	hɪlz:hɪps	3.5
	dɪlz:dɪmd	2.4
	pirs:pivz	2.5

C=Consonant V=Vowel R=resonant of complex nucleus x=small change X=large change

Appendix F

Experiment 4. Stimuli for the comparisons of near phoneme mismatches in various positions with mean PSJ for each (64 word pairs).

Simple Nucleus (CCVCC)			Complex Nucleus (CCVRC)		
CCVCx	klæsp:klæst	7.6	CCVRx	twɛlv:twɛlð	7.0
	flʌks:flʌkt	6.0		krʌmd:krʌmz	6.8
	blæst:blæsk	7.9		sport:spord	7.2
	drɪft:drɪfs	7.2		blerd:blert	6.2
CCVxC	klæsp:klæʃp	6.2	CCVxC	twɛlv:twɛrv	4.1
	flʌks:flʌts	7.0		krʌmd:krʌnd	6.1
	blæst:blæʃt	6.0		sport:spolt	5.0
	drɪft:drɪθt	8.9		blerd:bleld	4.6
xCVCC	klæsp:glæsp	7.4	xCVRC	twɛlv:dwɛlv	7.5
	flʌks:slʌks	6.9		(krʌmd:glʌmd)	*
	blæst:plæst	7.2		sport:ʃport	7.1
	drɪft:trɪft	7.6		blerd:glerd	7.5
CxVCC	klæsp:kræsp	7.3	CxVRC	twɛlv:trɛlv	7.2
	flʌks:frʌks	6.2		krʌmd:klʌmd	7.2
	blæst:bræst	7.1		sport:skort	6.9
	drɪft:dwɪft	7.0		blerd:brerd	6.6
CCVxx	klæsp:klæʃt	5.1	CCVxx	twɛlv:twɛrf	3.8
	flʌks:flʌgz	5.8		krʌmd:krʌnt	4.0
	blæst:blæʃk	5.6		sport:spold	3.3
	drɪft:drɪvd	5.9		blerd:blelt	3.4
CxVxC	klæsp:kræʃp	5.1	CxVxC	(twɛlv:trɛrf)	*
	flʌks:frʌts	5.1		krʌmd:klʌnd	4.8
	blæst:bræʃt	4.2		sport:stolt	1.9
	drɪft:dwɪθt	6.4		blerd:breld	3.1
xCVCx	klæsp:glæst	5.2	xCVRx	twɛlv:dwɛlf	5.0
	(flʌks:θrʌkt)	*		(krʌmd:grʌmp)	*
	blæst:glæsp	5.8		sport:ʃpord	4.5
	drɪft:trɪfp	4.6		blerd:plert	3.1
xxVCC	klæsp:træsp	6.0	xxVRC	twɛlv:drɛlv	6.2
	flʌks:vrʌks	6.4		krʌmd:plʌmd	6.2
	blæst:dræst	5.8		sport:ʃtort	5.3
	drɪft:twɪft	5.7		blerd:drerd	5.5

* Not included in analyses because changes were too large.

Appendix G

Experiment 4. Stimuli for the comparisons of rhyme changes with mean PSJ for each (30 word pairs).

One unit out of two changed:

VC:XC	it:ɔt	3.1	CV:CX	bi:ba	3.3
	ist:ɔst	4.0		kri:kra	3.8
	ed:ʌd	2.8		pe:pʌ	3.1
	eks:ʌks	3.7		sle:slʌ	3.6

One unit out of three changed:

CVCC:CXCC	bist:bɔst	4.0	CVRCC:CVCC	ʃarps:ʃaps	5.3
	krist:krɔst	4.2		birdz:bidz	4.0
	sleks:slʌks	5.3		porks:poks	4.1
	peks:pʌks	3.4		forst:fost	4.7
CVRCC:CVRXX	pʌrst:pʌrvz	4.4			
	bardz:barft	4.1			
	turz:turp	4.0			
	forst:fordz	4.9			

Two units out of three changed:

CVRCC:CVXX	forts:fozd	2.4	CVRCC:CXXX	bardz:bist	1.9
	pirst:pivz	2.9		porks:pæpt	1.5
	ʃarps:ʃagz	2.6		ʃarps:ʃɪvd	2.2
	lordz:lofs	3.1		birdz:boft	1.4
CVCC:CXXX	bist:bavd	1.7			
	pʌks:pezd	1.6			

C=Consonant V=Vowel R=/r/ X=large mismatch