AN EXPLORATION OF PSYCHOLINGUISTIC UNITS IN INITIAL READING

by

Duncan Hansen and Theodore S. Rodgers

TECHNICAL REPORT NO. 74

July 6, 1965

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

I. Introduction

Each generation has seen reading theory fall under the ascendancy of the current Zeitgeist. In our own time it appears that much of the impetus for new reading research emanates from the similarly new inter-disciplinary field of psycholinguistics. Linguists, psychologists, and educators have been actively collaborating in the formulation of propositions about the way a child's perception and assimilation of orthographic materials are controlled by features of the language, of the immediate stimulus situation, and of the child's behavioral history. Psycholinguistics may be defined as the scientific activity that seeks empirical confirmation or disconfirmation for such propositions. Before presenting the details of our paper which will report the activities of the Stanford Project on Computer-Based Instruction in Initial Reading, we would like to comment on some recent trends in psycholinguistics, which will, we hope, provide a background to activities of the project.

In any consortium of inter-disciplinary fellows, one comes to expect a wide variation in the units of behavior proposed to describe an activity like initial reading. Linguists, for example, have proposed several such units which customarily coincide with the formal units within the levels of structural analysis. Hence, the forwarding of the phoneme, the morpheme, and the sentence; these being the basic units of linguistic analysis at the phonological, morphological, and syntactic levels, respectively. A scanning of the current linguistic literature, however, reveals a disconcerting amount of disagreement among linguists as to the requirement and definition of the proposed units and levels. There is, if possible, even less concensus as to the explicit interrelations between the different levels (Chomsky, 1963).

Psychologists, in their turn, offer alternative views concerning potential units in language behavior, generally, and in initial reading, specifically (Jenkins, 1964). Extended S-R associational theories formulated from a tradition of experimentation with N-gram and word associations have been challenged by some psychologists (Miller, 1964) who stress the role of recognition and feedback processes. The report of the 1954 Conference on Psycholinguistics (Osgood and Sebeok, 1954) is replete with extensive discussions of the relative merits of proposed linguistic-behavioral units. In regard to reading per se, scholarly discussions of "the" appropriate reading sub-units, particularly the letter vs the word, date back to the 18th century or before (Fries, 1963; Morris, 1964).

Fully aware then of the sound and fury which has accompanied consideration of linguistic units variously championed by linguists, psychologists, and educators, we timorously offer a new contender to the ranks - a unit we shall call the "Vocalic Center Group." The body of the paper will both define the Vocalic Center Group and substantiate our reasons for considering the Vocalic Center Group (hereafter VCG) a viable unit for initial reading. Our confidence in this new construction rests in the systematic relationship between the VCG, linguistically defined, and the behavioral units one finds necessary to assume in order to account for the specific performance of children in initial reading.

Study of these behavioral units and particularly the mechanism implied in the utilization of such units represents a second facet of our investigations into reading behavior. Here the approach attempts to characterize the child-reader as a device capable of assimilating the systematic structures in natural language, as represented in either speech or reading, when exposed to specific stimulus situations. In this regard we find the work of Liberman and Cooper at Haskins Laboratory illuminating. Liberman, et al., (1959) have suggested a rule set and appropriate transmission-reception mechanisms that are sufficient for both production and perception of continuous speech sounds. From another vantage, we note that linguists working on generative and particularly transformational grammars have formulated sets of generative rules which when order-processed by a syntactic mechanism are capable of producing demonstrably proper sentence sequences (Halle and Stevens, 1963). Later in the paper we shall present a schema which may hopefully annotate the close relationship observable between speech production, speech perception, and initial reading.

American linguists have been rather outspoken in their criticism of educators who, it has been felt, fail to recognize the high degree of linguistic sophistication which children bring to their initial schooling (Carroll, 1964). The considerable language skills possessed by children prior to formal training in language arts has led us to consider the initial reading task as a transfer process that requires the mapping of an orthographic symbol system onto the already extensive language repetoires. This view suggests that the reading task be sequentially sub-divided, not only chronologically but methodologically.

Traditionally, the first stage consists of teaching the child that the orthographic symbols he sees on the printed page correspond in some way to his spoken language. The immediate task consists of pronouncing letters, words, or sentences aloud as a response to the orthography. This initial skill in decoding the orthography is assumed to be a necessary prerequisite to a higher level response in which the child comprehends the meaning of the segment and makes recall or inferential statements. In this paper we shall restrict ourselves to a discussion only of the first stage.

As theory and implementation have evolved in psycholinguistic research, some parallel developments have been taking place in computer science. The now-working knowledge of programmed learning coupled with the technological advances in computer console instrumentation have opened a bright new area for intensive individualized instruction and perhaps, more crucially, curriculum evaluation. These developments will offer important opportunities to linguists, psychologists, and educators alike. Linguists have here the tools for extensively tracking specific features of language acquisition. Psychologists may look to these developments as providing a promising setting for theoretical studies of such dependent measures as latencies. Educators may here find the means to make valid teacher-independent comparisons between different instructional methodologies. In the latter portion of this paper we will describe a computer-based instructional system that is now coming into existence at Stanford University.

We trust it is in keeping with the purpose of this symposium to have taken this discursive view of some recent developments in psycholinguistics. We shall impose on your good will in two further respects: first, we shall present our project report chronologically and second, we shall initiate our formal presentation with the discussion of a study considerably removed from initial reading. We hope this method of presentation may give you a sense of the way our own thinking has developed in our attempt to isolate the psycholinguistic units and behaviors of initial reading.

II. The Russian Study

At the Institute for Mathematical Studies in the Social Sciences, Professors Suppes and Crothers have been conducting a number of studies concerning the optimal list size for learning Russian-English vocabulary pairs (to be reported in Suppes and Crothers, book in preparation). The term block size refers to the average number of unique pairs intervening between the presentation of pair i on trial n and the re-presentation of that pair on trial $n + 1$. The task for the subject was as follows. A Russian item was pronounced from tape by a native Russian speaker. After five seconds a correct English translation for the item was given. If the subject recognized the Russian word from previous presentations, he was to write the English translation before it was pronounced on the tape. Subjects were called back after a week and given a recall test. In reviewing the primary data from the results of two groups of 20 subjects studying 300 Russian-English vocabulary pairs over a nine-day period, we discovered that some pairs were both acquired and retained with surprising ease by subjects regardless of block size whereas other pairs were never learned by any subject in any test condition.

The high inter-subject consistency as regards item difficulty encouraged us to attempt to determine those variables about the pair which seemed to facilitate or retard the learning. After investigating a number of possible variables in both the Russian stimulus words and the English response words, we found a form class schema to be the most illuminating. Using a class-frame sentence technique (Fries, 1952), we analyzed the 300 words into 12 classes. These results are presented in Table 1. The results indicate a regular monotonically decreasing relationship between the various classes and proportion of correct responses. Independent research has indicated a similar learning hierarchy for the major function classes (Glanzer, 1962).

In light of this finding we formulated the hypothesis that functional and possibly semantic properties of the English words provided a mediational linkage to the Russian associates. We then planned a further experiment in an attempt to substantiate this proposition. Selecting the 25 most-learned and the 25 least-learned of the 300 pairs, we reversed the order of the English responses to the Russian stimuli so as to form a set of spurious translation pairs (that is, Russian item 1 was now paired with English item 50, etc.). If our hypothesis concerning the dominating influence of the English response was correct, we expected to observe that the acquisition scores of this new set of

50 pseudo-pairs should be more or less the reverse of what was found in the prior experiment.

The subjects were given exactly the same instructions as in the prior experiment. They were presented the pairs in block sizes of 50 for 6 trials using the same equipment and procedure as for the earlier experiments. Six randomizations of the 50 items were used. The 6 trials were given to 10 subjects in one day with a rest break between the 3rd and 4th trials.

The results of the experiment are presented in Table 2. The first row reports the mean proportion of correct responses in the prior Russian study. The second row presents the mean proportion of correct responses in this experiment that reversed the Russian-English pairing. If our hypothesis were to be confirmed, one would expect to find that the means in row two would be the reverse of the prior experimental results presented in row one. To our chagrin the means did not reverse. We were thus forced to shift our attention from the function or meaningfulness of the English responses to factors inherent in the Russian stimuli.

Suspecting that the length of the Russian word might be a critical variable we performed a syllable count, the totals of which are shown in the third row of Table 2. The number of syllables for the two lists of 25 Russian words are approximately equal.

We then attempted to explore in greater depth a "covert rehearsal" hypothesis which many verbal learning experimenters posit as a necessary condition for success in rote learning tasks. It appeared obvious that rehearsal and pronounciability rating of a Russian word must be highly dependent. How then might one get a pronounciability rating of a Russian word by an American subject? Professors Greenberg and Jenkins (1964) have proposed a method for determining structural distance of any given syllable from the syllable canon of English. Unfortunately, the rating procedure is exceedingly arduous, requiring information about English syllable structure which must be hand-generated, and ultimately yields scores only for monosyllabic items. A simpler and more general procedure was obvious desirable. One set of linguistic features which appeared somewhat promising and at the same time relatively easy to isolate was the set of Russian consonant sequences. Using a phonetic transcription of the Russian stimulus words we sought to determine which of the Russian consonant sequences were phonetically and positionally (initially and finally) similar to consonant sequences in English. Having matched those Russian sequences which were phonetically similar to sequences in English, we then determined a consonant complexity score for each Russian word. A Russian word without a consonant sequence (a sequence is defined as two adjacent phomes of the same type--consonant or vowel) was assigned a complexity score of zero; an item containing a consonant sequence phonetically similar to an English sequence in the same position was assigned a score of one for each such sequence; an item with sequences dissimilar to any found in a comparable initial or final position in English was assigned a complexity score of 2 for each such sequence.

This rating scheme, you will note, automatically weights complexity scores for length of consonant sequences. The results of this rating yield the values in row 4 of Table 2. The consonant complexity score ranking is observed to be closely correlated with that of the learning score. Apparently English-like consonant sequences are easier to rehearse than non-English consonant sequences. It appears, additionally, that longer consonant sequences, English-like or not, are more difficult to rehearse than shorter consonant sequences.

These two propositions are anticipated by linguistic evidence on the restrictive subset of consonant sequences that occur in comparison to the large combinatorial possibilities. Greenberg (1965) notes in a discussion of consonant cluster universals in 104 languages:

> "1. For initial and final systems, if $\underline{x}$ is the number of sequences of length $\underline{m}$ and $\underline{y}$ is the number of sequences of length $\underline{n}$ and $\underline{m} > \underline{n}$, and $\underline{p}$ is the number of consonant phonemes, then $\frac{x}{p^m} \leq \frac{y}{p^n}$
>
> In other words, the proportion of the logically possible ambinates utilized decreases or remains the same with increasing length of the sequences. This may be illustrated for English initial clusters as follows: the number of consonant phonemes are 22. All of these except / $\underline{\check{z}}$ / and /ŋ / occur as single phonemes. The logically possible sequences of length 2 are $22^2 = 484$. Of these 28 occur. For length 3 the logically possible number of combinations is $22^3 = 10{,}648$. Of these only 6 occur...
>
> 2. For initial and final systems, if $\underline{x}$ is the number of sequences of length $\underline{m}$ and $\underline{y}$ is the number of sequences of length n, and $m > n$ and $n \geq 2$, the $\underline{x} \leq \underline{y}$ ... syllables containing sequences of $\overline{n}$ consonants in a language are to be found as syllabic types, then sequences of $\underline{n-1}$ consonants are also to be found in the corresponding position (prevocalic or postvocalic) except that $CV \rightarrow V$ does not hold ...

In general, the validity of 1 and 2, to which no exception was found in the 104 languages of the sample, provides objective evidence of the "difficulty" of clusters. This would seem to correlate with the diachronic tendency towards their simplification, since any simplification automatically reduces the number, both absolutely and proportionally, of sequences of the length subject to reduction and increases the number of shorter sequences."

Encouraged by the linguistic-universal evidence as to the length-difficulty of clusters and by our own evidence as to the role consonant clusters assume in rehearsal difficulty, we entertained the conjecture that some phonemic sequences might be easier for initial readers to rehearse than others. More specifically, that rehearsal difficulty of consonant clustered units may be hierarchically determined, in that each pronounceable sub-unit, of a larger unit is a) permissible in the language, and b) less difficult to rehearse (Greenberg's evidence suggests that the CV unit may represent a lower bound on this generalization.). One anticipates that there may be evidence of rehearsal or pronunciation "ease" reflected in all of the dominant as opposed to the recessive phonemic patterns in language.

The relationship between the Russian study just discussed and reading research may appear highly tenuous. We have discussed the Russian experiment in this detail because we feel it demonstrates several necessary if not sufficient conditions which we have imposed on our present investigations in beginning reading.

We first examined the Russian data in the light of several interdependent hypotheses known to psychologists under rubrics such as meaningfulness, frequency of occurrence, availability, familiarity, concreteness, etc.. These "hypotheses" often have appeared as "criteria" in choosing the initial vocabulary and sequence of reading materials. Thus, we find vocabulary chosen on the basis of frequency (McKee, et al., 1963), meaningfulness (Fries, 1963), and situational availability (Russell and Ausley, 1959). In the Russian study these hypotheses appeared intuitively reasonable and, in one instance, gave a fairly accurate account of the initial data. All failed to yield valid predications as to the results of the follow-up experiment. We will attempt to demonstrate in what follows that such hypotheses also fail to adequately characterize observations of initial reading behaviors.

Secondly we have attempted to indicate, again on the basis of the Russian experiment, some phonetic features of linguistic response which appear to strongly influence the ease or difficulty of language learning tasks. Two such features, composition and length of consonant sequences, which strongly determined the success of pair-learning in the Russian experiment, will be re-examined in our discussion of the sequence of beginning reading material.

In order to understand why certain phonetic sequences present greater difficulty than others, one would like to develop some sort of operational characterization of speech production at the phonological level. We would now like to turn our attention to some relevant research in this area of speech production.

III. Factors in Speech Production.

Our hypothesis concerning the role of rehearsal in the acquisition of oral reading behavior led us to a serious re-examination of the literature on speech production with particular reference to pronouncability of vocabulary items (See Greenberg and Jenkins, 1963, or Underwood and Schultz, 1960, for examples of adult rating of pronouncability). Such an approach seems of limited experimental use in working with children and leaves undefined the principal problem of clarifying the independent variables. A more promising source of information as to the relative difficulty of pronunciation units was found in the experimental research on speech synthesis. Several types of speech synthesizers are currently being explored (See for example, Rosen, 1958 or Peterson, 1958). All of these minimally require some set of basic discreet units and some sort of rules for converting these discreet units into a speech-like continuum. We felt that a measure of pronounciability might be deduced from the nature of the base units and the number and type of rules required to synthesize the pronunciation in question.

Rules for synthesis involve considerably more than mere concatenation over phonemes or any other phonetic element. Independent evidence from linguists (Hockett, 1955) and acousticians (Harris, 1953; Liberman, et al., 1959) emphatically suggests that the speech signal can in no way be equated to a string of phonemic beads. Hockett pictures the speech signal as a set of broken phonemic Easter eggs, yolking into one another, if you will. Among the more revealing yolking procedures are those suggested by Liberman, Ingemann, Lisker, Delattre, and Cooper (1959) of the Haskins

Laboratories in their rules for speech synthesis. These rules are specified in terms of acoustic features, which, when reproduced as a spectographic pattern, can be played back as recognizable speech. The synthesis rules are sub-divided into three groups according to the function they play in speech production. The first set of rules specifies the core characteristics of a given phoneme, i.e., the form and frequencies and/or loci which typify the three principal formants of that phoneme. The second group of rules specifies the sequential transitions between core centers necessary to smooth the discreet core characteristics into a speech continuum. The third group of rules called "position modifiers" provide alteration of the core rules for a given phoneme on the basis of adjacent phonemes. In essence, these "position modifiers" determine the allophonic variations of the phonemes in the phonetic string. An example from Liberman, et al., (1959) may help clarify the nature of the "position modifiers." The example is for the syllabic unit /glu/. The formant frequencies, intensities, durations and transition characteristics for /g/, /l/, /u/, are specified and then the authors note,

> "Now a rigid application of the basic rules for the phonemes constituting the syllable /glu/ yields an ultimate acoustic output of less than tolerable intelligibility. A marked improvement is achieved if the basic rules for each phoneme are modified as follows: /g/ before /l/ requires only a burst of specified frequency; /l/ before /u/ has the frequency of its second formant lowered somewhat; /u/ following /l/ has a second formant which first rises from the second formant frequency of /l/ and then, after a specified duration, shifts at a given rate to the normal steady-state frequency for /u/."

Similarly, a successful human rehearsal must operate over units such that the outputed forms are of tolerable intelligibility. A simple concatenation of phonemes or even allophones will not, as Liberman suggests, produce such a tolerably recognizable output. (It is recognized that additional requirements such as those for syllabic stress are also implied by our intelligibility criterion. Liberman, et al., do suggest a criterion for determing vowel characteristics under primary stress.)

The required detail of the rule specification shown in the example above suggests that speech activity conceived in these terms must require manipulation of a ponderous collection of such rules. Due largely to the overlapping of subphonemic features many of the redundant specifications in phonemic classification can be eliminated in the speech synthesis model. In fact, the Liberman, et al model utilizes 17 rules for consonantal, and 14 rules for vocalic articulation. This compares favorably with the English phonemic inventory considered to be about 40 phonemes. In order to form connected speech of tolerable intelligibility (not a requirement of phonemic descriptions) twelve position modifiers, a stress modifier, and a meshing routine are also required.

There is a notable isomorphism between the articulation feature sets and the synthetic formant sets. For example, one applies a single articulatory rule for aspiration in initial stop consonants. Similarly, to synthesize aspiration, a single rule specifies 50 millisecond aperiodic second and third formant transitions to steady-state vowel formants. In both cases, the rules apply in the same fashion to the same phonemes.

In terms of our present interest, the categorical nature of rule specification suggests to the linguist economical statements of phonotactic constraints; offers to the psychologist an inference set of complexity measures; and presumably, simplifies for the language-user perception and production of speech sequences. An example of the use to which rule specifications of this type might be put is given by Saporta (1954). Saporta has hypothesized that an intermediate number of feature shifts between successive consonants optimizes both speech perception and production. Such a statement is linguistically and psychologically testable.

We interpret this whole line of linguistic and psycholinguistic evidence to suggest that there is a continuum of preference as to pronunciation and perception of speech units and that the learning effects due to rehearsal are directly related to this continuum. The task now becomes one of relating the generic properties of the rules for speech synthesis to an appropriate psycholinguistic perception-production unit and secondly, relating the processes of speech perception and production to the processes of initial reading. We now turn our attention to the first of these tasks.

IV. Vocalic Center Group

The psycholinguistic unit for initial reading that we propose in the Vocalic Center Group is an elementary structure resulting from the integration of phonemic elements into a minimal pronunciation unit. The Vocalic Center Group is a structure in the sense that it is the optimally minimal sequence within which all necessary rules of phonemic co-occurrence can be stated. Such rules are commonly referred to as phonotactic rules. By integration we refer to the process whereby phonemes are positionally modified so as to form a phonotactically permissable and tolerably intelligible pronunciation. The VCG is marked by one vocalic element (which is not necessarily a vowel). Non-vocalic (consonantal) or semi-vocalic elements may occur preceeding or following the vocalic center. The "complexity" of phonotactic rules governing the phonemic combinations within the VCG, we hypothesize, are intimately related to the "difficulty" of speech production, speech perception, and we will claim, initial reading behaviors.

It may have occured to some of you that our discussion of the VCG thus far is in fact a discussion of that unit which has been traditionally called the syllable. What justification can we offer for adding "the Vocalic Center Group" to the already overloaded lexicon of psycho-linguo-educational neologisms?

In reviewing the literature on the syllable (Haugen, 1954), one finds a vast array of definitional diversity and argumentation as to the acceptability of the syllable. Attempts to associate syllables with the breath pauses physiologically observable in the intercostal muscles have been generally successful only in giving quantitative accounts of the number of vocalic nuclei (Stetson, 1945). No valid criteria for determining syllable cuts between adjacent consonant groups at the syllabic margins are offered. Definitions based on distributional procedures (O'Connor and Trim, 1951) offer little insight as to the classification of these distributional observations and leave unresolved the problem of marking syllable boundaries in polysyllabic units. Other linguistic theorists have attempted to define the syllable on the basis of higher level, supra-segmental features; thus the attempts to identify syllabic nuclei with stress (Pike, 1947) or syllabic boundaries with juncture (Harris, 1951). The limitations of these proposals have been discussed extensively in the linguistic literature. Attempts to identify the orthographic syllable have presented similar problems. Contradictions between the phonological, morphological, and historical criteria used in determining lexographic syllabification have been bitterly bewailed by the very lexicographers who perpetuate the system. The unfortunate syllable has fallen heir to the calumny and confusion of its definitions. For reasons then, both political and theoretical, we felt a different terminology advisable.

Perhaps the most valid reason for introducing the new terminology, however, was our intention to assign to the VCG properties of an entirely different nature than those previously defined into the syllable. First, as we have noted, the VCG is the minimal construct defined over a set of phonotactic rules. These rules define constraints not only between continuous elements (consonant-consonant, consonant-vowel, vowel-consonant) but over discontinuous elements (pre- and post-vocalic consonants) as well. Second, ambiguities encountered in syllable cuts at consonantal boundaries are recognized and accounted for in terms of VCG priority rules. These priority rules are used as a basis for discussing individual variations encountered in syllabification of speech. Third, the rule structure of a particular VCG is employed to make predictions concerning the facility of rehearsal of the VCG and, ultimately, the facility for association of the phonotactic to the orthographic pattern. Finally, the structure of the VCG requires us to make both quantitative and qualitative predictions as to the type of errors that initial readers will make in forming pronunciations for a particular graphic segment.

In presenting the rules governing the VCG, we shall refer to the Vocalic Center by the letter V and the consonant cluster by the letter C with superscripts i, f, or m, standing for initial, medial, or final positions. Descriptive listings and examples will be held to a minimum as one can find a reasonably complete description of the permissable consonant sequences in English in Bloomfield (1932), Whorf (1940) and Harris (1951). What follows below are examples of the various phonotactic rule types for Monovocalic Center Groups.

Case 1 Rules Governing Initial Consonant Clusters of the Form C^i ... Where $i = 2, 3$.

A. Consonants having the same manner of articulation do not occur in the same cluster. The phonemic classes of stop, fricative, nasal, liquid, semi-vowel may be represented no more than once in any cluster. Thus one finds /str-/ but not */sθr-/ or */stp-/ or */slr-/.

Case 2 Rules Governing Initial Consonant Clusters and Following Vowel of the Form C^iV ... Where $i \geq 2$.

Clusters containing the semi-vowel /y/ can be followed only by the vowel /u/. Thus, one has /m y u wt/ mute but not */m y a wt/.

Case 3 Rules Governing Final Consonant Clusters of the Form ... C^f Where f = 2, 3, 4.

Two consonants having the same manner of articulation may cluster only if the second member is articulated in the alveolar region. Thus, one finds /fs, θs, pt, kt, bd, gd, rl, dz, ɣ z, bz, gz, vz/# but not */sf, sθ, tp, tk, db, dg, lr, etc./#.

Case 4 Rules Governing Vowel and Following Consonants of the Form ... VC^f.

A. In VC^f if V = /i, e, æ , u/, then C^f must be greater than zero. Thus we find final vowel /ɔ/ in law, /aw/ in now, /a/ in ma, /ə/ in sofa, /iy/ in me, /ay/ in my, /ɔy/ in boy, /ey/ in bay, /ow/ in bow, and /uw/ in do but not $*C^i$ + /i, e, æ , u/#.

B. In VC^f if V ≠ /i, e, æ , ə , u, a, ɔ /(i.e., simple vowels), then C^f can not be of the form liquid (/r,l/) plus any consonant except /d/ and /z/. Thus we find /part/ part, /help/ help, /silk/ silk, /mə lč/ mulch, /fayld/ filed, /fiyld/field, /feyld/ failed, /fɔ yld/ foiled, /fawld/ fouled, /fuwld/ fooled, /fowled/ foaled, but not */fayrt/, */feylp/, /*fiylk/, */fuwlc/, etc.

Case 5 Rules Governing Initial and Final Clusters of the Form C^i and C^f Where i = 2, 3, and f = 2, 3, 4.

A. All consonants which are members of a voiced-voiceless pair set will be all either voiced or voiceless in the same cluster. Thus one finds /s t u w/, stew, /r æ f t s/ rafts but not */s d u w/, */r æ v t s/, */r æ f d s/,* /r æ f t z/.

B. Any linear partition of a consonant cluster will yield sequences permissible in the same pre- or post-vocalic position as the principal cluster. Thus, initial /s t r-/ is divisible into the permitted initial sequences /st/+/r/, /s/+/tr/, /s/+/t/+/r/; or /-ŋ kθs/ is divisible into the permitted final sequences /ŋ / + /kθs/ or /ŋ kθ/ + /s/ or /ŋ k/ + /θs/ or /ŋ / + /k/ + /θs/ or /ŋ / + /k/ + /θ/ + /s/, etc.

B' Any medial complex consonant cluster (poly-syllabic words where $m \geq 2$) is divisible into a permitted final plus initial cluster. For example, /ekstrə/ extra must be divisible into at least two simpler initial-final clusters /eks + trə/, /ek + strə/, /ekst + rə/.

Case 6 Rules Governing Cross-Vocalic Consonant Clusters of the Form C^iVC^F Where $i \geq 2$ and $f \geq 1$.

A. Any sonant (/m, n, l, r/) may occur either in initial or final clusters but the same sonant may not occur in both initial and final clusters in the same VCG.

Thus one finds slit and still but not *slill; bread and beard but not *breard; small and slam but not *smam, snug and shun but not *snun.

Phonotactic constraints of this type have not, to the best of our knowledge, been previously noted. They provide the strongest justification for our claim that the VCG is the minimally optimal unit within which rules of phonemic co-occurrence can be stated. This restriction does not apply to form C^iVC^f where $i \leq 1$. One does find lull, rare, mum, and none. We believe other phonotactic constraints apply, less absolutely, to cross-vocalic consonant clusters. We call such forms "unfavored" VCG's. We postulate a broad continuum of such VCG forms that range from "non-permissible" to "favored."

From a psycholinguistic point of view, the ranking of a particular VCG form along this continuum may be inferred from the number and type of speech synthesis rules one must apply to generate a pronunciation. The larger the number of rules required, the more difficult the rehearsal (The less formal "Principle of Least Effort" has long been a popular explanantion for ease of articulation (Zipf, 1949)). As the consonant clusters are the only part of the VCG that may vary in phonemic length, we hypothesize that longer consonant clusters will result in more difficult rehearsal and consequently will be acquired with greater difficulty in, say, the context of initial reading. We also speculate that for clusters of equal length the rules for some sequences will be more complex than for others. One apparent form of sequential complexity at this level is measured in terms of the phonetic similarity of adjacent phonemes. This suggests that a sequence of the form apt/æpt/ containing a consonant-phoneme pair identical in manner of articulation and voicelessness would be more difficult, in the sense discussed above, than a form like art/art/ where the consonant phonemes share no phonetic features.

Desiring an empirical test of these assumptions, and more specifically, of the viability of the VCG in a reading situation, we designed a learning experiment in which certain complex monosyllabic words and certain simple disyllabic words would be taught as reading units. In an attempt to avoid confounding the results with semantic factors we constructed pseudo-words from occurrent phonotactic patterns. A minimal disyllabic form that avoids adjacent syllabic vowels is the pattern CVCVC. This represented the disyllabic pattern. We assumed stress on the first vowel. The monosyllabic forms consisted of two CCCVC, two CCVCC, and two CVCCC patterns. We also wished to test pre-vocalic and post-vocalic cluster difficulty to see if we could partially replicate the studies of children's speech perception indicating that post-vocalic errors are twice as frequent as pre-vocalic errors (Templin, 1943). Controlling for phoneme length and phoneme frequency where possible, we constructed the list of twelve forms presented in Table 3. We tested eight-year old children in order to have subjects who had only partially mastered the initial reading process (i.e., subjects capable of rehearsing words but limited in practice on the orthographic consonant sequences found in the list).

A paired-associate anticipation method was employed. The children, tested individually, were shown a 3" x 5" card on which a given word was printed in 3/4" letters. A four second interval was given for a response. After a response or at the end of four seconds, the experimenter pronounced the word and the child overtly rehearsed it once. The cards were reshuffled at the end of each trial (a trial consisted of once through the deck). Twenty children were run for ten consecutive trials.

The mean total errors for each pseudo-word are presented in Table 3. The criterion subjects were defined as children who had two or more consecutive errorless trials. A learning curve for each group of words is presented in Figure 1. An analysis of variance yielded a highly significant difference ($F = 19.67$, $P < .001$) in favor of disyllabic pseudo-words. The summary statistics are presented in Table 4.

In regards to individual behavior, the relative difficulty of the consonant clustered monosyllabic words was even more prominent. All subjects initiated errorless sequences earlier for the disyllabic words than for the consonant cluster words (i.e., the first instances of the trial of the last error always had a higher proportion of disyllabic words). There also was a higher proportion of correct disyllabic word pronunciations in the trial of the first success. As evidenced by the item means presented in Table 3, there was not a perfect correlation between syllable length and learning efficiency (i.e., the subjects did not learn all the disyllabic words prior to learning the consonant cluster words). As regards the second of our original hypotheses, the results clearly indicate that post-vocalic clustered forms present significantly ($P < .01$) greater learning difficulties than initially clustered or non-clustered forms. Thus the learning results support the inferences that follow from the VCG conceptualization.

Now if the VCG and the behavioral propositions posited for it have psychological validity, one would wish to provide an account of reading errors as well as of efficient reading behavior. First, we would predict that errors should be intimately connected with failures to handle a unit VCG or a combination of VCG's. Consequently we would expect to find that the frequency of permutation, insertion, and reduction errors would be of the same magnitude as substitution (individual phoneme replacement) errors.

There were, in fact, 809 errors committed by the children. We categorize these as follows, using the items <u>hulig</u> and <u>borst</u> as examples for each error type:

1) <u>Omissions</u>: failure to emit a pronunciation.
2) <u>Permutations</u>: shifting a pair of consonants about the vocalic center (e.g., <u>luhig</u> and <u>brost</u>).
3) <u>Consonant Reductions</u>: deleting a consonant from the item (e.g., <u>huli</u> and <u>bost</u>).
4) <u>VCG Deletions</u>: deleting an entire VCG from the item (e.g., <u>lig</u>). Deletion of the VCG from the single VCG's would be considered an omission.
5) <u>VCG Insertions</u>: inserting a phonemic element to form a distinctly new VCG (e.g., <u>hoglig</u> and <u>borsit</u>).
6) <u>Familiar Words</u>: substituting a word (e.g., <u>holy</u> and burst).
7) <u>Vowel Substitutions</u>: substituting a single vocalic element (e.g., <u>hulog</u> and <u>birst</u>).
8) <u>Consonant Substitutions</u>: substituting a single consonantal element (e.g., <u>dulig</u> and <u>horst</u>).
9) <u>Miscellaneous Errors</u>: errors that we could not categorize or we were unable to analyze due to unintelligibility.

The results are presented in Table 5. We analyzed the data separately for the criterion and non-criterion groups as there were marked differences in total errors and relative percentages of errors in given categories between the two groups of subjects. If the categories of permutations, consonant reductions, VCG deletions, and VCG insertions are combined as representing errors directly related to VCG formation, one accounts for 50 percent of the results. This is, perhaps, more significant when one observes that single phoneme errors (vowel substitutions and consonant substitutions) constitute only 10 percent of the errors.

The categories of omissions and familiar words, representing 1/3 of the errors, suggest a major gap in the formulation to this point. These children, and speakers in general, appear to edit potential speech segments in some manner so as to exclude outputs that are not well-formed (Loban, 1963). The determination of well-formedness, whether in terms of phonological, grammatical or semantic criteria, is largely a product of prior language experience. It has been suggested that inherent characteristics may also play a part in such determination (Lenneberg, 1964). The concept of covert evaluation of potential speech units prior to pronunciation has been formulated into a behavioral feedback process called the "TOTE" by Miller, Galanter, Pribram (1960). The "TOTE" process, in essence, requires an internal recognition response prior to the emission of a pronunciation. If there is no recognition response, the would-be speaker re-cycles his processing of the potential output until either a recognition response does occur or the response interval is exceeded so that the potential output is represented as an omission.

In terms of our experiment, the familiar word errors occurred, we hypothesize, because certain children required a "TOTE" at the word level and so incorrectly generated a unit that was recognized as a word. Omissions may have occured through a similar strategy; instead of generating a familiar word that would be obviously incorrect, the "correct" pronunciation was aborted when no lexical match was found.

V. Heuristic Algorithm for Initial Reading

At this juncture, we wish to combine the formulations about the VCG and the recognition process into a set of ordered statements that, we believe, characterize a typical behavioral sequence in initial reading. This characterization, for the purpose of specificity, is written as a set of programmatic steps such as one might use for computer instructions. We will, in fact, utilize some of the terminology from computer programming. We disclaim any notion that there is an isomorphic relationship between computational hardware and/or routines and human neurological structure and/or functions known to be utilized in initial reading. We present this "heuristic algorithm" to illustrate how theory can be explicitly implemented but not to claim that either the theory or the implementation is presently a sufficient model for characterizing observed reading behaviors. We do contend that initial readers might use a heuristic algorithm of this nature. We have found such characterizations useful in generating hypotheses about initial reading and in planning a set of curriculum materials for initial reading.

We will list the steps in the algorithmic routine and then illustrate the routine by analyzing the words <u>fasting</u> and <u>pasting</u>.

<u>Step 1</u>. Input the orthographic unit (e.g., that string of orthographic symbols between spaces) into an immediate memory buffer. If the buffer is empty at any point of re-cycling, re-input the orthographic unit.

<u>Step 2</u>. Mark the vowels (e.g., a, e, i, o, u, y) in the unit.

<u>Step 3</u>. Count the vowels not followed by vowels. If the count is zero, branch to Step 1. If the count is one, proceed to Step 4. If the count is greater than one, apply the following syllabification rules to the medial consonants as follows:

a) ... VCV ... maps into V + CV mATIng → mA + TIng
... VCCV ... maps into VC + CV mATTIng → mAT + TIng
... VCCCV ... maps into VC + CCV thIRSTY → thIR + STY

If recycled to Step 3, apply these syllabification rules as follows:

b) ... VCV ... maps into VC + V BUSIng → BUS + Ing
... VCCV ... maps into V + CCV tASTY → tA + STY
... VCCCV ... maps into V + CCCV pASTRY → pA + STRY

<u>Step 4</u>. Decode each VCG separately and sequentially. Translate the orthographic code onto the code utilized in analyzing and processing individual VCG speech units. (This involves specifying the phonetic components for both the vocalic center and the pre- and post-consonantal clusters according to the constraints specified for any VCG of this articulatory type. In essence, this step translates the orthographic reading code onto the phonotactic speech code.

(mA → /mey/)

(TIng → /tiŋ/)

(mAT → /mæt/)

(TIng → /tiŋ/)

(thIR → /θɚr/)

(STy → /stiy/)

and so forth.

Step 5. Determine the category in the long-term memory area that is to be searched for a recognition response (that is, decide whether to search the word listings, the listings of experienced VCG forms, or, perhaps, the index of phonotactically permissible VCG forms). If a word listing search is to be performed, assemble the VCG units if there is more than one.

Step 6. Search for a match in the assigned long-term memory area. If a match fails to occur, branch to Step 3 and apply the re-cycling rules for syllabification.

Step 7. Put the matched VCG into the articulation buffer. Assemble the sequence of VCG's. Apply the rules for integration of phonetic units at VCG boundaries. Execute the motor commands for pronunciation.

It would perhaps be useful, before proceeding, to comment in a bit more detail on the notions underlying Steps 3 and 4. We have suggested previously and now state as theorems the following:

Theorem 1: A consonant cluster of length n-1 is simpler to pronounce than one of length n.

Theorem 2: An initial sequence of length n is simpler than a final sequence of length n.

At Step 3 we cut the orthographic string according to the criterion of simplicity. Thus, for orthographic strings of the form C^iVC^mV ... where $\underline{i} = 0, 1, 2, 3$ and $m = 1$, we cut $C^iV + C^mV$ ... For orthographic strings of the form C^iVC^mV ... where $\underline{i} = 0, 1, 2, 3$, and $m \geq 2$, we cut $C^i\ VC^F + C^i\ V$ ... such that $C^i \geq C^f$. Note that these are similar to traditional orthographic rules for syllabification based on the "length" of the first vowel and its relative stress. Note also that words containing doubled consonants will be cut between the consonants - *happy* → *hap* + *py*; *running* → *run* + *ning*. Note finally that a single segmentation procedure will not permit us an initially correct solution for, say, both *cathode** (*cat* + *hode*) and *cathouse* (*cat* + *house*); thus the recycling requirement within the heuristic algorithm.

In Step 4 we speak of the mapping rule for individual VCG. In our examples the strings to the left of the arrow are graphemic inputs while the right hand strings are the VCG outputs of this step. We conceive of the actual processing within Step 4 in terms of a recursive generator which, taking one graphic element at a time, assigns a set of phonetic features to that element. The application is recursive in that as each new element is introduced, the total sequence is reprocessed. This expansion processing continues until no graphic inputs remain. By definition the product of this processing is a well-formed VCG.

Note that this mechanism provides resolution of graphic difficulties of several different types. For instance since neither the full phonetic expansion of *kn-* → */kn/ (e.g., *knock*) nor *-mb* → */mb/ (e.g., *thumb*)) is tolerated in English phonotactics, the recursive rules yield a phonetic realization of only the element nearest the vocalic center; similarly for doubled final consonants - tiff, pass, hill, etc. Re-writes of the form th → */th/, sh → */sh/ ph → */ph/ are similarly prohibited. In these latter cases, however, the realization is phonemically distinct from either member. Phonetically it may be economical to consider that the non-initial graphemic h signals manner ("friction") and that its preceding consonant signals place (labial, dental, etc.) of articulation. Exactly what grapheme is to be considered first in this processing and exactly what direction the processing should assume is a matter of some conjecture at this point in our research. We should also point out that these rules are not the ones that will determine which of two well-formed phonetic interpretations is correct for an ambiguous sequence (e.g., read → /riyd/ or /red/). One would require, we assume, that the possibility of ambiguity in this string would also be signaled by the generator.

Now let us process the words fasting and pasting in order to illustrate how the algorithm operates. After inputting the orthographic image for fasting into the immediate memory buffer, one marks the two vowels. Step 3 is applied and fasting is segmented into fas + ting. After mapping the orthographic code onto the phonotactic code by Step 4, yielding /fæs + tiŋ/, it is decided to search the word listings for a match. For most adults, a match occurs and the word will be pronounced The word pasting, in the first cycle, yields a partition into the units pas + ting. At Step 4, this will be given the phonologic shape /pæs + tiŋ/. Obviously this form will fail in the word listing search. If the original unit is re-cycled, the word will now be cut pa + sting yielding /pey + stiŋ/ at Step 4. This phonologic form will lead to a match, and consequently, a pronunciation.

A direct inference from our heuristic algorithm is that words like pasting will show longer latencies in processing than will words like fasting. At the moment, no such latency data is available. We do have experimental data from five-year old children that document the close relationship between speech or orthographic decoding and speech production that we assume in the heuristic algorithm. According to the model, if a child has difficulty in detecting a mispronounciation of a word, then he should have difficulty in learning to read the word since Steps 4 and 5 of the algorithm require a mapping and matching of well-defined pronunciation units.

We pre-trained 20 five-year old children to accurately point to cells containing pictures of a top, a man, a pair of shoes, and a wall according to which word was pronounced. The children pointed to a blank cell for all words other than top, man, shoes, and wall. We then systematically interspaced twelve correct pronunciations with twenty-nine mispronunciations that involved a minimal feature shift of either a vocalic or a consonantal element (e.g., /tɔp/, /n · en/, /suwθ/, /yɔl/, etc.). The child had unlimited time in which to point to one of the pictures or to the blank depending on his interpretation of the given pronunciation.

None of the children committed an error on the twelve correct pronunciations. Moreover, they committed all of their errors by pointing to the picture associated with the given mispronunciation rather than to the blank square. As evidenced in row one of Table 6, the children performed significantly better than chance (P = .20) in detecting mispronunciations.

One week later we taught the same children to read the four words using the paired-associate anticipation paradigm described earlier in this paper. In row two of Table 6, you will find the mean total errors for each item. Notice the monotonic relation between the mean perception errors and the mean reading errors. Moreover the correlation between the mean error proportions from the perceptual task and the mean total errors from the reading task are substantial (see row 3, Table 6). We interpret the relationship between the two types of data to indicate that the children employed common codes and behavioral algorithms in speech processing and initial reading. This is, of course, the contention of our heuristic algorithm for initial reading.

Let us re-emphasize that the experimental evidence strongly suggests an independent or detachable phonologic processing loop with a VCG-like unit constituting its highest construct. Some contemporary linguistic theorists have claimed that any input perceived as language will necessarily be processed at the highest semantic and syntactic levels; others claim, less extremely, that any such input must be processed at the morphemic level. Postman and Rozenzweig (1956) conducted a series of experiments in which nonsense-syllables (e.g., int), morphemic units (e.g., ing), and words (e.g., ink), were tachistoscopically shown to a group of adult subjects. In the summary of results, they observed: " ... the failure of the English words to yield lower thresholds than the nonsense syllables suggests that S is no less ready to use syllables as response units than he is English words of comparable linguistic frequency." This is not to claim that reading should or could be restricted to processing at this level. It is to claim that any reading system which ignores the existence of such processing must fail to efficiently utilize an important language capacity of the child reader.

Even if persuaded of the viability of the heuristic algorithm, one might legitimately inquire if the requirement for the VCG can be justified at psycholinguistic levels presumably higher than that of phonologic decoding and encoding. More specifically, does the evidence on search processes in the long-term memory structure, be it conceived as a word association hierarchy or a conceptual space structure, reveal any manifestation of the VCG?

Recently, Roger Brown and David McNeil (1965) have investigated the behavioral nature of near word recall phenomena where a person cannot exactly recall a given word but feels that the word is on the "tip of the tongue." Using low frequency words from the Lorge-Thorndike list, the experimenters read definitions of these words to college students and asked them to try to recall the words. The subjects were requested to state the number of syllables plus the initial letter for those words that they couldn't quite recall but felt were on the "tip of their tongues." For the proven instances of near recall, the college students were able to give the correct number of syllables with 60 percent accuracy and the correct initial letter with 57 percent accuracy. Brown and O'Neil interpret these findings as indicating that the structural properties of a word (number of syllables, etc.) are part of the code the subjects use in locating words. We maintain that these structural properties can best be characterized by properties of the VCG. It would appear, in any case, that some VCG-like unit exists at these higher levels of long-term memory.

A few parenthetical comments about our research strategy will, perhaps, illuminate our reasons for developing such a simplified formulation of initial reading. We feel that employing a minimal amount of theoretical machinery allows one to be explicit as to the behavioral sequence. Moreover, one has the added benefit of being able to formulate detailed hypotheses about sequencing in the reading curriculum. The great disadvantage resides in knowing that further theoretical formulations would offer greater scope to the description and application. Thus we know that prior to developing an algorithm for sentence comprehension we might want to consider at least three implementations: 1) a prosodic coding system that would account for intonational information like stress, 2) a syntactic analyzer in the form of a generative and/or transformational rule set that would allow for "chunking" of the word units contained in the sentence, and 3) a semantic processor that would relate these units to the child's conceptual structure. Since we have restricted ourselves here to considering only teaching children to "read" words in the pronunciation sense, we consider these future tasks which we ultimately will attempt to formulate and resolve.

If one accepts our views about the viability of the VCG and the heuristic algorithm as psycholinguistic units and processes in initial reading, one might legitimately inquire if such a conceptualization provides any reasonable implications for specifying the nature and internal sequence of a reading curriculum. We believe there are many direct applicational inferences that, moreover, are empirically testable. First, the VCG provides a new and explicit way of classifying the initial vocabulary of reading according to consonant-vowel-consonant sequences in the vocabulary (i.e., one can classify the words according to the number of VCG's, the structure and position of consonant clusters within the VCG and the type of phonotactic constraints evidenced in the VCG's). Although consonant-vowel pattern classification schemes are far from novel in reading pedagogy (Bloomfield and Barnhart, 1959), our hypothesis about the learning difficulties of a given pattern provides a way of ordering the materials along a continuum of relative learning efficiency.

Perhaps an example treating a within-pattern analysis will illustrate our point. The ubiquitous consonant-vowel-consonant (CVC) pattern dominates the first hundred words of most current materials. How might one order these CVC words in order to maximize transfer effects (i.e., the behavior of generalizing from acquired words to reading new words) while one is commuting seemingly known consonant letters about the five simple vowels? If one happens to start with the words tap, nag, and man, what commutations would optimize the amount of transfer? We predict that those letter commutations that result in a minimal change in the VCG constraint statements of the acquired words should optimize transfer. Moreover initial consonant-vowel changes should be simpler than vowel-final consonant changes. Therefore we would predict the following orders of difficulty:

Acquired Words		Least Difficult				Most Difficult		
tap		tag		nap		gap		Pat
nag	→	tan	<	map	<	mag	<	Pam
man								

tag and tan would be least difficult because the children have already mapped orthographic codes onto the CVC forms and the commutation involves only the initial consonant-vowel relationship. Pat and Pam would be most difficult because the new vocalic-consonant combinations require major relearning or rewriting of the necessary VCG constraints. At the very least, the VCG conceptualization offers a testable hypothesis about optimal sequencing of such forms.

We are now preparing reading lesson materials that are designed to evaluate the hypotheses derived from the VCG conceptualization (i.e., the CVC transfer hypothesis and the consonant cluster hypothesis) and the heuristic algorithm. The experimental work will take place within the confines of a computer-based instructional system. The computer system provides for complete control of the stimulus material (audio and visual) via specified performance criteria embedded in the decision network of the program. A detailed recording of all student responses and response times is made on magnetic tape for off-line data analysis. These fine-grained behavioral histories will be of sufficient number and detail that the proposed hypotheses or alternative hypotheses can be evaluated. Let us for a moment consider some of the present and planned capabilities of this system.

VI. Description of the Stanford Computer-Based Instructional System.

This system consists of a medium-sized computer and six instructional booths. Each booth is a small 7' by 8' room that contains three input-output devices: (1) an optical display unit, (2) a cathode-ray tube display unit, and (3) an audio system. The main computer controls the presentation of the visual and auditory materials; the students respond with either a light-pen device, a typewriter keyset, or a microphone. The computer evaluates the responses and selects new audio and visual material according to the outcome of the evaluation.

The optical display unit is a rapid, random-access projection device that presents visual material to the student on a 10" by 13" screen. The source of the materials (any 8-1/2" x 11" page of text) is photographed on microfilm and is stored in a small projector cell that has a capacity of 256 pages. Since each display unit has two projectors, each instructional unit has a total capacity of 512 individual pages. Moreover, additional combinations are possible by constructing composite images from both projectors on the common screen or by using the shutter system that divides the screen into eight equal sections by various masking arrangements. The student responds to the display and sends information to the computer via a light-pen. As the pen is touched to the screen, the coordinates of the position are sent to the computer for evaluation according to predefined/redefined areas for correct and incorrect responses. Most of the evaluation operations of the optical display unit will occur in approximately one second.

The cathode-ray tube display unit, commonly called a "scope," can present any of 120 prearranged alpha-numeric characters or line vectors on a 10" x 10" screen. A light pen is available for sending information to the computer for evaluation via the specified coordinate system as described for the optical display device. In addition, a typewriter keyboard is attached to each scope and may be used to send information from the student to the computer.

The random-access audio system can play any prerecorded message to the student. The messages are recorded on a 6" wide magnetic tape. Two tape transports are available to each instructional booth. Each transport has a capacity of approximately 17 minutes that can be divided into any combination of message lengths from 1020 one-second messages to one message of 17 minutes duration. The student may record onto the tape and then have the recording played back for comparison purposes or retained as response data. The random-access time to any stored message is less than two seconds.

The various control and switching functions between the different input-output terminal devices are handled by a medium sized computer that has a 16,000 word core memory. An additional 4000 word core can be interchanged with any of 32 bands of a magnetic memory drum. All input-output devices are processed through a time-sharing program that services them only when necessary. Two high speed data channels permit simultaneous computation and servicing of terminal devices. Additional back-up in computational power, storage, and increased input-output speed is obtained through connections to a larger computer system located at the Stanford Computational Center.

We are also in the process of developing a similar computer-assisted instructional system that will be located in an elementary school. Sixteen instructional stations will be available on the school site. We plan to provide instruction for approximately 90 students per day with this system starting in the Fall of 1966.

Does a sophisticated computer system like the one just described provide the means for answering practical pedagodical questions? The number of repetitions required for mastery of a word represents one of many seemingly simple, direct, but unanswered questions in reading pedogogy. Gates (1930) attempted to answer the question in the typical classroom comparison design but the host of uncontrolled variables - teacher competency, student aptitude, etc., now known to affect these experimental situations casts serious doubt on his recommendation of "35" repetitions as the optimal number. Even to attempt an answer to the basic question requires many qualifications and specifications.

One must know the stimulus materials, i.e., the nature and difficulty of the words involved. One must know the nature of the presentation and reinforcement procedures. One must follow the course of acquisition, marking for further study the point of the first correct response and the points and nature of all subsequent errors. One must also account for individual differences in the acquisition process. Thus, one really requires a detailed description of the course of learning.

Mathematical learning models provide a method for concisely describing such a fine-grained sequential learning process. One might ask the practical question, do such models provide a way of telling us how many repetitions a given child will require to reach a given criteria of, say, "one correct cycle through the list of words without an error"? And, more important, can we gain the information about required repetition early enough in the lesson sequence to aid us in our scheduling of presentations?

One must first find a mathematical model that accurately describes the course of learning of a list of words. We hoped that the data from the disyllabic-consonant cluster paired-associate experiment might provide us with an exemplar. We fitted a number of mathematical models to the data and found one, the "one-element model," that yielded a reasonable account of the course of learning. It should be emphasized that the requirement for such models is that they fit the data. There is no suggestion that such a "fit" represents an "interpretation" of the data. Neither is it implied that the model will fit similar experiments with the same degree of accuracy.

The one-element model (OEM) is a special case of the more general models of Stimulus Sampling Theory (Atkinson and Estes, 1963). The model (OEM) specifies that the behavioral association between the given orthography of a word and its pronunciation is assumed to be in one of two states, unlearned (UL) or learned (L). Given a word that is in the unlearned state (UL), there is a constant probability, c, that the word is learned during each correction-rehearsal event. While in the unlearned (UL) state, the probability of a correct response is dependent on the number and availability of the set of responses or required pronunciations. In our calculations we set the guessing probability equal to one over the number of VCG's in the word list (i.e., the guessing probability was set equal to 1/18). This simple all-or-none learning process can be represented by the following transition matrix and response probability vector:

$$\begin{array}{ccc} & & \text{Trial } n+1 \\ & & \begin{array}{cc} L & UL \end{array} \\ \text{Trial } n & \begin{array}{c} L \\ UL \end{array} & \begin{bmatrix} 1 & 0 \\ c & 1-c \end{bmatrix} \end{array} \qquad \begin{array}{c} \text{Prob. of Correct Response} \\ \begin{bmatrix} 1 \\ 1/18 \end{bmatrix} \end{array}$$

In order to make a comparative evaluation of a number of learning models, an analysis was made of the sixteen possible sequences of correct and incorrect responses that result from considering blocks of four trials. These sixteen possible sequences, presented in the first column of Table 7, encompass a sufficient number of observations to be considered reliable and sensitive to the course of learning. Furthermore, closed theoretical expressions for these sixteen event sequences can be derived for a large number of learning models. A simple method for estimating the parameters for a given model consists of minimizing the X^2 function associated with the sixteen outcome (O) events. Let $Pr(O_i; c)$ denote the probability of the event O_i which, of course, is a function of the parameter c. Then one can define the function:

$$X^2(c) = \sum_{i=1}^{16} \frac{[TP_r(O_i; c) - N(O_i)]^2}{TP_r(O_i; c)}$$

where $N(O_i)$ denotes the observed frequency of outcome O_i and $T = N(O_1) + \ldots + N(O_{16})$. Using a high-speed computer that is programmed to systematically scan grids of possible parameter values, one selects the estimate of c that minimizes the X^2 function. Under the null hypothesis this minimum X^2 has the usual limiting distribution with the degrees of freedom equal to 16 minus the number of estimated parameters minus one.

The model (OEM) assumes that all items are stochastically independent and identical (i.e., all pairs are of equal difficulty and all subjects learn at the same rate). Knowing that the disyllabic and consonant cluster words were significantly different, we performed separate analyses for each list. The results for trials 1 through 4 are presented in Table 7. The X^2 values with 14 degrees of freedom are non-significant so that one accepts the null hypothesis that there is no statistical difference between the observed frequencies and the frequencies predicted by the one-element model. The analyses of trials 5 through 8 yielded even lower X^2 values. The fit of the model is surprisingly good considering the assumption of no individual differences in the rates of learning. Thus we have at hand a quantitative model that will allow us to predict the necessary amount of repetition required for list mastery.

The appropriate measure for determining the required repetitions is the mean trial of the last error. We predicted the mean trial of the last error for each subject for each of the two types of words by the following procedure:

1) Find the trial of first success (i.e., the trial that first indicated possible learning).

2) Estimate the learning parameter, c, by the

equation $$\hat{c} = \frac{(\text{Proportion Correct}) - g}{1-g}$$

(This equation follows directly from the one-element model).

3) Use the estimate of c to predict the trial of the last error by the equation

$$E(\text{Trial of the Last Error}) = \frac{\frac{(1-g)}{(1-c)(1-g) + c}}{c}$$

(This expectation was derived by Bower (1961) for the one-element model).

4) Calculate the observed mean trial of the last error.

The mean of the absolute value of the difference between the observed and predicted mean trial of the last error yields 1.31 trials for the disyllabic words and 2.65 trials for the consonant cluster words. The item means (see Table 4) indicate the greater heterogeneity of the consonant cluster words; consequently one would expect a poorer fit between the observed and predicted values for these items. Given this reservation, one can predict with reasonable accuracy the number of trials a child will need to cycle through a list of words in order to reach the mean point of mastery, (i.e., the mean trial of the last error).

Now all of these calculations can be performed on-line by a computer while a child is learning a list of words. If one should use a static optimatization scheme that calls for recycling a child through a list until the predicted mean trial of the last error plus one standard deviation of this statistic are reached, would the child attain a criterion of mastery such as one complete errorless cycle of the list? In the disyllabic-consonant cluster paired-associate experiment twelve out of the thirteen criterion subjects would have fulfilled this mastery criterion. Moreover, five of these twelve subjects would have been branched out of the task exactly at the end of the trial where they attained the criterion.

We wish to reiterate that this is only an example of how a computer-based instructional system can be used to solve practical pedogological questions. Extensive exploration will be necessary before we will have an adequate mathematical model of initial reading that can handle problems of retention over delay periods, interference between differing lists of words, and so on. We remain optimistic, however, that computer-based instructional systems will ultimately provide a setting where practical problems of initial reading can be solved by quantitative methods.

VII. Summary

Our lengthy chronicle ceases at this point. If this report appears a bit kaleidoscopic in nature, this ensues from the inter-disciplinary backgrounds and interests of our group. Each of the representatives from linguistics, psychology, and education bestows insights from his particular fund of knowledge. This interaction leads to specific enterprises that range on Hilgard's (1964) suggested continuum for educational research from pure scientific discovery to new pedogogical innovations. The substance of the paper reflects these diverse but, we believe, coordinated efforts.

The Russian study provided us with our initial hypothesis about the hierarchy of rehearsal difficulty, resulting from consonant clusters. We then investigated the factors in speech perception and production that might account for this hierarchy of rehearsal difficulty. The Liberman, et al., (1959) work on speech perception, provided us with insights into the requirements for "meshing" successive phonemes in order to produce intelligible pronunciations. Using implications of the "rules for speech synthesis" and of our investigations into the phonotactics of English, we formulated the "Vocalic Center Group" as the psycholinguistic unit for initial reading. After investigating an implicit assumption regarding rehearsal and reading performance, we used an analysis of the reading errors observed in this investigation to formulate a "heuristic algorithm" specifying a processing system and an informational flow to simulate an initial reading performance. Syllabification, orthographic to phonemic translating, VCG construction and concatenation, memory storage, and recognition-matching were assumed operational components in the "heuristic algorithm." We then verified our inference from this schema that speech recognition errors for given words should be monotonically related to initial reading errors for these same words.

After citing some findings from Brown and McNeil suggesting the requirement for VCG-like units in long-term memory processes, we illustrated how the VCG provides some detailed hypotheses about constructing and sequencing the beginning reading curriculum. We described a computer-based instructional system that provides us with an automated pedagogical environment within which to rigorously test our predictions. We concluded by citing an example of how the computer-based instructional system and quantitative models of learning might offer answers to specific questions as to, for example, the amount of word repetition necessary for a specific criterion of mastery.

Our interests and activities do span a number of disciplines but our enterprise retains focus, we feel, by the task we have set for ourselves: the development of the most simple but adequate learning model for initial reading. Just as models of a language learner require less complexity than models of a competent language user, so our own model attempts specification of the minimal capabilities of the beginning reader without speculation as to the maximal capacities of the mature reader. The complexity of our theoretical formulation is still several orders of magnitude removed from an adequate "Theory of Reading." But to quote from a philosopher of science, Brodbeck (1963), such an approach warrants its user

> " ... that instead of helplessly gazing
> in dumb wonder at the infinite complexity
> of man and society, he has knowledge,
> imperfect rather than perfect, to be sure,
> but knowledge not to be scorned nonetheless ... "

TABLE I

FORM CLASS ANALYSIS OF ENGLISH TRANSLATIONS OF THE 300 RUSSIAN WORDS

Form-Class	Number of Words	Mean Proportion Correct per Item	Examples of Russian English Pairs	
Definite pronouns	5	.93	ВЫ	you
Colors	4	.67	БЕЛЫЙ	white
Concrete Nouns	84	.58	ЛОБ	forehead
Possessive Pronouns	3	.42	ОНА	her
Prepositions	11	.42	ЧЕРЕЗ	across
Adjectives	57	.41	ДИКИЙ	wild
Numbers	8	.37	ПЯТЬ	five
Abstract Nouns	34	.36	ВЛАСТЬ	authority
Interrogatives	7	.31	ПОЧЕМУ	why
Adverbs	22	.23	ПОЧТИ	almost
Conjunctions	6	.22	ИЛИ	or
Verbs	59	.21	ДЕРЖАТЬ	to hold
	300			

TABLE II

GROUP STATISTICS FOR THE RUSSIAN-ENGLISH REVERSAL STUDY

Group Statistics	Russian Words Low Difficulty	Russian Words High Difficulty
Mean Proportion of Pair Items Correct in the Suppes-Crothers Experiment	.454	.015
Mean Proportion of Pair Items Correct in Reverse Pair Experiment	.299	.118
Number of Syllables in the Russian Words	46	48
Sum of the Consonant Cluster Complexity Scores	22	50

TABLE III

THE MEAN TOTAL ERRORS COMMITTED BY THE CHILDREN ON THE

TWELVE ENGLISH PSEUDO-WORDS

Disyllabic Item	Total Group Mean Total Errors N = 20	Criterion Group Mean Total Errors N = 13
hulig	3.10	1.54
molin	1.85	.62
devan	2.30	.92
renad	3.50	1.69
nateb	2.55	1.23
fegom	2.45	.85
Consonant Cluster Items		
strem	3.25	1.39
splog	2.90	1.46
brind	3.85	1.46
flesk	3.25	1.23
borst	4.05	2.39
vimpt	4.30	3.08

TABLE IV

GROUP STATISTICS FOR DISYLLABIC AND CONSONANT CLUSTER ITEMS

	Disyllabic Words	Consonant Cluster Words
Mean Total Errors S.D.	2.79 3.24	3.60 3.33
Trial of the Last Error S.D.	3.06 3.40	3.95 3.71
Successes Before the Last Error S.D.	.27 .67	.35 .89
Mean Errors Before the First Success	2.43	3.04

TABLE V

CATEGORIZATION OF THE READING ERRORS FOR THE DISYLLABIC AND CONSONANT CLUSTER ITEMS

Error Categories	Disyllabic Items	Criterion Group (N = 13)	Percentage of Errors	Non-Criterion Group (N = 7)			Total Group
	Disyllabic Items	Consonant Cluster Items	Percentage of Errors	Disyllabic Items	Consonant Cluster Items	Percentage of Errors	Percentage of Errors
Omission	3	2	2%	58	63	21%	13%
Permutations	8	9	10%	8	11	4%	4%
Consonant Reductions	3	30	14%	8	41	8%	10%
VCG Deletions	8	0	3%	23	0	4%	4%
VCG Insertions	30	36	29%	92	97	33%	32%
Familiar Words	4	45	21%	41	91	23%	23%
Vowel Substitutions	11	8	8%	15	9	4%	5%
Consonant Substitutions	8	6	6%	6	8	2%	5%
Misc. Errors	10	5	7%	4	2	1%	4%
Number of Errors	85	147	232	255	322	577	809

TABLE VI

MEAN RESULTS OF THE SPEECH PERCEPTION-INITIAL READING EXPERIMENT

Group Statistics	top	man	shoes	wall
Mean Proportion of Detection Errors of Mispronunciation	.192	.283	.293	.490
Mean Total Errors	1.45	2.00	2.70	4.50
Coefficient of Correlation Between Perception Errors and Reading Errors.	.73	.60	.75	.85

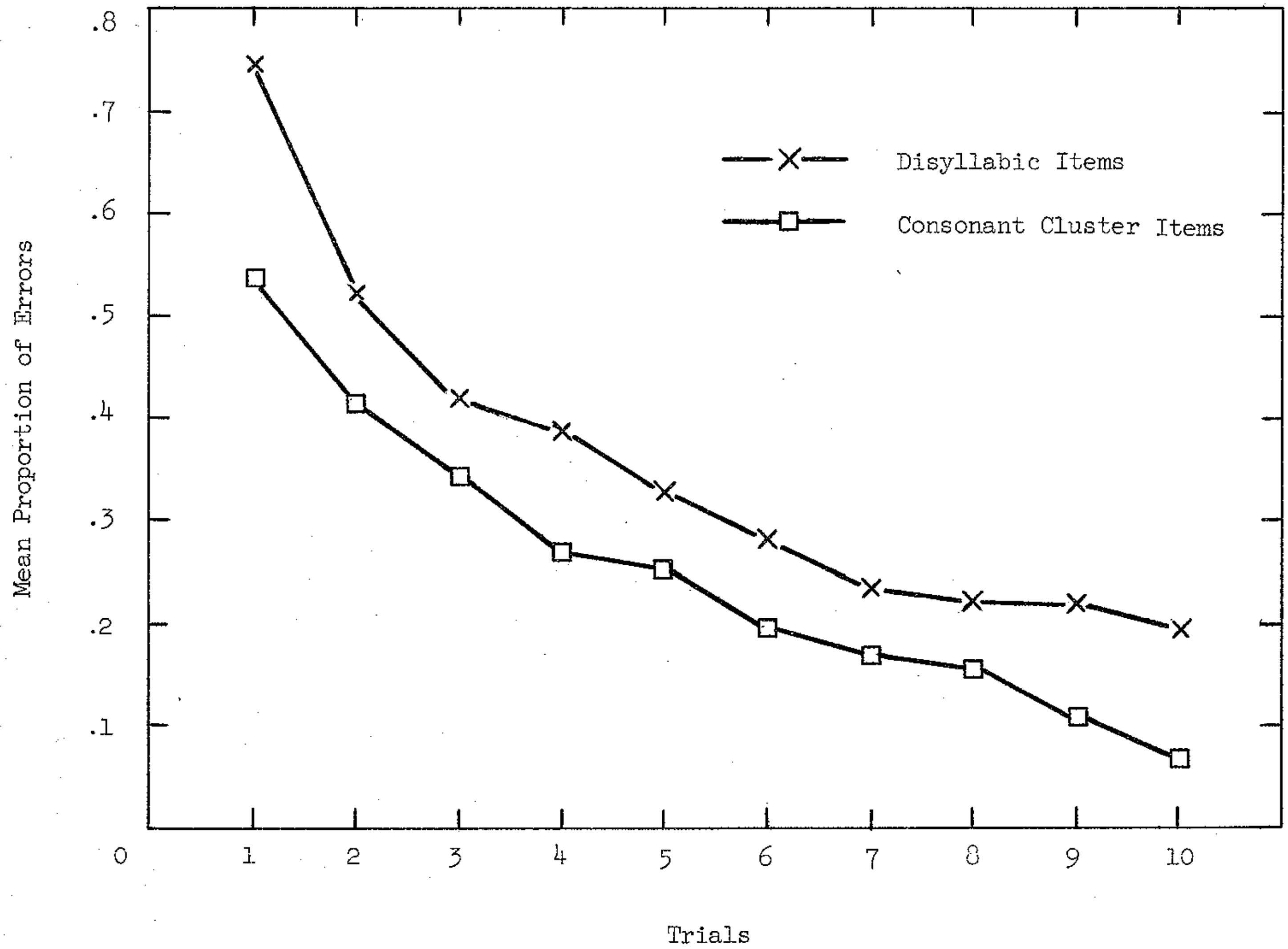
TABLE VII

THE FIT OF THE ONE-ELEMENT MODEL TO THE DISYLLABIC CONSONANT CLUSTER PAIRED-ASSOCIATE DATA

Event Sequence	Disyllabic Words		Consonant Cluster Words	
	Observed Frequency	Predicted From OEM	Observed Frequencies	Predicted From OEM
cccc[1]	38	31.3	21	24.5
ccce	0	.0	0	0
ccec	0	.0	0	0
ccee	0	.1	0	.1
cecc	2	.9	0	.8
cece	0	.1	0	.1
ceec	1	.7	0	.7
ceee	0	1.8	0	2.4
eccc	20	22.2	22	18.7
ecce	1	.1	1	.1
ecec	2	.7	0	.7
ecee	3	1.8	1	2.4
eecc	17	15.8	21	14.3
eece	3	1.8	5	2.4
eeec	6	12.4	12	12.6
eeee	27	30.2	37	40.0
	$\sum = 120$	$X = 20.34$ $c = .250$	$\sum = 120$	$X^2 = 18.65$ $c = .195$

Figure 1

Learning Curves for Disyllabic and Consonant Cluster Pseudo-Words



REFERENCES

Atkinson, R. C. and Estes, W. K., Stimulus sampling theory. In Luce, R., Bush, R. and Galanter, E., Handbook of Mathematical Psychology, Vol. II. New York: Wiley, 1963, Chapter 2.

Bloomfield, L., Language. New York: Holt, 1933.

Bloomfield, L. and Barnhart, C., Let's Read: A Linguistic Approach. Detroit: Wayne State Univ. Press, 1961.

Bower, G., Applications of a model to paired-associate learning, Psychometrika, 1961, 26, 255-280.

Brodbeck, M., Logic and scientific method in research on teaching. In Gage, N. (Ed.), Handbook of Research on Teaching, Chicago: Rank McNally, 1963, 44-94.

Brown, R. and McNeil, D., The "Tip of the Tongue" phenomena, Beh. Sci. in press.

Chomsky, N., Formal properties of grammars. In Luce, R., Bush, R. and Galanter, E., Handbook of Mathematical Psychology, Vol. II. New York: Wiley, 1963, 269-322.

Chomsky, N., Current issues in linguistic theory. In Fodor, J. and Katz, J. (Eds.), The Structure of Language. Englewood Cliffs, New Jersey: Prentice-Hall, 1964, 50-118.

Carroll, J., The analysis of reading instruction: perspectives from psychology and linguistics. In Hilgard, E. (Ed.), Theories of Learning and Instruction, N.S.S.E. 63 Yearbook. Chicago: Univ. of Chicago Press, 1964, 336-353.

Fries, C., The Structure of English. New York: Harcourt, Brace, and World Inc., 1952.

Fries, C., Linguistics and Reading. New York; Holt, 1963.

Gates, A., Interest and Ability in Reading. New York: Macmillan, 1930.

Glanzer, M., Grammatical category: a rote learning and word association analysis, J. Verb. Lrn. Verb. Beh., 1962, 1, 31-41.

Greenberg, J. and Jenkins, J., Studies in the psychological correlates of the sound system of American English, Word, 1964, 20, 157-177.

Greenberg, J., Some generalizations concerning initial and final consonant sequences, Voprosy Jazyhoznanija, in press.

Halle, M. and Stevens, K., Speech recognition: A model and a program for research. In Fodor, J. and Katz, J. (Eds.), The Structure of Language. Englewood Cliffs, New Jersey: Prentice-Hall, 1964, 604-612.

Hilgard, E., A perspective on the relationship between learning theory and educational practices, Theories of Learning and Instruction, N.S.S.E. 63 Yearbook, Chicago: Univ. of Chicago Press, 1964, 402-415.

Harris, C., Study of the building blocks of speech, J. Acoust. Soc. Amer., 1953, 25, 962-969.

Harris, Z., Structural Linguistics. Chicago: Univ. of Chicago Press, 1951.

Haugen, E., The syllable in linguistic description. In *For Roman Jakobson*, The Hague: Mouton, 1956, 213-221.

Hockett, C., *A Manual of Phonology*. Baltimore: Waverly Press, 1955.

Jenkins, J. and Palermo, D., Mediation processes and the acquisition of linguistic structure, *Monogr. Soc. Res. Child*. Dev., 1964, 29, 141-168.

Lenneberg, E., The capacity for language acquisition. In Fodor, J. and Katz, J. (Eds.), *The Structure of Language*. Englewood Cliffs, New Jersey: Prentice Hall, 1960, 579-603.

Liberman, A., Ingemann, F., Lisker, P., Delattre, P., Cooper, F., Minimal rules for synthesizing speech, *J. Acoust. Soc. Amer*., 1959, 31, 1490-1499.

Loban, W., *The Language of Elementary School Children*. Champaign, Illinois: National Council of Teachers of English, 1963.

McKee, P., Harrison, M., McCowen, A., and Lehr, E., *Teacher's Guide for Tip*. Boston: Houghton Mifflin Company, 1963.

Miller, G., Galanter, E., and Pribram, K., *Plans and the Structure of Behavior*. New York: Holt and Co., 1960.

Miller, G., Some preliminaries to psycholinguistics, *Amer. Psychol*., 1965, 20, 15-20.

Morris, R., *Success and Failure in Learning to Read*. London: Oldbourne Book Co., 1963.

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

Osgood, C. and Sebeok, T., (Ed.), Psycholinguistics: A Survey of Theory and Research Problems. Baltimore: Waverly Press, 1954.

Peterson, G., Wang, W., and Sivertsen, E., Segmentation techniques in speech synthesis, J. Acoust. Soc. Amer., 1958, 30, 739-742.

Pike, K., Phonemics. Ann Arbor: Univ. of Mich. Publ., 1947.

Postman, L. and Rosenzweig, M., Practice and transfer in the visual and auditory recognition of verbal stimuli, Amer. J. Psychol., 1956, 69, 209-226.

Rosen, G., A dynamic analog speech synthesizer, J. Acoust. Soc. Amer., 1958, 30, 201-209.

Russell, D. and Ousley, O., Manual for Teaching the Pre-Primer Program. New York: Ginn and Company, 1957.

Saporta, S., Certain characteristics of phoneme sequences. In Osgood, C. and Sebeok, T., (Ed.), Psycholinguistics, Baltimore: Waverly Press.

Stetson, R., Bases of Phonology. Oberlin, Ohio: Oberlin Press, 1945.

Suppes, P. and Crothers, E., Second Language Learning, in preparation.

Templin, M., A study of sound discrimination ability of elementary school pupils, J. Speech Disorders, 1943, 8, 127-132.

Underwood, B. and Schulz, R., Meaningfulness and Verbal Learning. Chicago: Lippincott and Co., 1960.

Whorf, B., Linguistics as an exact science, Technol. Rev., 1940, 43, 61-63, 80-83.

Zipf, G., Human Behavior and the Principle of Least Effort, Cambridge, Mass. Addison-Wesley, 1949.