DOCUMENT RESUME

ED 028 415 AL 001 764

By-Brown, Dennis Lee
Identification of American English Initial /l/ and /r/ by Native Speakers of Japanese.
Pub Date Feb 69
Note-85p.; Expanded version of paper presented at the Third Annual TESOL Convention, Chicago, Illinois, March 5-8, 1969.
EDRS Price MF-$0.50 HC-$4.35
Descriptors-*Acoustic Phonetics, *American English, *Auditory Perception, Consonants, English (Second Language), *Interaction Process Analysis, *Japanese, Phonology, Second Language Learning, Syllables, Vowels

Six hundred and sixteen CVC stimulus syllables beginning with /l/ and /r/ were presented by tape recording to six Japanese foreign students at Indiana University. Analysis of variance showed the effect of the final consonant variable on the identification of initial /l/ and /r/ was significant at the .05 level; the effect of the CV interaction was significant at the .01 level. Response scores for the CV interaction pairs correspond closely to the magnitude of the third formant shift from /l/ and /r/ to the various syllable nuclei determined by Lehiste. Specific results, correlations with frequency of occurrence and acoustic data, and implications for improving /l/ and /r/ identification by native speakers of Japanese are discussed. (Author/DO)

IDENTIFICATION OF AMERICAN ENGLISH INITIAL /l/ AND /r/ BY NATIVE SPEAKERS OF JAPANESE

FEBRUARY 1969

Dennis Lee Brown

IDENTIFICATION OF AMERICAN ENGLISH INITIAL /l/ AND /r/ BY NATIVE SPEAKERS OF JAPANESE

Dennis Lee Brown

Southwest Regional Laboratory
11300 La Cienega
Inglewood, Calif. 90304

This research was submitted in partial fulfillment of the M.A. in Linguistics, Indiana University, February 1969.

ACKNOWLEDGMENTS

I would like to thank Dr. Fred W. Householder of Indiana University for advice and criticism in the writing of the thesis. For advice and consultation on statistical design and processing I would like to thank Dr. Vincent DiLollo formerly of Indiana University and Dr. Ludwig Mosberg of the Southwest Regional Laboratory.

CONTENTS

Identification of American English
Initial /l/ and /r/
By Native Speakers of Japanese

Abstract

616 CVC stimulus syllables beginning with /l/ and /r/ were presented by tape recording to six Japanese foreign students at Indiana University. Analysis of variance showed the effect of the final C variable on the identification of initial /l/ and /r/ was significant at the .05 level; the effect of the CV interaction was significant at the .01 level. Response scores for the CV interaction pairs correspond closely to the magnitude of the third formant shift from /l/ and /r/ to the various syllable nuclei determined by Lehiste. Specific results, correlations with frequency of occurrence and acoustic data, and implications for improving /l/ and /r/ identification by native speakers of Japanese are discussed.

Introduction

Extensive knowledge of the phonology, syntax, culture, and values of the society makes much of the speech information transmitted between two native speakers of a particular language redundant. Because of redundancy, total comprehension can take place when the surface constructions are ambiguous and acoustic interference or attenuation seriously distorts much of the information originally transmitted by the speaker. The loss of a single phonemic or syntactic distinction would rarely have any effect on the comprehension of a native speaker.

The non-native listener does not have the same mastery of phonology or syntax, nor the knowledge of the culture and the values of society. Thus, much of the speech information that would have been redundant for a native listener would not be so for a non-native listener. The non-native listener is either unable to decode, or decodes incorrectly much of the phonological and syntactic information. In addition, the non-native speaker's lack of knowledge of the culture and values of the society make more information necessary to assure total comprehension of the message. As a result, surface ambiguities (ambiguous surface constructions) and loss of information through acoustic interference and attenuation cause even more difficulty in comprehension for the non-native listener than for the native listener. The loss of even a single phonemic or syntactic distinction could readily lead to a comprehension breakdown.

The ultimate goal of the second-language learner desiring mastery would be the perception and interpretation of the same number of cues in the same way as the native speaker. Since it is a very long-term project to provide the non-native speaker with extensive experience and knowledge of the culture and values of the society, initial efforts to improve comprehension mastery are better concentrated on the perception (decoding) of phonological and syntactic cues. Basic to the decoding of syntactic cues, is the decoding of phonological cues. Even though syntactic cues may be used to reconstruct phonological units as well as visa versa, there must first be some phonological cues in order to detect any syntactic information.

The approach to designing a second-language instructional program is often limited to:

1) The preparation of a comparative analysis of the two languages in search of areas of possible interference. Formerly, this was done more often by comparison of phonemic analyses and syntactic taxonomies of the two languages. More recently, investigators have taken to comparing the various phonological and syntactic transformational grammars of each particular language.

2) Alternatively, "skeptics" of comparative analysis not recognizing a great contribution to instruction through

such an approach, have often limited themselves to uncontrolled superficial observations of certain second-language learning problems in the classroom. Such casual observations as, "The Japanese cannot pronounce /l/," might be expected.

Although it is understandable that "skeptics" find little practical use for comparative analysis in devising and designing a meaningful instructional program for these problems, at least as comparative analysis is commonly applied, it is frankly difficult to see how casual recollections of classroom experience could be of any greater value.

While comparative analysis arrives at a prediction of problems and problem areas in second-language learning, these predictions are obviously worthless if they do not correspond to the problems the students actually have. The prediction from a comparative analysis that speakers of German would experience confusion between American English (AE) /t/ and /θ/ when in fact a few minutes observation of a learner would indicate a confusion between AE /s/ and /θ/, is plainly a waste of time.

The implications of various parts of a confirmed prediction are also unclear. The comparison of two analyses carried out according to certain theory-specific adequacy and economy criteria deleting certain redundant information, cannot be expected to arrive at conclusions which are the result of the interaction of all of the learner and language facts.

On the other hand, casual observations of the "skeptics" afford no indications of underlying factors of the learning problems. They seem only to provide fertile seed beds for the growth and spread of personal peeves and irresponsible speculation.

These remarks are not to be taken by any means as a total condemnation of the proponents of comparative analysis or the "skeptics." Proponents of comparative analysis can provide the linguistic facts and the system for utilizing these facts; the "skeptics" can provide the observations which can validate or reject basic premises and thus activate the system for utilizing these facts. As pointed out by Anderson (1964), investigation which considers the basic linguistic facts and experimentally takes the learner into account each step of the way will avoid the pitfalls of counterfactive projection and overgeneralization from superficial observation, and thus can contribute valuable information for use by instructional program designers.

Those coming in contact with Japanese who speak English as a second language readily notice that the Japanese often pronounce the AE /l/ and /r/ the same to AE ears. Less obvious to non-Japanese is the difficulty the speakers of Japanese experience in perceiving AE /l/ and /r/. Initial pilot tests indicated that the Japanese had considerably less success in identifying /l/ and /r/ produced by a speaker of AE than the speakers of AE had in identifying /l/ and /r/ produced by the speakers of Japanese. Nakajima (1957) explicitly mentions that it is "next to impossible for Japanese speakers of English to discriminate them in hearing..."

Because of the high degree of confusability of AE /l/ and /r/ for Japanese exchange students, the relatively high frequency of occurrence and information load, and the fact that most of the Japanese exchange students can scarcely tolerate such loss of information due to problems in syntax and lack of experience in the culture, extensive testing of the identification of AE /l/ and /r/ by Japanese was chosen first for investigation.

The use of "allophone" in this paper has been limited to those phoneme variants, distributionally described, that are inherently different from one another in terms of a feature. Thus, the final [l] and initial [l] are two allophones of phoneme /l/, distributionally predicted, that differ with respect to certain, let us say, articulatory features. Yet, they are certainly the same in many others.

In addition to "distributional" allophones there are "contextual" variants. As pictured here, a "contextual" variant would be the manifestation of a particular "distributional" allophone in contiguity with other phones. The (target) articulatory features at the point of maximum closure would be the same for two "contextual" variants. As the context for two "contextual" variants differs by definition, the transition either from the preceding phone or to the following phone would necessarily be different. For example, the transitions from [l] to [o] are quite different from those to [e] (cf., Figure 13). Yet, the starting Fs for the two are very nearly the same (cf., Figures 1 and 12).

Since the transitions to and from a particular consonant phoneme, such as /l/, are considered part of that phoneme, and play a valuable role in the detection and perception of the consonant phoneme, it follows that there are variants from context to context, independent of the articulatory position of maximum closure. (In the case of stops such as [t], there is at most only the sound of air turbulence--nothing at all if the [t] is unaspirated. The only perceivable portion in that case is the transition (Liberman, Delattre, Cooper, & Gerstman, 1954).

Initial pilot checks indicated that Japanese students seemed to have less trouble in identifying final /l/ and /r/. /l/ and /r/ in cluster, including the most extreme allophonic variation, seemed to cause all the students so much trouble that there would probably be little variation of identification scores from context to context. In order to maintain the complexity of the experimental design within reasonable limits, the investigation was limited to initial

/l/ and /r/ until the effects of context had been determined.

Such investigation would entail controlled experimentation whereby a number of Subjects (Ss) would be chosen at random from a target population of Japanese exchange students. AE initial /l/ and /r/ would be presented in various contexts to the Japanese Ss for phonemic identification. The across-S differences in identifiability for the various contexts would be compared with acoustic characteristics and frequency of occurrence of the stimuli. Such comparison of across-S differences with linguistic facts of the stimulus leads to a more basic understanding of the events surrounding the /l/ and /r/ identification problem, and can therefore ultimately lead to more successful instructional designs.

The next three sections contain a short articulatory description of AE initial /l/ and /r/ and Japanese /r/ and a somewhat more extensive acoustic description of the variants of AE initial /l/ and /r/ in various contexts according to the Lehiste (1964) data. These are followed by a short section summarizing the implications of the acoustic and articulatory evidence for the identification of AE initial /l/ and /r/ by speakers of Japanese.

Articulatory Description of AE /l/ and /r/

AE initial allophone of /l/ is a voiced apico-alveolar lateral vocoid. The initial variety begins with the tongue position approximately that for AE [t], [d], [n], but with the sides of the tongue lowered so that air passes over and out the sides without friction. The end of the initial /l/ is characterized by opening of a center channel for air passage and transition to the position of the following vowel.

AE initial allophone of /r/ is a voiced apico-alveolar retroflex semivowel. According to Francis (1958), the sides of the tongue are against the back teeth in contrast with the initial /l/. "The blade and tip are turned upward and withdrawn a bit toward the back of the mouth, the tip points to the extreme back of the alveolar ridge where it joins the palate, considerably back of the position of contact for the alveolar consonants [t], [d], and [n]. From this position the apex flicks rapidly forward and down into the position for the following vowel."

Prator (1951) notes as well that the lips are open and comments more fully on the manner of articulation. "In whatever direction the movement may end, it always begins by a motion toward the back of the mouth. More than any other factor, it is this retroflex (toward the back) motion that gives the English [r] its typical sound. The tongue tip rises a little and is curved backward, while the sides of the tongue slide along the back part of the tooth ridge as along two rails." Similarly Jones, quoted in Toyoda (1957), describes the AE /r/ articulation as, " a general retraction of the whole body of the tongue with simultaneous lateral contraction."

Articulatory Description of Japanese /r/

Roughly corresponding to AE /l/ and /r/ is the Japanese /r/ described by Bloch (1950) as a short alveolar flap. Bloch reported the Japanese /r/ to consist of two allophones in most speakers, three in some. The [r], a short voiced alveolar flap occurs before [e, a, o, u]. The [r], a short alveolar palatalized flap, occurs before [i,y]. For some speakers [l] stands in free variation with [r] before [e,o].

The allophones of both AE /l/ and /r/, while similar with respect to the point of articulation for the allophones of Japanese /r/ are quite different in manner of articulation. The initial AE and Japanese [l] are both alveolar, lateral, and voiced. The AE [l] is a vocoid, the Japanese [l] a flap. The initial AE and Japanese [r] are both alveolar and voiced, but the AE [r] is also retroflexed and a semivowel, the Japanese [r], a flap. (Kimizuka (1962) considers the point of articulation of the Japanese /r/ and AE /r/ to be different. Kimizuka contends the Japanese /r/ "is produced by the movement of the tip of the tongue touching the palate and releasing it." Elsewhere he makes the general statement that the sound of Japanese /r/ is "somewhere between the American /r/, /1/ and /d/.")

Acoustic Description of AE Initial /l/ and /r/ According to the Lehiste (1964) Data

Figures 1, 2, and 3 contain the mean formant frequencies of AE initial /1/ and /r/ when followed by 11 different vowel nuclei, as well as the steady state frequencies of the vowel nuclei themselves. These various mean formant frequencies were computed from the spectograms of the pronunciation of five native speakers of AE by Lehiste.

Figures 1 and 2 contain the mean frequencies of F_1 and F_2. Figure 1 contains /1/ in the various contexts; Figure 2 contains /r/. The direction of increasing frequency has been adjusted so that the vowels assume positions approximately those in an articulatory or perceptual triangle chart (Hanson, 1960). Thus, fronting is toward the left, backing toward the right. High is toward the top of the sheet, low toward the bottom.

The straight line drawn from the point of F_1 and F_2 for the initial consonant, $C(F_1,F_2)$, to the point of F_1 and F_2 for the vowel nucleus, $V(F_1,F_2)$, does not mean the transition actually followed the shortest path between those two points. The straight line serves to visually locate the two points of a phonemic sequence on the chart.

The mean of all the F_1 and F_2 for /1/, $\overline{/1/(F_1,F_2)}$, is

FIGURE 1

Initial and Final F_1 and F_2 Frequencies for 11 [l + vowel] Sequences, Lehiste (1964) data

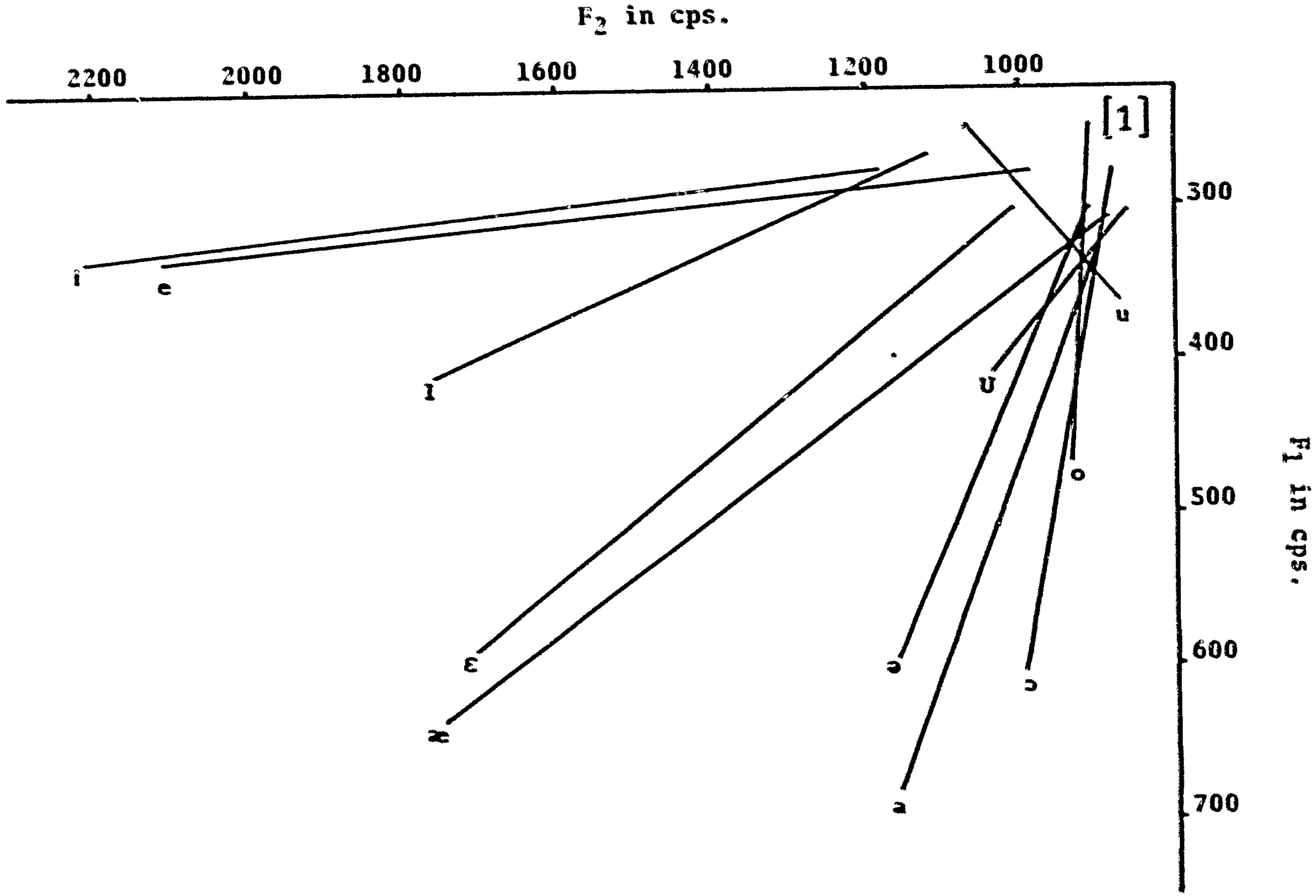


FIGURE 2

Initial and Final F_1 and F_2 Frequencies for 11 [r + vowel] Sequences, Lehiste (1964) Data

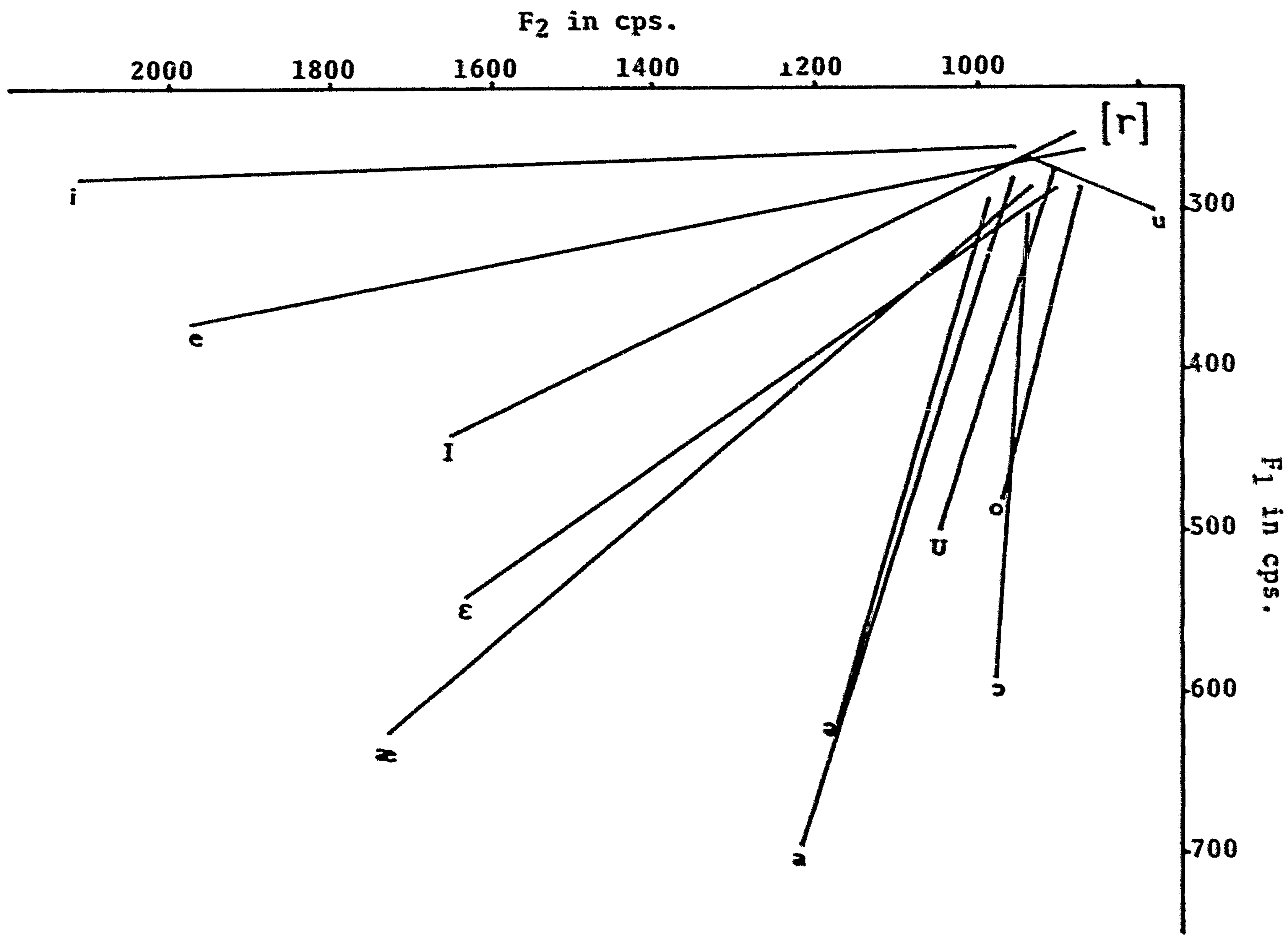
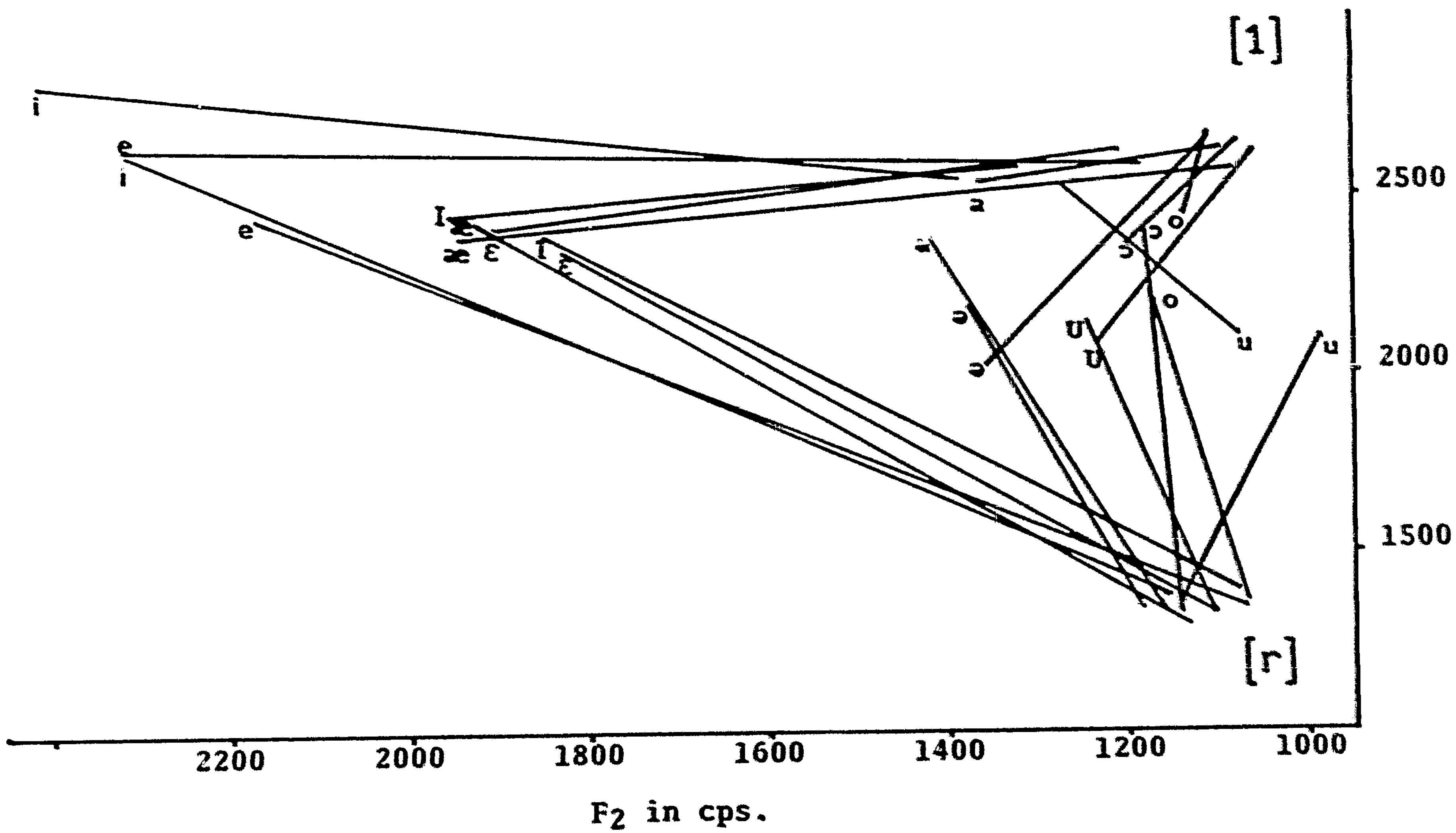


FIGURE 3

Initial and Final F_2 and F_3 for Sequences of /l/ and /r/ + 11 Different Vowels, Lehiste (1964) Data



not very different from $\overline{/r/(F_1,F_2)}$. But $/1/(F_1,F_2)$ varies much more with respect to the following vowel context, than does $/r/(F_1,F_2)$. The values for $/1/(F_1,F_2)$ are contained within an area bounded by $250<F_1<330$cps. and $850<F_2<1200$cps., while the values for $/r/(F_1,F_2)$ are contained in an area bounded by $255<F_1<305$cps. and $870<F_2<990$cps. There are some differences in the $V(F_1,F_2)$ preceded by /1/ from those preceded by /r/, but no general trend could be easily defined.

Figure 3 contains the mean frequencies for F_2 and F_3 for both AE /1/ and /r/ preceding 11 vowel nuclei. As an aid to reading the graph, F_2 was plotted along the horizontal axis increasing to the left as it wa^r for Figures 1 and 2. F_3 is plotted along the vertical axis increasing towards the top of the graph. The straight lines drawn from $C(F_2,F_3)$ to $V(F_2,F_3)$ serve to visually locate the two points of a phonemic sequence on the graph.

There was little variation for $/1/(F_3)$ and $/r/(F_3)$ with respect to the following vowel context. In most instances the $V(F_3)$ preceded by /1/ is slightly lower than $V(F_3)$ preceded by /r/. The effect is minimal however. The variation of $/1/(F_2,F_3)$ seen in Figure 3 is due to the variation of $/1/(F_2)$ pointed out in Figure 1.

The only striking difference between the formant frequencies of /1/ and /r/ occur in the third formant as has been indicated by Joos (1948); O'Connor, Gerstman, Liberman, Delattre, & Cooper (1957); Lisker (1957), and others as well as Lehiste herself.

Summary of Acoustic and Articulatory Evidence on Identification

Native speakers of Japanese would not be expected to phonemically distinguish between lateral and non-lateral sounds since according to Bloch (1950),

1) the lateral and non-lateral alveolar flaps are in free variation for some speakers; and

2) the lateral flap does not even exist for most speakers of Japanese. Since all the allophones of Japanese /r/ are flaps, the Japanese would not be expected to be sensitive to duration and abruptness cues in connection with AE /l/ and /r/. Of course, the Japanese Ss already speak English and would therefore be expected to identify AE /l/ and /r/ correctly to some extent.

The identification of AE /l/ and /r/ was expected to be affected by allophonic variation. Each major allophonic variation constitutes a different stimulus which must be perceived and relegated to the appropriate phoneme. This was borne out in pilot checks where identification scores for AE /l/ and /r/ in cluster were considerably lower than the scores for the initial allophones. AE /l/ and /r/ in cluster were not included in this study in order to maintain the complexity of the experimental design within reasonable limits until the effects on the non-allophonic-varying context had been examined.

Phonological context would be expected to affect identification through other than strict allophonic variation. Results in French and Steinberg (1947), Luescher and Zwislocki (1949), Moser and Dreher (1955), Hardy (1956), and Calearo and Lazzaroni (1957), along with others would tend to indicate that initial /l/ and /r/ might not be perceived only in terms of inherent (self-contained) characteristics, but also in comparison with (in contrast with) contiguously occurring phonemes. Results of this study would therefore not only point out troublesome phonemic sequences, but would have bearing on which context dimensions had an effect on identification and which features of the stimuli the Ss seemed to be perceiving.

Experimental Design and Equipment

There were three stimulus variables, the first of which was I, the syllable initial consonant variable with two levels: /l, r/. The second, V, was the vowel nucleus with 14 levels: /i, I, e, ε, æ, a, ɔ, o, U, u, ə, aI, aU, ɔI/. The third, C, was the final consonant variable with 22 levels: /p, b, t, d, k, g, f, v, θ, ð, s, z, š, ž, č, ǰ, m, n, ŋ, l, r, ∅/. (∅ represents "without final consonant.")

The 616 stimulus syllables consisted of all possible combination of levels for the three variables. In addition to the original list of stimulus syllables in phonemic transcription, a second list of the syllables in conventional English orthography was prepared for use on the response sheet.

A randomized list of the 616 stimulus syllables was prepared and then read onto tape by the experimenter. The syllables were pronounced at 3-second intervals. After the pronunciation of every fifth syllable on the list, the number of the next stimulus syllable was given.

For each oral stimulus recorded on tape, two syllables appeared on the response sheet: the conventional English spelling of the stimulus syllable and its /l/ or /r/ counterpart. The left-right position of the /l/ syllable was assigned randomly, the /r/ syllable taking the opposite position.

For example:

1.	lace	race	
2.	rid	lid	etc.

The six Japanese Ss, who were students at Indiana University, were tested individually. Each S was placed at a table equidistant from the two high quality loudspeakers. The distance from the speakers and the sound level were the same for all six sessions.

The instructions were presented to the S in both written (cf., Appendix A) and oral form. The instruction sheet was separate from the response sheets and contained six examples. The oral form was played back from tape through the same system and under the same conditions as for the presentation of the stimulus syllables. None of the six Ss requested further explanation after presentation of the instructions.

All recordings were made at the language laboratory recording studios in Ballantine Hall, Indiana University on Ampex recorders. The tapes were recorded at 7.5 inches/second (19cm./sec.) half-track. Playback for testing was accomplished on a Telefunken 96k tape recorder connected to a Fisher Model 800 radio-amplifier system. The transmission was judged clear after testing two native speakers of English. Out of 616 responses, there were four errors for one S, and six for the other. The two native Ss had no error in common, and therefore the loss in the system was considered to be negligible.

Experimental Results

The summary of the analysis of variance can be found in Table 1. The means for the various levels of each of the four variables can be found in Table 2.

This experiment was considered to be an expansion of the A x B x S design in Lindquist (1953) to an A x B x C x S design. As in Lindquist, the error term used for each F test was the interaction of the variable (or variables) to be tested with the Ss.

The analysis of variance showed the C, the ending of

TABLE 1

Summary of Analysis of Variance

Grand Mean = 0.73404

Source of Variation	Degrees of Freedom	Mean Squares	F	P
I	1	6.84010	1.52161	N.S.
V	13	0.15337	.87590	N.S.
C	21	0.21904	1.75922	<.05
S	5	14.84118		
IV	13	0.61457	2.57142	<.01
IC	21	0.28738	1.31368	N.S.
IS	5	4.49529		
VC	273	0.17356	1.12054	N.S.
VS	65	0.17510		
CS	105	0.12451		
IVC	273	0.17174	1.02470	N.S.
IVS	65	0.23900		
ICS	105	0.21876		
VCS	1365	0.15489		
Residual	1365	0.16760		
Total	3695			

I = Syllable Initial Consonant
V = Vowel Nucleus
C = Final Consonant
S = Subject

TABLE 2

Marginal Means

Variables	Levels	Means	Variables	Levels	Means
I	l-	0.69102		-f	0.72619
	r-	0.77706		-v	0.75595
V	-i-	0.67803		-θ	0.66071
	-I-	0.72727		-ð	0.66071
	-e-	0.73485		-s	0.75595
	-ɛ-	0.76515		-z	0.74405
	-æ-	0.72727		-š	0.66667
	-a-	0.75000		-ž	0.70238
	-ɔ-	0.74621		-č	0.76190
	-o-	0.70455		-ǰ	0.74405
	-U-	0.73485		-m	0.73810
	-u-	0.74242		-n	0.76190
	-ə-	0.76515		-ŋ	0.73810
	-aI-	0.75758		-l	0.78571
	-aU-	0.73106		-r	0.75000
	-oI-	0.71212		Ø	0.77976
C	-p	0.71429	S	1	0.80357
	-b	0.70238		2	0.62987
	-t	0.72619		3	0.91558
	-d	0.75000		4	0.54058
	-k	0.76190		5	0.62825
	-g	0.76190		6	0.88636

I = Syllable Initial Consonant
V = Vowel Nucleus
C = Final Consonant
S = Subject

the stimulus syllable, to have a statistically significant main effect on the perception of the syllable-initial /l/ and /r/. ($F = 1.76$; 21, 105 d.f.; $P < .05$).

Figure 4 shows the means for the 22 levels of variable C in the order of decreasing proportion correct for the criterion task.

The Duncan procedure as described in Winer (1962), was applied to determine which of the means of the final consonants were (statistically) significantly different from the others. Table 3 contains a summary of analysis and the group score difference matrix used to determine whether the scores for one group are significantly different from the others. The results are shown in Table 4.

Schematically the results in Table 4 might be summarized as follows:

Group Number	1	2	3	4	5	6	7	8	9	10	11	12
	θ	š	b	p	t	m	z	d	s	g	Ø	l
	ð		ž		f	ŋ	ǰ		r	n		
									v	č		
										k		

Those levels included in groups underlined by a common line do not differ (significantly); levels included in groups not underlined by a common line do differ. Thus, levels for /l/ and Ø differ from levels for /θ, ð, š/ but not /b, ž, p, t, f, m, ŋ, z, ǰ, d, s, r, v, g, n, č, k/, and similarly /θ, ð, š/ differ from /l/ and Ø but not from /b, ž, p, t, f, m, ŋ, z, ǰ, d, s, r, v, g, n, č, k/. It

FIGURE 4

Levels of C Grouped According to Total Number Correct in Preparation for Use of Duncan Procedure

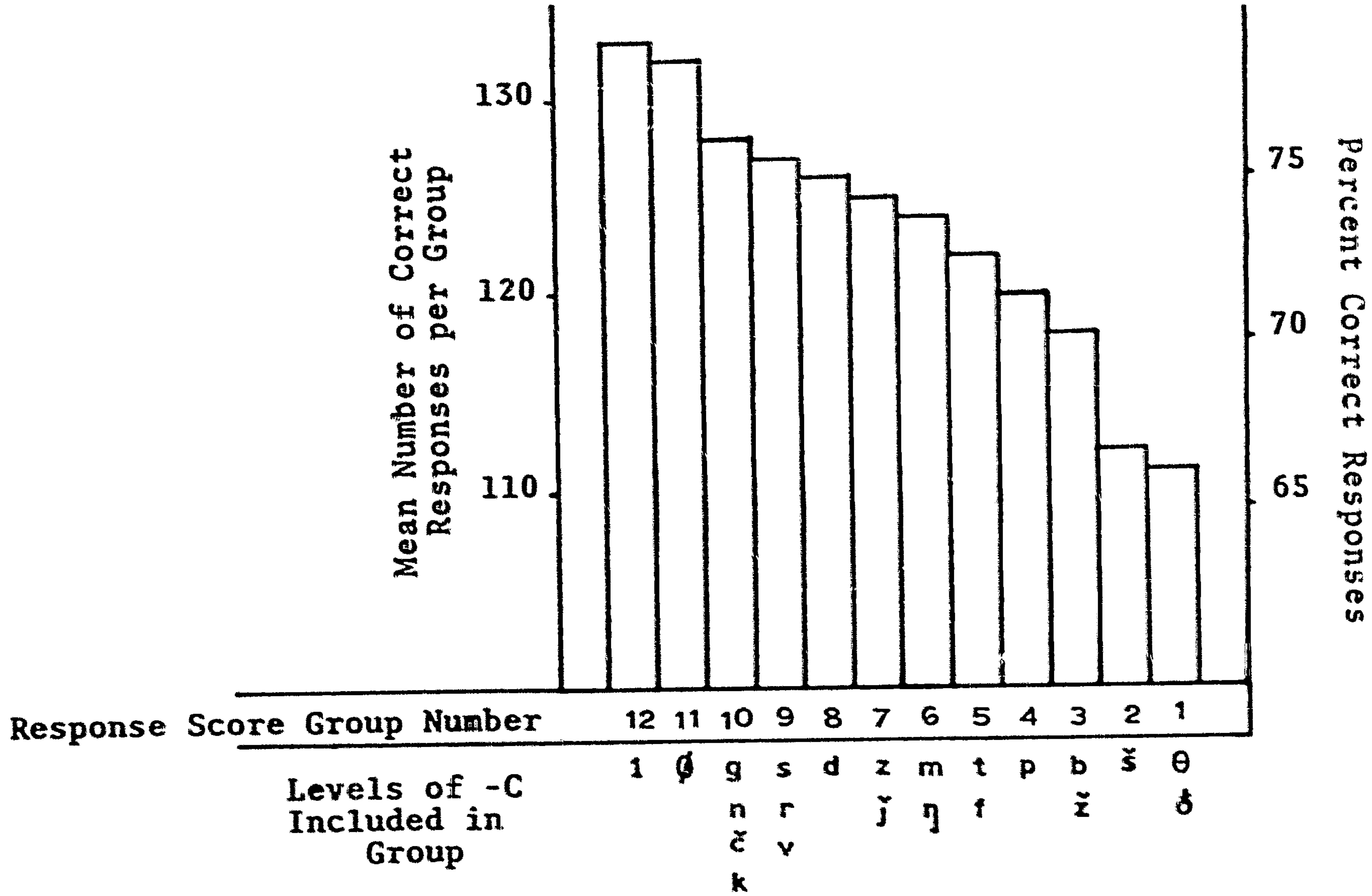


TABLE 3

Tests on All Ordered Pairs of Means, Duncan Procedure, (k = 22, n = 168)

Source		SS	df	MS	F
Treatments	C	4.59984	21	0.21904	1.75922
Experimental Error	CS	13.07386	105	0.12451	
Total		17.67370	126		$F_{.95}$ (21, 105)* = 1.66

Order	1	2	3	4	5	6	7	8	9	10	11	12
Treatment in order of T_j	θ ð	š	ž b	p	t f	m ŋ	ǰ z	d	s r v	g n č k	ø	l
T_j	111	112	118	120	122	124	125	126	127	128	132	133

	1	2	3	4	5	6	7	8	9	10	11	12
1	-	1	7	9	11	13	14	15	16	17	21	22
2		-	6	8	10	12	13	14	15	16	20	21
3			-	2	4	6	7	8	9	10	14	15
4				-	2	4	5	6	7	8	12	13
5					-	2	3	4	5	6	10	11
6						-	1	2	3	4	8	9
7							-	1	2	3	7	8
8								-	1	2	6	7
9									-	1	5	6
10										-	4	5
11											-	1
12												-

*interpolated

TABLE 4

(k = 22, n = 168, $\sqrt{n\ \text{MSerror}}$ = 4.573)

Truncated Range r	2	3	4	5	6	7	8	9	10	11	12
q.99 (r, 105)*	3.717	3.874	3.982	4.060	4.124	4.175	4.219	4.256	4.289	4.319	4.344
q.99 (r, 105) $\sqrt{n\ \text{MSerror}}$					18.86	19.09	19.29	19.46	19.61	19.75	19.87

	θ ð	š	ž b	p	t f	m ŋ	ǰ z	d	s r v	g n č k	ɸ	l
θ ð											**	**
š											**	**
ž b												
p												
t f												
m ŋ												
ǰ z												
d												
s r v												
g n č k												
ɸ												
l												

*interpolated

is predicted that the probability of such differences occurring as a result of random error would be less than 1%.

The null consonant, Ø, indicating the absence of final consonant, /C/, corresponded to relatively high perception scores as predicted. Although not necessarily greater in duration, the stimulus syllables of three phonemes are 1/2 longer in terms of phoneme length than those of two phonemes. Additional stimulus length is associated with increased response difficulty unless the additional length provides information needed for solution to criterion.

No helpful information would be provided by the addition of a /-C/ in most instances. In fact the addition of a /-C/ may distract and interfere with the necessary information in the first two phonemes. A final consonant probably has a slight effect on the preceding vowel, which in turn may have a slight effect on the transitions from the points of major constriction for /l/ and /r/ to the vowel nucleus. As the transitions are of great importance for the perception of /l/ and /r/, any additional change in them would introduce an additional souce of variability with the effect of static or interference in the system.

As the Ss could regularly discriminate between syllable final /l/ and /r/ in pilot tests and were given the final consonant context on the score sheet, it is conceivable that Ss could benefit from the presence of a final /l/ or /r/ by means of comparison with the initial

/l/ or /r/. The stimulus syllable would be viewed as a set of three stimuli and the task would be to establish whether the initial stimulus was the same as or different from the final one. Apparently the Ss could identify the final [l] as the phoneme /l/ better than [r] as /r/, since there were higher identification scores for syllables ending in /l/.

The Japanese Ss nearly always indicated a final [r] in pilot pronunciation tests of the stimulus syllables by [a]. (This observation is made in Kimuzuka (1962) as well.) We might presume the pronunciation of /-or/ as [-oa] for example, to be indicative of perception difficulties. If the Ss thought they had heard Ø instead of final /-r/, they would employ the same hypotheses for perception of /IV/ + /r/ as for /IV/ + Ø resulting in a decrease in performance scores. The mean proportion correct for stimulus syllables /IV/ + /-r/ was .762, somewhat below that of .786 for /IV/ + Ø and .792 for /IV/ + /-l/. None of these differences were statistically significant however.

The low scores for /θ, ð, š/ could not be explained on the basis of occurrence in Japanese or phonological characteristics. The /θ/ and /ð/ corresponding to a score of .661 do not occur in Japanese, but neither does /v/ which corresponds to a score of .756. On the other hand, /š/ is a moderately frequent consonant phoneme in Japanese and corresponds to a score of only .667. No simple grouping according to articulatory position or manner as well as distinctive features seemed to work well, because a group including /θ, ð, š/ would also include /s, z/ corresponding to scores of .756 and .744 in the mid-high portion of the scale. (/θ, ð, š/ differ from /s, r, v, k, č, n, g, ∅, 1/ at the .05 level according to the Duncan procedure.)

In the absence of a phonological grouping that would correspond to the grouping of response scores, it was felt the perception of "real" as opposed to "non-real" stimulus would have to be examined. A "real" stimulus syllable would presumably be one that was a real word in one English dialect or another, or at least occurred in a real word. The point of comparing responses to "real" with "non-real" is derived from the assumption that the Ss will have been exposed more often to the "real" stimulus syllables. Some kind of estimate of the frequency of occurrence in English would provide further information as to exposure.

Therefore, a list of word examples was compiled meeting the following criteria (cf., Appendix B). A standard pronunciation of the word

a) is a stimulus syllable.

b) is a stimulus syllable plus another consonant.

(Clustering of the measure variable /1/ or /r/ in initial position was not permitted, as the /1/ and /r/ in cluster have rather different phonemic shapes which give rise to substantially lower response scores in pilot tests.)

c) is a stressed syllable in a polysyllabic word.

Proper names were excluded from this list. The corresponding Rinsland (1945) total-count defined frequency of occurrence was included as a rough estimate of the Ss' exposure to those words meeting criterion (a). Due to the large number of words that could possibly meet criteria (b) and (c), it was not possible to estimate the frequency of occurrence.

(The words in the Rinsland list were generated by children in Grades I-VIII throughout the United States primarily as written material. Only the Grade I tabulation was supplemented with oral material. The range of total running words was from 350 thousand for Grade I to 1+ million for Grade VII, the range of total different words from 5+ thousand for Grade I to about 18 thousand for Grade VIII. The total-count defined frequency represents the numerical total of the occurrence for each grade even though the number of words collected for each grade was not constant. This weights the sum toward Grade VIII usage where the total number of words was greatest. At best the use of the total-count defined frequency is a rough estimate of the exposure or possible exposure to these English words.)

On the basis of the information in this list the following counts and comparisons were made:

Group A:

The tabulacion for Group A included the responses to stimulus syllables that were the standard dialect pronunciations of words meeting criterion (a) and having a Rinsland frequency greater than 0. Thus, for Group A only the responses to /lak/ pronunciation of "lock" were included in the tabulation, while for other counts the responses to both /lak/ and /lɔk/ were included. The response scores for the "real" stimulus syllables were tabulated separately from the rest of the syllables, termed "non-real," for each S and then compared using the dependent measures t-test.

Table 5 contains the mean proportion correct S-by-S for the "real" versus the "non-real" stimulus syllables as defined above, the computed t-score and P. Even though the t-test yields a relatively small P, i.e., is a relatively non-conservative test, P is still greater than .10.

Group B:

The tabulations for Group B included the responses to stimulus syllables used by many native speakers of English at the University. Thus, both /lak/ and /lɔk/ were considered pronunciations of "lock." The mean proportion correct for each of the five groups of words was computed for each of the six Ss. The t-score was computed for every possible pair.

> Stimulus syllables Group 1 consisted of pronunciations to all words in the list satisfying criterion (a) regardless of Rinsland frequency. These syllables might be regarded as "real" words.
>
> Stimulus syllables Group 2 consisted of pronunciations to words satisfying criteria (b) or (c) but not (a). The syllables in this group occurred as sequences of phonemes in stressed position, but were not monosyllabic words in themselves.
>
> Stimulus syllables Group 3 consisted of all those syllables for which there was no word satisfying criterion (a), (b), or (c). These syllables were neither "real" words nor sequences.

TABLE 5

T-score for Non-Independent Measures Between "Real" and "Non-Real" Stimulus Syllables, Computation Group A

Mean Correct for	S_1	S_2	S_3	S_4	S_5	S_6
Real	.805	.590	.900	.545	.628	.870
Non-real	.803	.710	.949	.535	.644	.917

d.f. = 5	for α =	t_α =
	.5	.727
	.4	.920
	.3	1.156
	.2	1.476
	.1	2.015
	.05	2.571

$t_{computed} = 1.58$ $\quad .1 < P < .2$

Stimulus syllables Group 4 consisted of stimulus syllables Groups 2 and 3. All these syllables did not occur as English monosyllabic words themselves.

Stimulus syllables Group 5 consisted of stimulus syllables Groups 1 and 2. All these syllables occurred as phonemic sequences in the stressed position of English words regardless of the number of syllables in the word.

A matrix containing the t-scores for all possible paired combinations of stimulus syllables Groups 1 to 5 can be found in Table 6. The largest t-score between independent groups with no stimulus syllables in common was 1.873 for stimulus syllables Groups 1 and 2--"real" words and occurring sequences. The smallest t-score was between stimulus syllables Groups 2 and 3--occurring and non-occurring sequences. For all comparisons P was greater than .10, therefore no significant differences were found in the responses to various stimulus syllable groups.

Therefore, the results for both tabulation Groups A and B are for the most part inconclusive. However, it must be pointed out, that there is little to indicate that the occurrence of a stimulus syllable as a phoneme sequence in the stressed position in itself, contributed to the perceptability of the initial /l/ and /r/ if the stimulus syllable was not a "real" word in itself.

The various words in the list grouped by criterion were broken down further into groups on the basis of the final consonant. The numbers of words in the 22 final consonant groups were counted for each of the criterion

TABLE 6

T-scores for Non-independent Measures between Pairs of Stimulus Syllable Groups for 6 Subjects, Computation Group B

Stimulus Syllable Groups	1	2	3	4	5
1	-	1.873	1.249	1.563	1.871
2		-	-0.019	-0.019	-1.874
3			-	0.018	-0.861
4				-	-1.155
5					-

d.f. = 5 for α = .2 t_α = 1.476
= .1 2.015

.1 < P for all pairs

TABLE 7

Pearson-R Correlations Matrix for Score Groups 1 through 6

Groups	1	2	3	4	5	6
1	-	0.262	0.513	0.546	0.123	-0.552
2		-	0.647	0.699	0.007	-0.636
3			-	0.906	0.189	-0.908
4				-	-0.039	-0.888
5					-	-0.425
6						-

groups of words, and then compared with the total number of correct responses and Rinsland frequency for monosyllabic words in each of the 22 final consonant groups.

Table 7 contains the Pearson-R correlations for the following paired observations:

1. Total score correct out of 168 possible for all stimulus syllables.

2. Rinsland total-count defined frequency of words meeting criterion (a).

3. Number of monosyllabic words satisfying criterion (a) and having a Rinsland frequency greater than zero.

4. Number of stimulus syllables corresponding to words in the list that satisfy criterion (a) regardless of the frequency.

5. Number of stimulus syllables corresponding to words in the list satisfying criteria (b) or (c), but not (a).

6. Number of stimulus syllables which correspond to no words in the list, i.e., those not included in 4 or 5.

Although there is no significant difference in the mean proportion correct for various types of "real" versus "non-real" or "non-occurring" stimulus syllables, there is a positive correlation between the number of correct responses and the number of monosyllabic words satisfying criterion (a).

As mentioned above, the numbers of words represented by the stimulus syllables, the frequency of these in English, and the number of stimulus syllables that represent words are imperfect measures of what we might like to call the

S's exposure to the stimulus syllables in his everyday experience with English. The number of stimulus syllables representing words, the numbers of words (which are highly related to the first), and the frequencies of the words are very low for stimulus syllables /IV/ + /θ, ð, š/ for which the lowest response scores were observed.

Figures 5, 6, and 7, contain graphs of the relations between variables 1 and 2, 1 and 3, and 1 and 4.

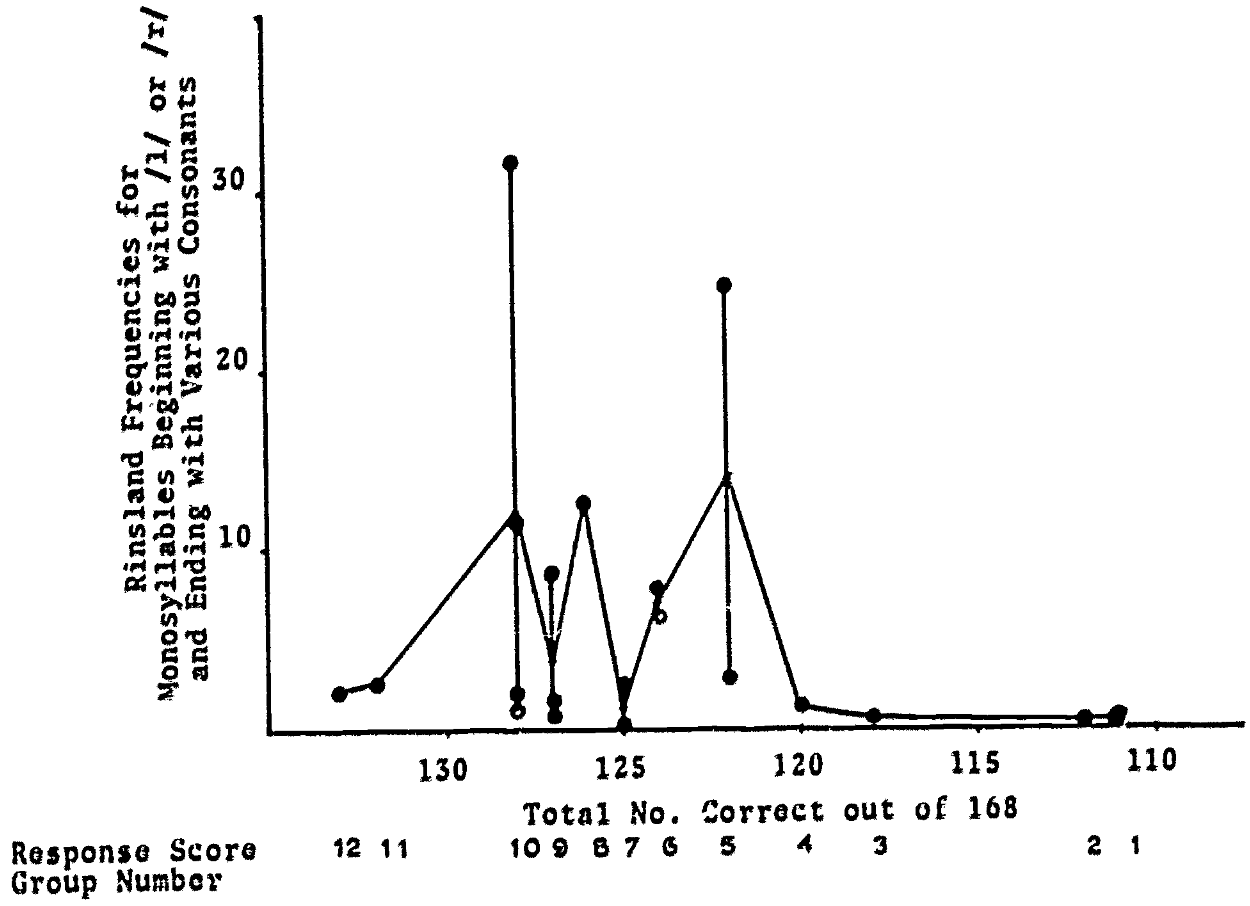
The Pearson-R correlation is actually not appropriate in the sense that any correlational deviation from a straight line reduces the size of R as well as high variability and small range. Threshold jump is a distinct possibility here. It is evident that the relation between 1 and 2 is low especially because of the high error variance. The relation of 1 and 4 is approximately zero for the majority of the final consonant groups while the few groups for which there are few stimulus syllables that are monosyllabic words have the lowest response scores.

Although there was no general effect for "probable exposure" to the individual syllable as was apparent from the low t-scores, there seems to be a possible effect on the response scores for a final consonant group if the probable exposure to stimulus syllables of that group is sufficiently low.

The experience the Japanese Ss have had with a particular stimulus syllable prior to the experiment is in

FIGURE 5

Sum of Rinsland Total-Count Defined Frequencies for the 12 Response Score Groups as a Function of Total Number Correct



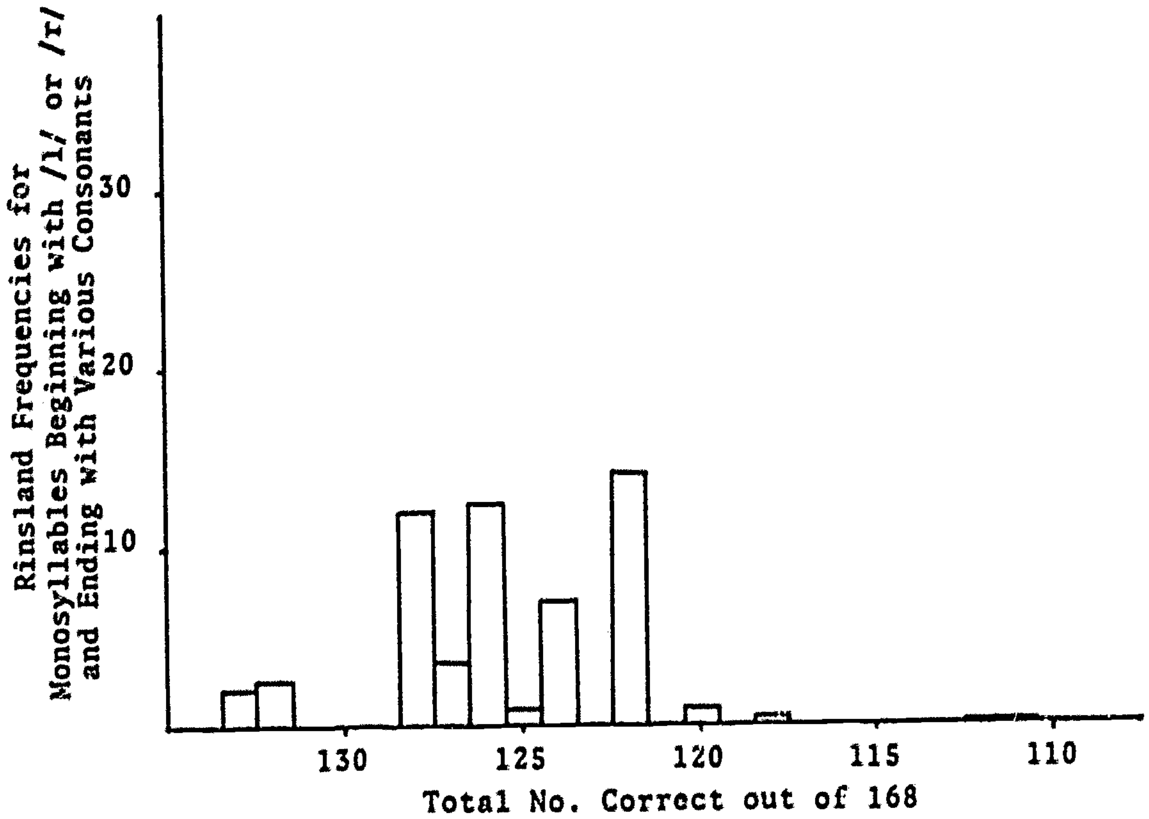
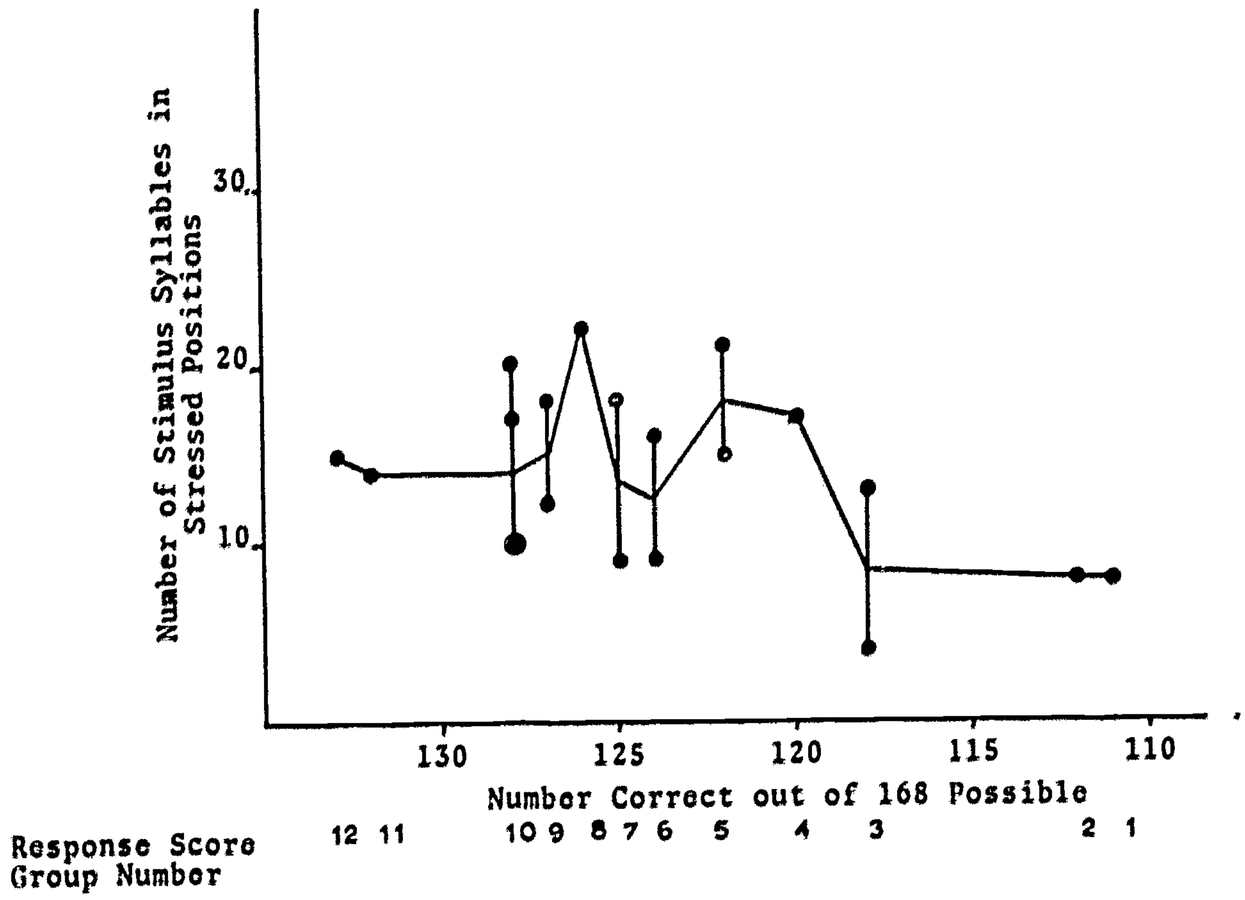


FIGURE 6

Number of Stimulus Syllables Occurring in Stressed Positions of "Rea." Words as a Function of Total Number Correct



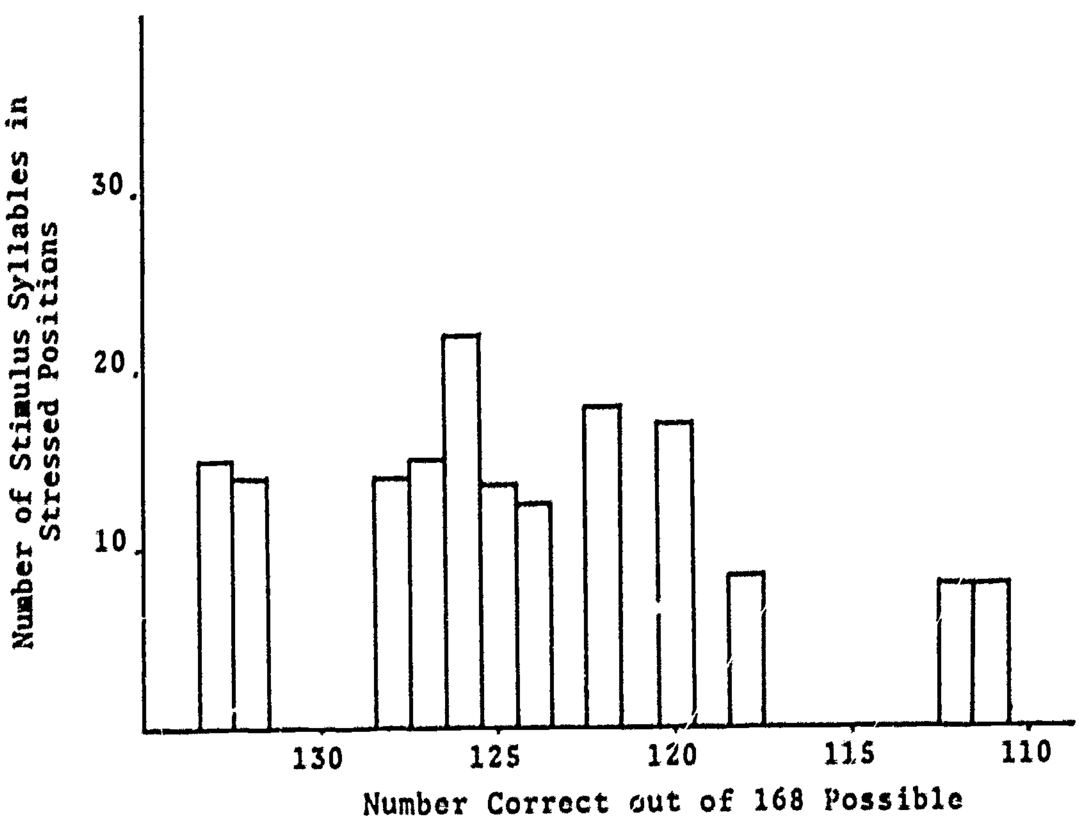
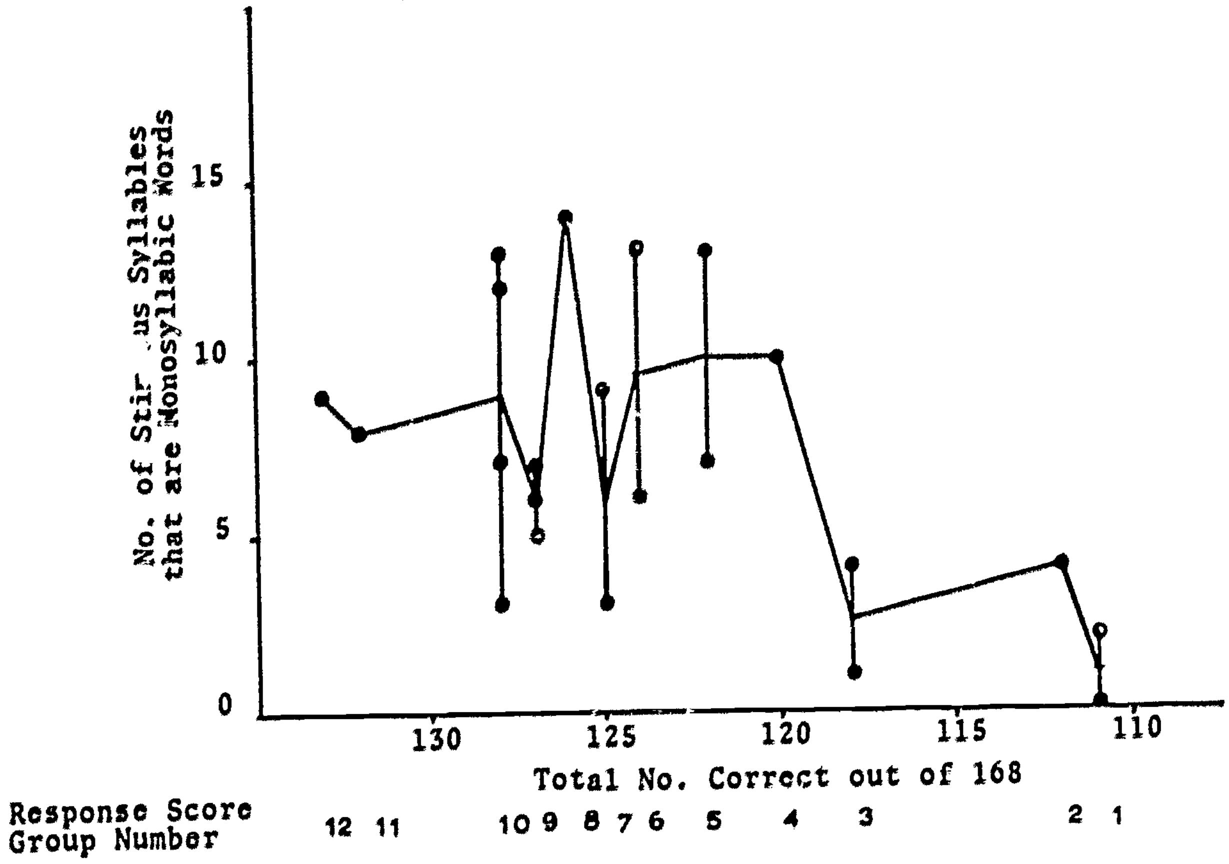
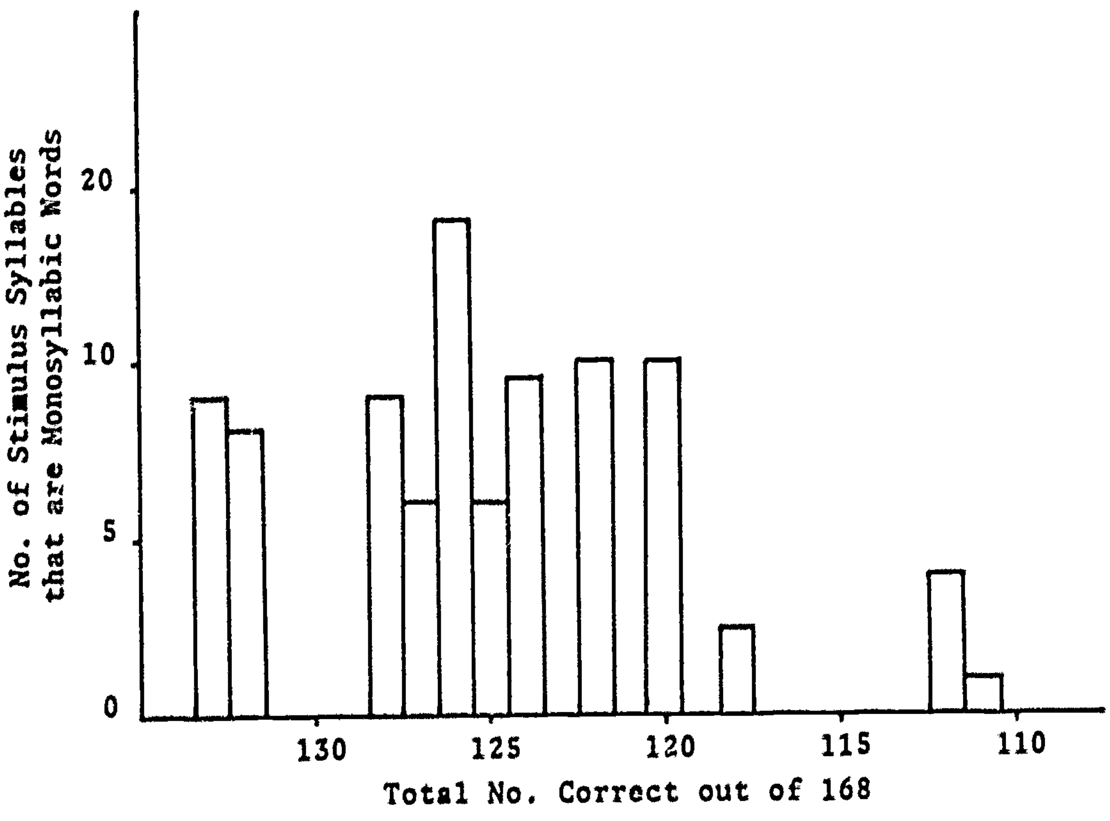


FIGURE 7

Number of Stimulus Syllables that are English Monosyllabic Words as a Function of the Total Number Correct





part a function of his experience with English, if it occurs in English and how often it occurs in English. Therefore, it would be expected that the performance of the Ss on groups of stimulus syllables would be a function of the number of stimulus syllables experienced and the frequency of experience with each.

One plausible conjecture is that lack of exposure to a particular type or group of stimulus syllables prevents the S from making some of the generalizations needed for successful initial consonant identification. If the S has been exposed to several of the stimulus syllables of that group fairly frequently, he arrives at some generalizations aiding perception. Further exposure might make for little improvement in performance once the threshold has been crossed.

Naturally as the overall appropriate experience in English is obtained, more and more of the types are mastered until, as was apparently the case with the native speakers of English, there is no longer any difference in the total correct for various contexts.

As Japanese students apparently scored significantly lower in this task only for types that rarely occur in English, there is little to encourage designers of instructional programs to drill them. The rarer types are encountered less often, are for the most part less important, and are therefore mastered only later in the experience of the second-language learner.

The analysis of variance also showed the A x B interaction, the interaction between the syllable initial consonant, I, and syllable-medial vowel, V, to be highly significant. ($F = 2.57$; 21, 65 d.f.; $P<.01$)

Figure 8 shows in graph form the score correct out of a possible 132 correct for each of the 28 interaction pairs. The vowels of the pairs are placed at even intervals along the horizontal axis in the order of decreasing difference between corresponding /l/ and /r/ interaction pairs. The difference was taken to be the score for the /r/ interaction pair minus the score for the /l/ interaction pair.

$$\text{Difference Score} = \text{Score}_{/r+V_i/} - \text{Score}_{/l+V_i/}$$

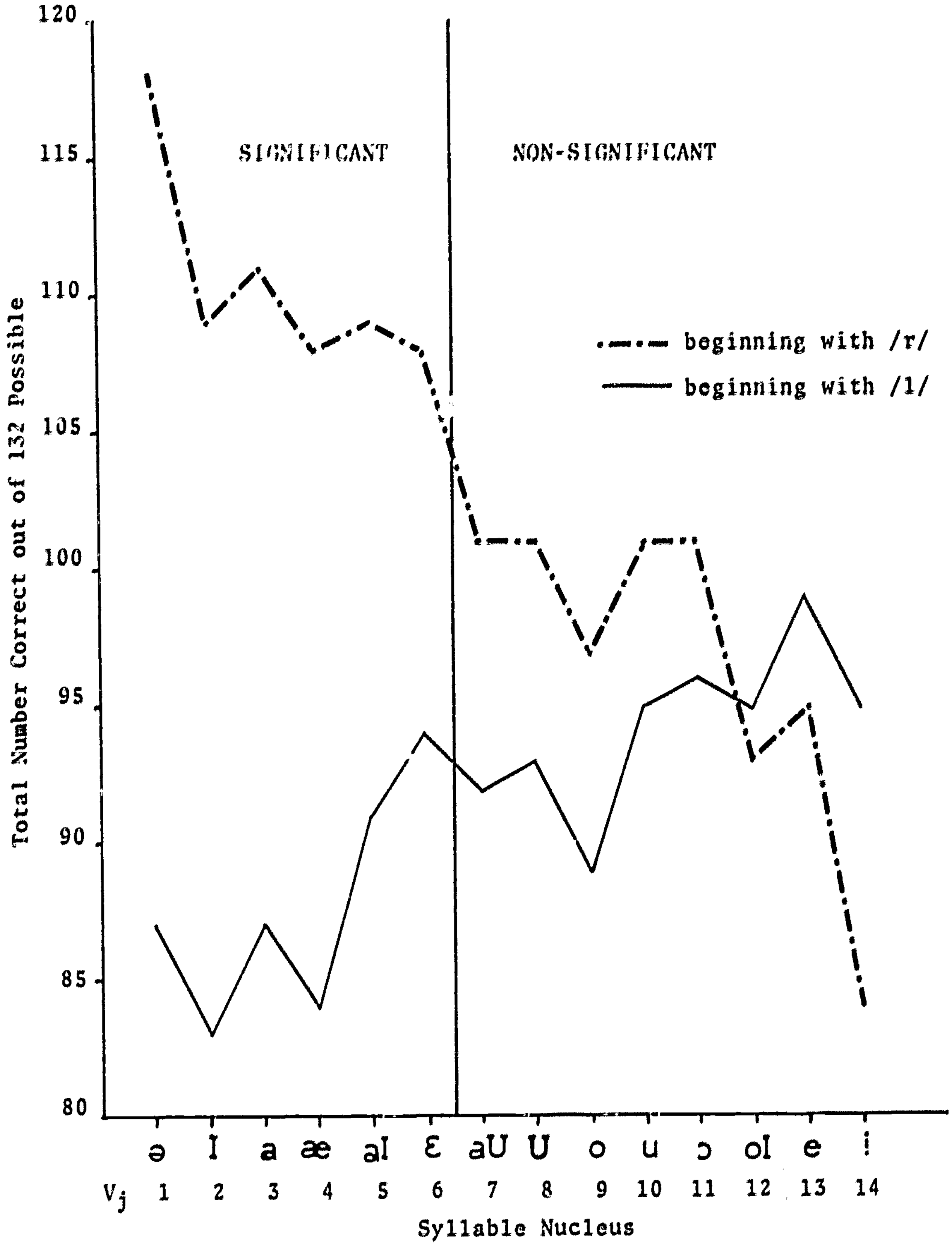
The difference scores for vowel nuclei /oI, e, i/ is therefore negative, the rest are positive. (For convenience of notation, the vowels have been numbered in order of the decreasing difference score (cf., Figure 8).

Scheffé's Test
(Edwards, 1960)

The difference in total correct between $/lV_{1-6}/$, (/lə, lI, la, læ, laI, lɛ/) and $/rV_{1-6}/$, (/rə, rI, ra, ræ, raI, rɛ/), was significant at the .01 level as determined by Scheffé's test for multiple comparisons. First, the difference between $/lV_1/$ and $/rV_1/$ (/lə/ and /rə/) was

FIGURE 8

Score Correct in Total and Percent as a Function of CV Interaction Pair



tested and found to be non-significant. For each subsequent test, the next $/lV_{n+1}/$ was added to the totals for $/lV_{1-n}/$. The highest score, A = 11.85, was obtained for the difference in total score correct between $/lV_{1-6}/$ and $/rV_{1-6}/$. Upon the inclusion of $/lV_7/$ and $/rV_7/$ the magnitude of A decreased, because the additional difference failed to increase the value of the numerator, d^2, sufficiently to offset the increase of two of the a_i term in the denominator.

Articulatory Description

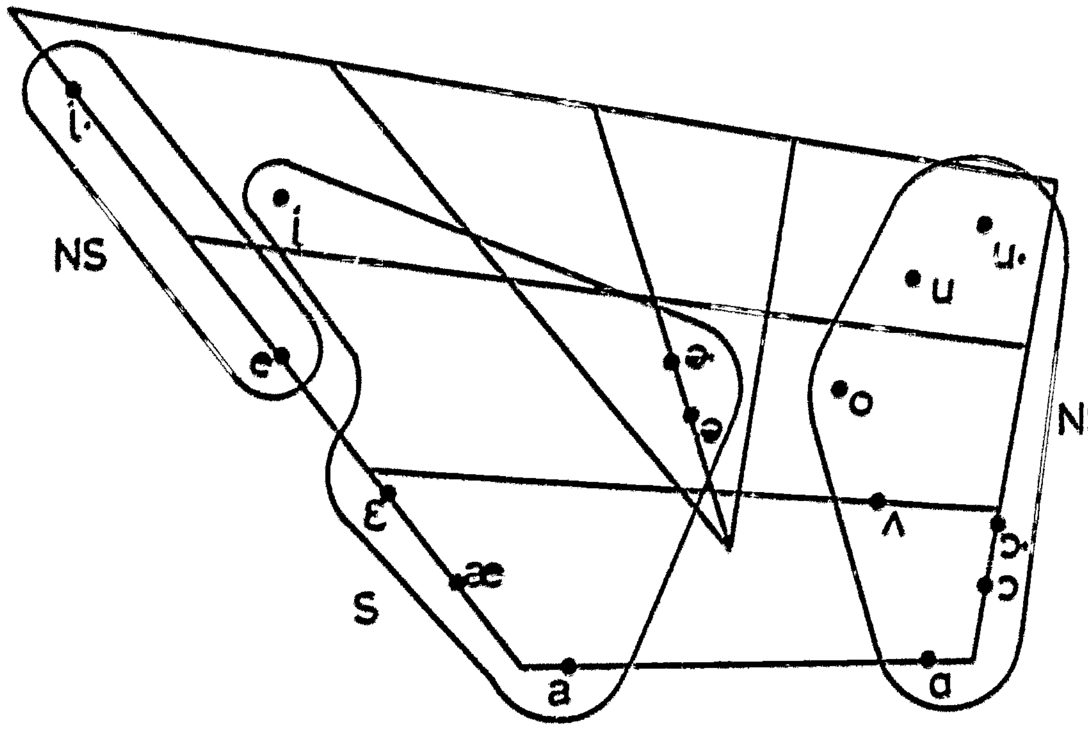
Vowels $/V_{1-6}/$, (/ə, I, a, æ, aI, ɛ/), corresponding to large statistically significant difference scores, are all unrounded central and central-front vowels. The remaining vowels corresponding to smaller non-significant difference scores are divided into two groups, $/V_{7-12}/$, (/aU, U, o, u, ɔ, oI/), which either are or include a rounded back vowel and $/V_{13-14}/$, (/e,i/), which are unrounded high-front long diphthongized vowels (cf., Figure 9). The difference scores for diphthongs /aI, aU, oI/ correspond more closely to that expected for the first phoneme rather than the second. Thus, /oI/ corresponds to a very small difference score, while /aU/ to one somewhat higher and just below the difference score for /ɛ/. The difference score for /aI/ falls well within the range of the significant group.

One S reported hearing an "o-like" sound in connection with /r/. If the S really did hear a type of rounding in connection with /r/ and not with /l/, higher identification scores would be expected for /r/ before unrounded vowels than for /r/ before rounded vowels. The lack of "rounding" for /l/ should therefore lead to lower identification scores for /l/ before unrounded vowels, higher scores for /l/ before rounded vowels. This much is borne out in general by the data. Yet, the absence of the two long high front unrounded vowels from the unrounded group, in fact their occurrence at positions expected for vowels of maximum rounding remain totally unexplained. (Attempts to relate the Japanese [r], palatalized alveolar flap before /i, y/, seem highly contrived.)

The location of $/V_{1-6}/$ and the remaining vowels in articulatory (perceptual) space (Björkhagen, 1956) is shown in Figure 9. Figure 10, likewise shows the location of the various vowels in articulatory (perceptual) space, but includes information as to the magnitude of /l/ versus /r/ syllable difference scores for each of the various vowels. Thus, the innermost enclosed area contains the vowel /ɵ/ associated with the greatest response score difference between /l/ and /r/ syllables (cf., Figure 8), the next highest difference score and so on.

FIGURE 9

Grouping of English Vowels According to lV-rV Response Score Difference

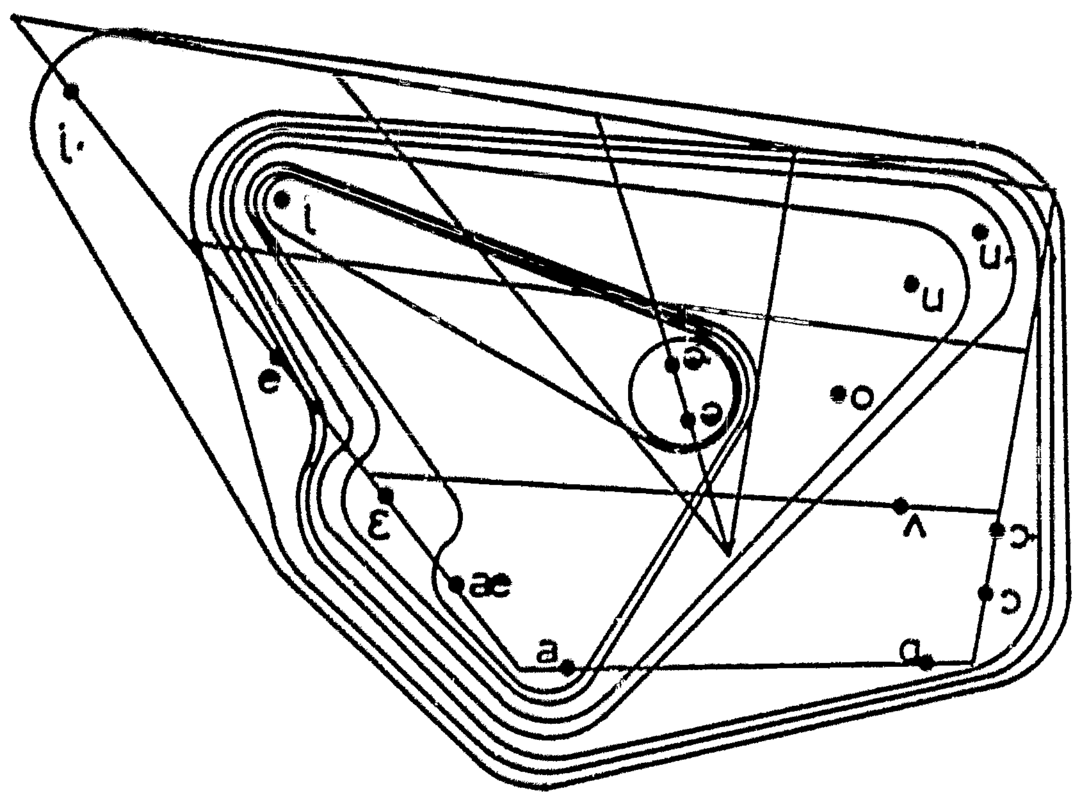


S Syllable nuclei where there was significant difference between RS for syllables beginning with [l] and those beginning with [r]

NS Non-significant

FIGURE 10

Syllable Nuclei Grouped According to [r-] Minus [l-] Response Score Magnitude



Discussion of Some Acoustic Properties of Stimulus Syllables in Relation to Response Scores

Description

F_1 frequencies are generally higher for /V_{1-6}/ than for the remaining vowels. F_2 frequencies for /V_{1-6}/ range from 11 to 18 hundred cps. (cf., Figure 11). F_2 frequencies for /V_{8-11}/ lie below 11 hundred, those for /V_{13-14}/ above. F_3 frequencies for /V_{1-6}/ range from 22 to 26 hundred cps. F_3 frequencies for /V_{8-11}/ tend to lie below those for /V_{1-6}/ while F_3 frequencies for /V_{13-14}/ tend to lie above those for /V_{1-6}/. It must be noted however that the range and differences of F_3 frequencies for the various vowels is realtively small compared with the range and differences of F_2 frequencies for the same vowels.

F_1 frequencies of initial /l/ and /r/ are essentially the same regardless of which vowel nucleus follows (cf., Figure 12). The range and variability of F_2 frequencies are greater for initial /l/ than for initial /r/, but there is no difference between F_2 frequencies of /l/ and /r/ followed by /V_{1-6}/ and those followed by the remaining vowels. The major difference between the formant frequencies of initial /l/ and /r/ lie in F_3. As was the case for F_2 frequency of /l/, the F_3 frequencies of /l/ have a greater range and variability than the F_3 for /r/ in the same contexts. This variability is very small compared to the F_3 frequency differences between initial /l/ and /r/ however.

FIGURE 11

F_1, F_2 and F_3 for Vowels Following AE Initial /l/ and /r/, Lehiste (1964) Data, as a Function of CV Interaction Pair

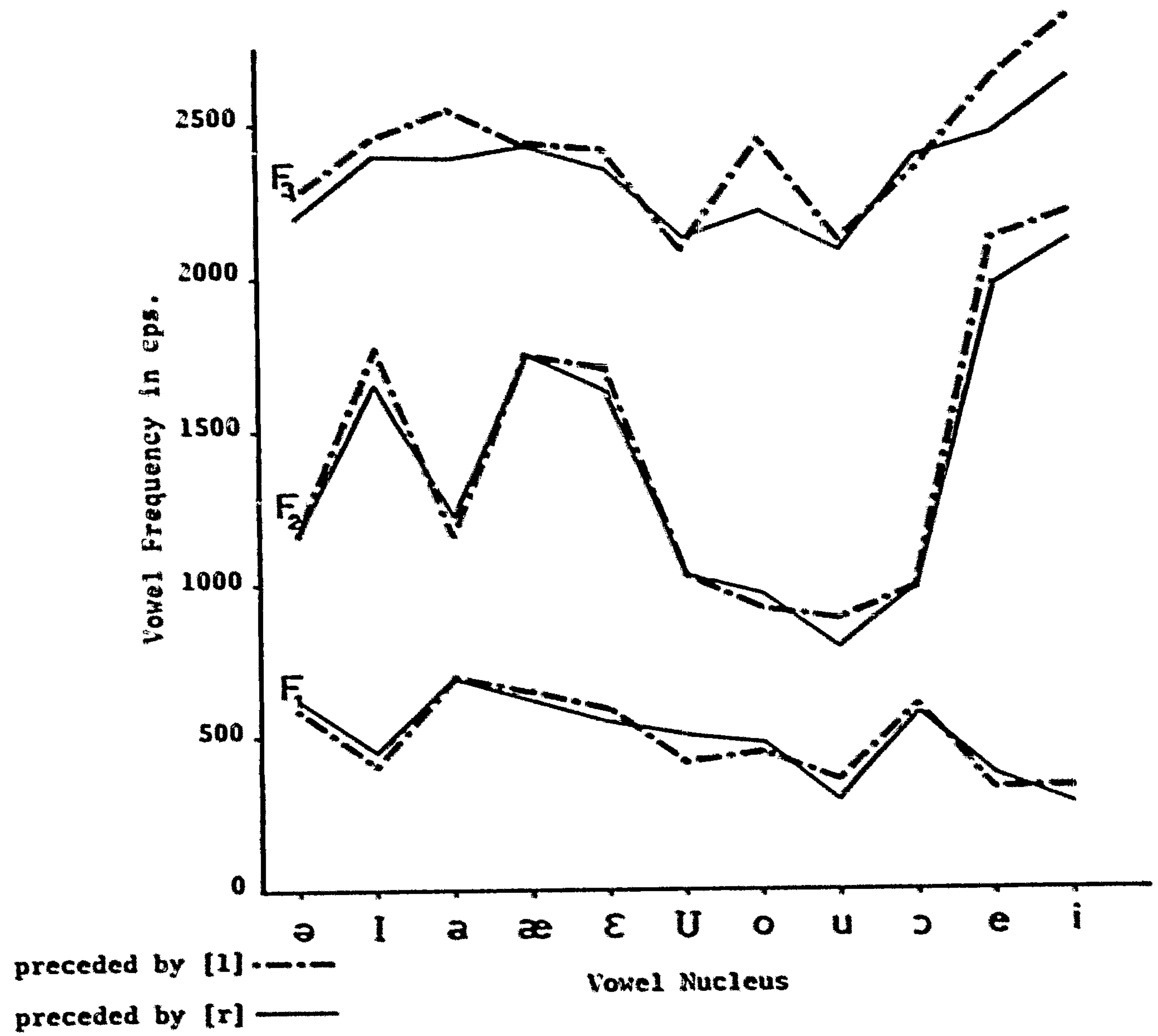
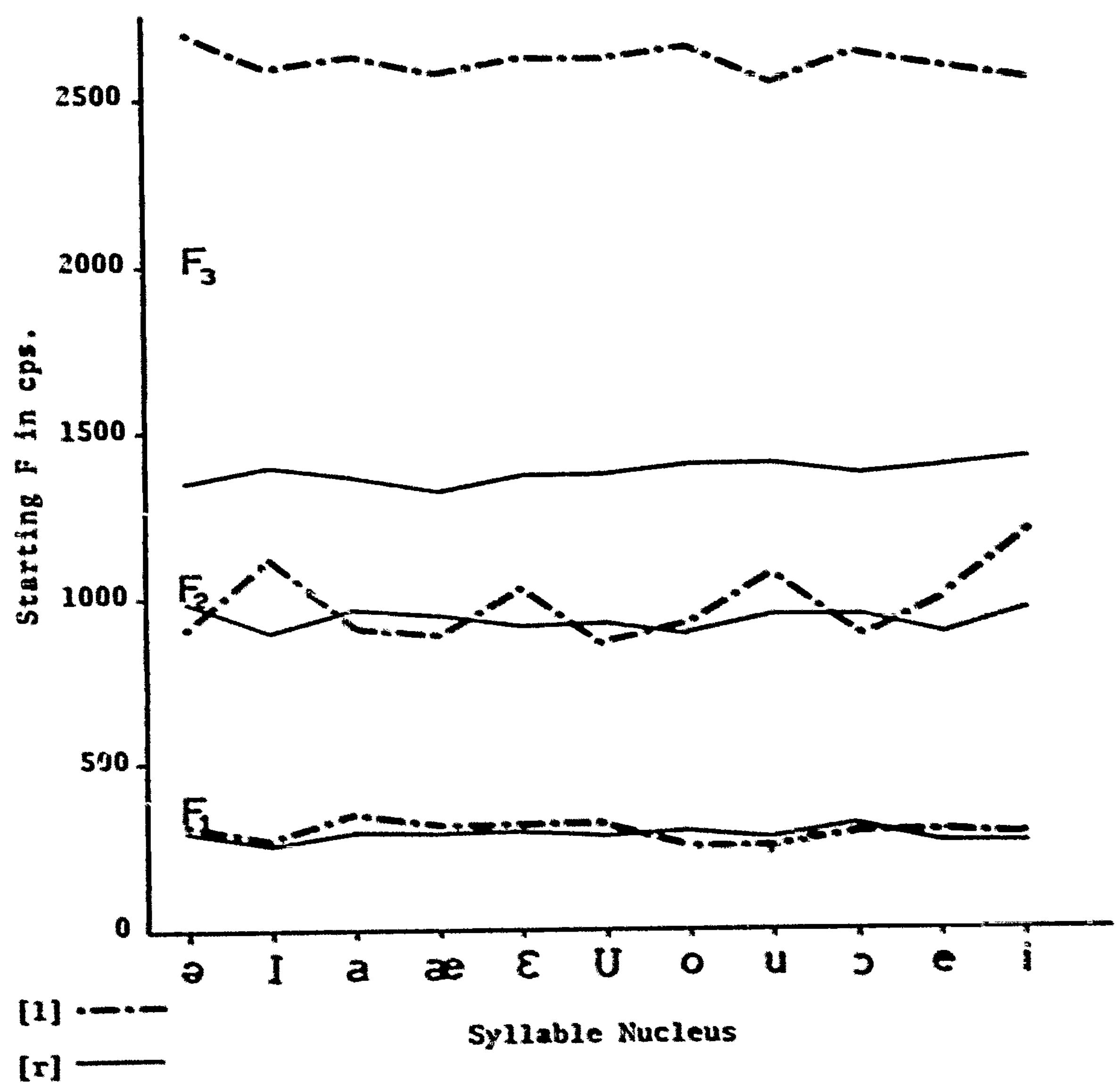


FIGURE 12

Starting F_1, F_2 and F_3 for AE Initial /l/ and /r/ Followed by Various Vowels, Lehiste (1964) data, as a Function of CV Interaction Pair



Any hypothesis of acoustic parameters that corresponds to the response scores (RS) would have to account for high RS for initial /r/ followed by /V_{1-6}/, low RS for initial /l/ followed by /V_{1-6}/ and intermediate scores of less difference between the RS for /l/ + /V_{7-14}/ and /r/ + /V_{7-14}/ (cf., Figure 8).

While the F_1 frequencies are generally higher for /lV_{1-6}/, Figure 11, than for the remaining vowels, and this corresponds to the magnitude of the difference scores, the F_1 frequencies for /V_{1-6}/ preceded by /l/ are the same as F_1 frequencies for /V_{1-6}/ preceded by /r/. Therefore, the F_1 frequencies fail to correspond to the difference in the RS for stimulus syllables beginning with /l/ from those beginning with /r/.

The same argument may be used for the F_2 frequencies of /V_{1-6}/ as compared with those of the remaining vowel nuclei. While a certain range of F_2 frequencies from 11 to 18 hundred cps. contains F_2 frequencies for /V_{1-6}/ and the F_2 frequencies for the remaining vowels lie either above 18 hundred or below 11 hundred cps., there is nothing to account for the CV interaction observed.

F_1 and F_2 frequencies for initial /l/ and /r/ do not differ on the average although the F_2 frequencies for the various [l] varied more than those for [r]. The greater variability of the F_2 frequencies for [l] allophones would account for slightly lower RS for syllables beginning with

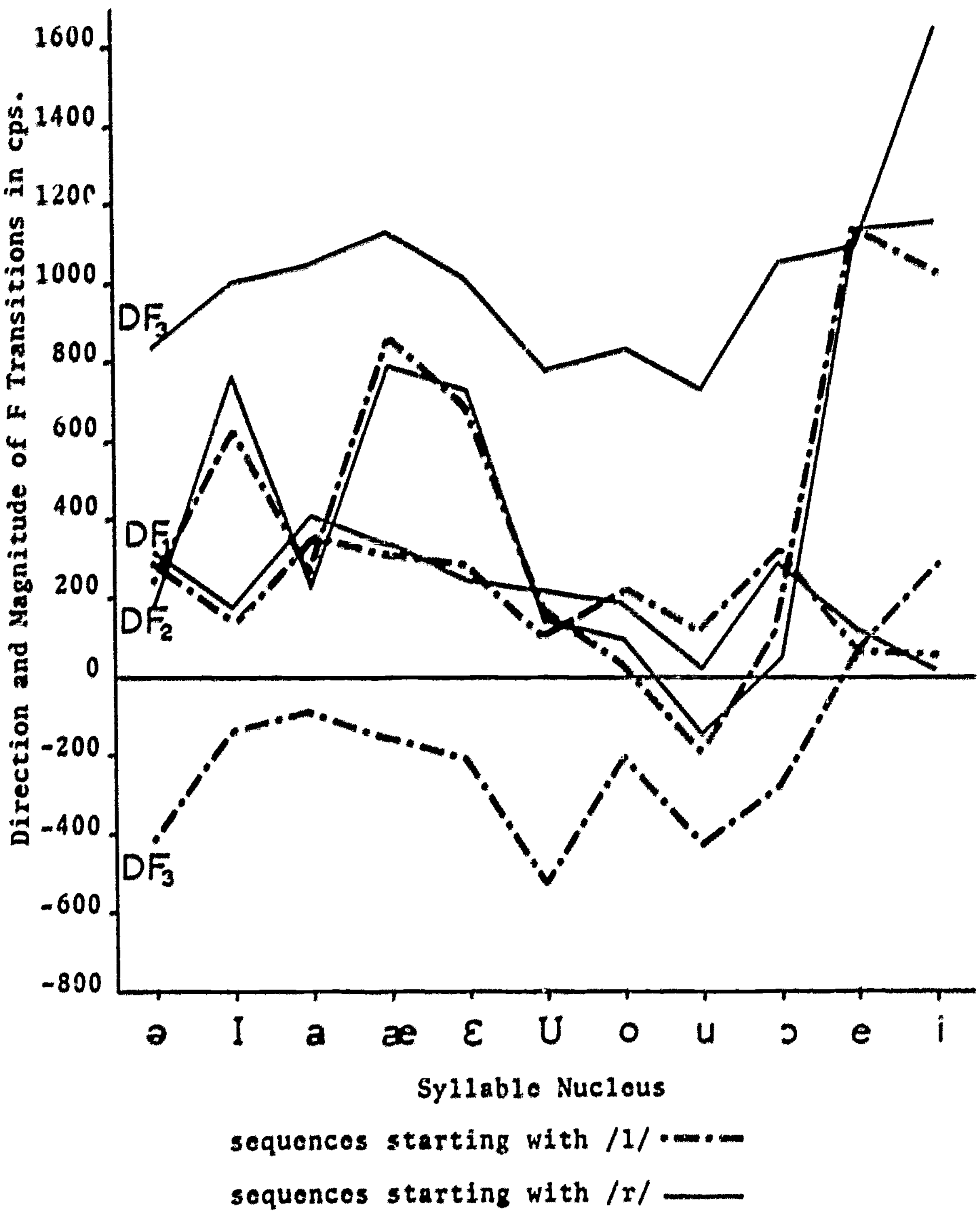
/l/, but would not account for the other variations. While the F_3 frequency for /l/ is considerably higher than the same for /r/, this does not suggest which of the two would have the higher RS; neither does it account for other variability. Any one of the acoustic parameters for C or V alone would only account for a main effect of C or V respectively.

In order to account for the CV interaction where there are significant differences between RS for /l/ and /r/ syllables containing /V_{1-6}/ and nonsignificant differences between /l/ and /r/ syllables containing the other vowels, the acoustic parameter of the various levels of C must be compared with those of the various levels of V. Appropriate dimensions for such a comparison might be the direction and magnitude of the various F frequencies from a particular level of C to a particular level of V. The direction and magnitude of the transitions are shown in Figure 13.

Since the F_1 and F_2 frequencies for the various levels of C and V did not differ systematically in terms of magnitude and direction with regard to the /l/ and /r/ syllables, the F_1 and F_2 frequency transitions do not account for the differences in RS of /l/ versus /r/ syllables. (The broken line representing the /l/ syllables corresponds closely to the solid line representing /r/ syllables for DF_1 and DF_2).

FIGURE 13

Direction and Magnitude of Transition for Initial /l/ and /r/
to Various Following Vowels as a Function
of CV Interaction Pair, Lehiste (1964) Data



The DF_3, the F_3 transitions from C to V, do however. For all V except /e, i/ the F_3 transitions are opposite in direction and differ in magnitude from CV pair to CV pair (cf., Figure 14). For /r/ + /V_{1-6}/ the F_3 transitions are positive and large, for /l/ + /V_{1-6}/ they are negative and small. For /r/ + /V_{7-12}/ the F_3 transitions are positive, but less in magnitude than for /r/ + /V_{1-6}/. For /l/ + /V_{7-12}/ the F_3 transitions are negative and greater in magnitude than those for /l/ + /V_{1-6}/. The F_3 transitions for /r/ + /V_{13-14}/ differ in magnitude, but not in direction. While the difference in magnitude is rather great, it is expected that the absence of the directional cue would reduce the RS scores. As seen in Figure 8 the RS differences lie in the non-significant group.

Figure 15 shows the magnitudes of the various F_3 transitions where the directions are opposite, i.e., including /V_{13-14}/. The order of the vowel nuclei is the same as in Figure 8. For the sake of comparison, Figure 16, a graph of the various RS is presented immediately below. The mean, DF_3, F_3 transitions for /l/ and /r/ syllables of the NS group are more different than the corresponding mean RS for NS /l/ and /r/ syllables. As seen in Figure 17, the exceptionally high score for the CV pair /rə/ would not have been expected on the basis of the magnitude of the F_3 transition. The students may have been given /r/

FIGURE 14

Direction and Magnitude of F_3 Transition for Initial /l/ and /r/ to 11 Following Vowels, Lehiste (1964) Data, as a Function of CV Interaction Pair

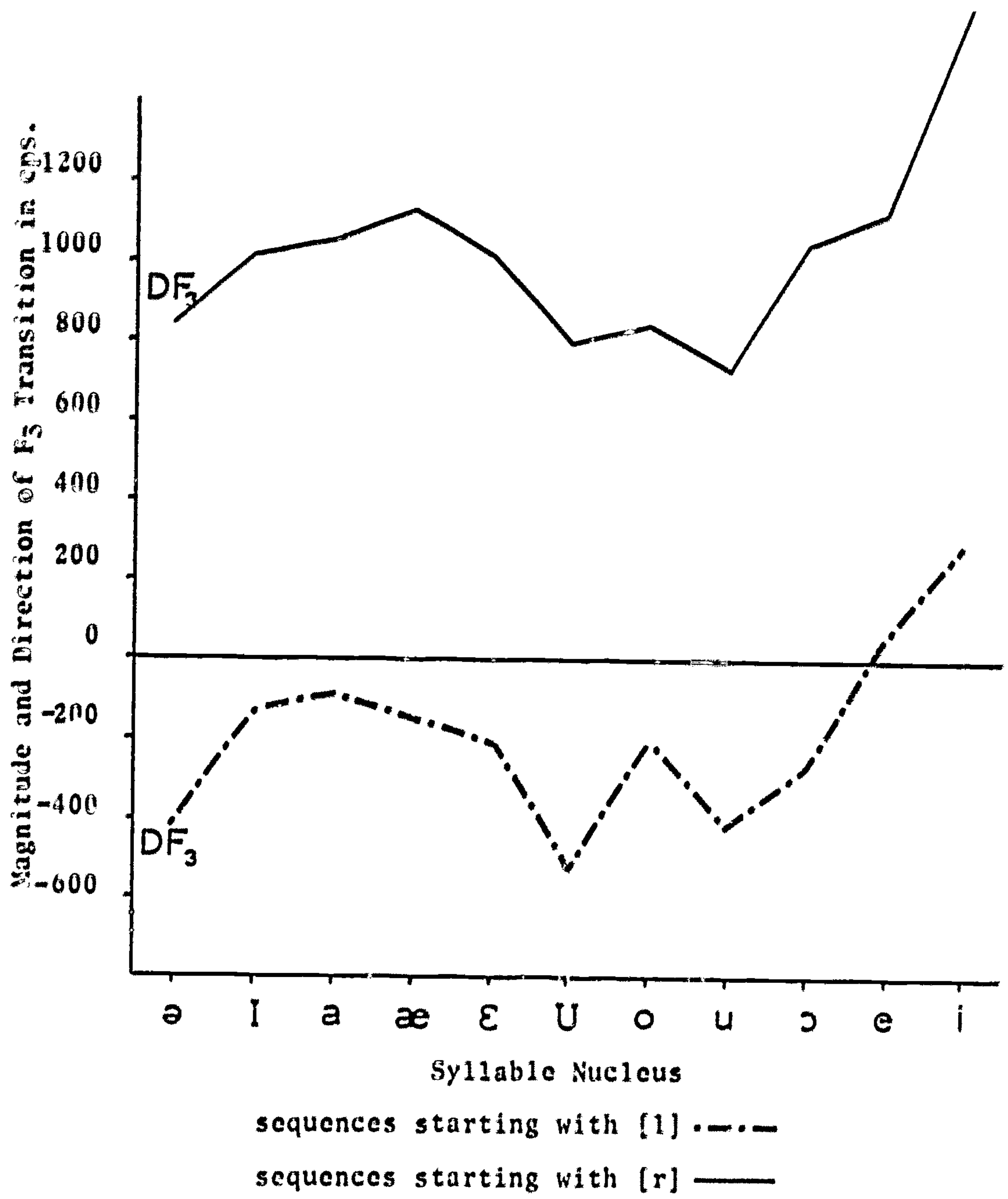


FIGURE 15

AE Initial /l/ and /r/ to Nine Following Vowels
Where the transitions from /l/ and /r/
are Opposite in Direction as a Function of CV Interaction Pair

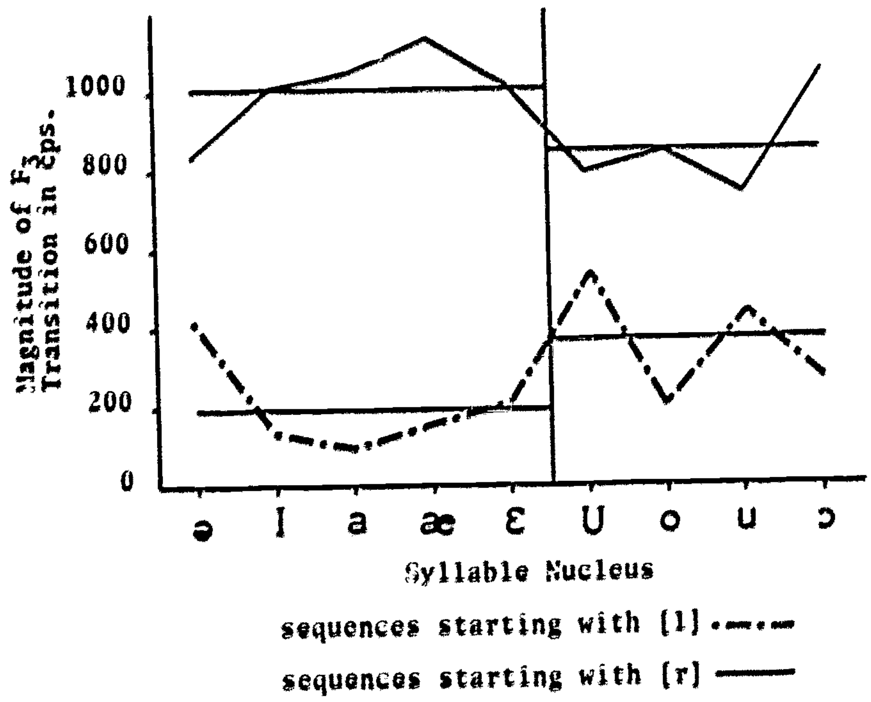


FIGURE 16

Response Scores in Percent Correct as a Function
of the Same CV Interaction Pairs
as Shown in Figure 15

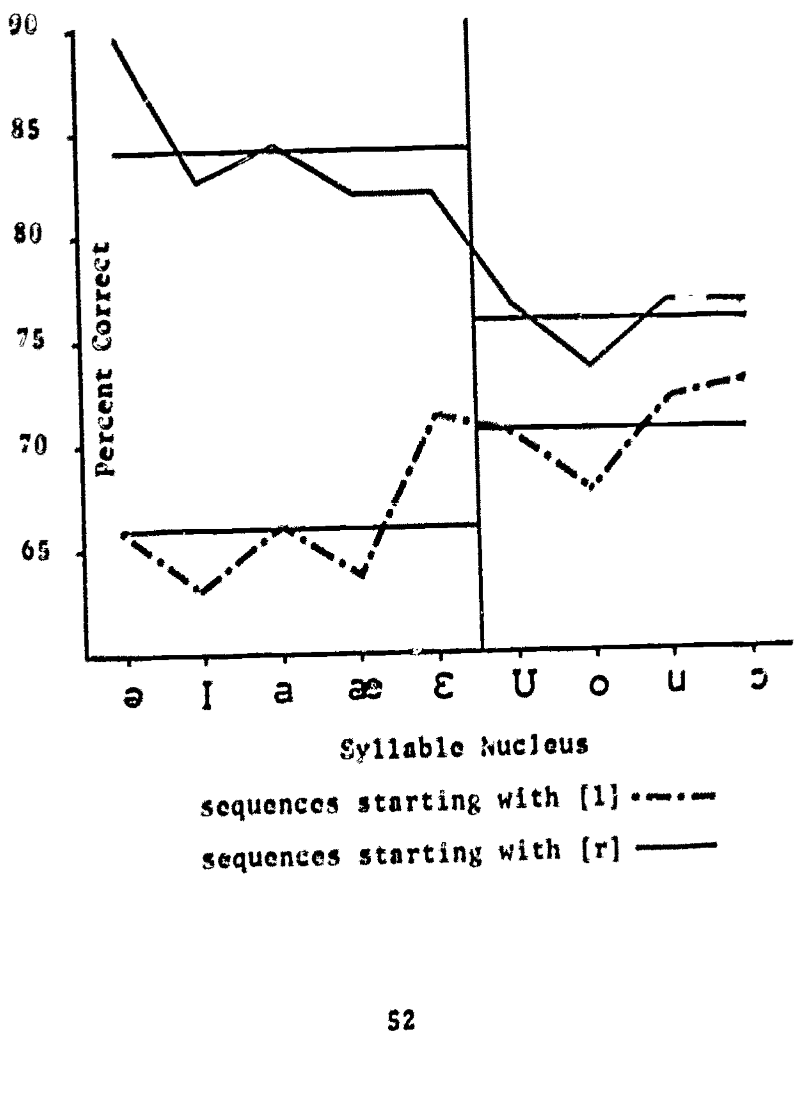
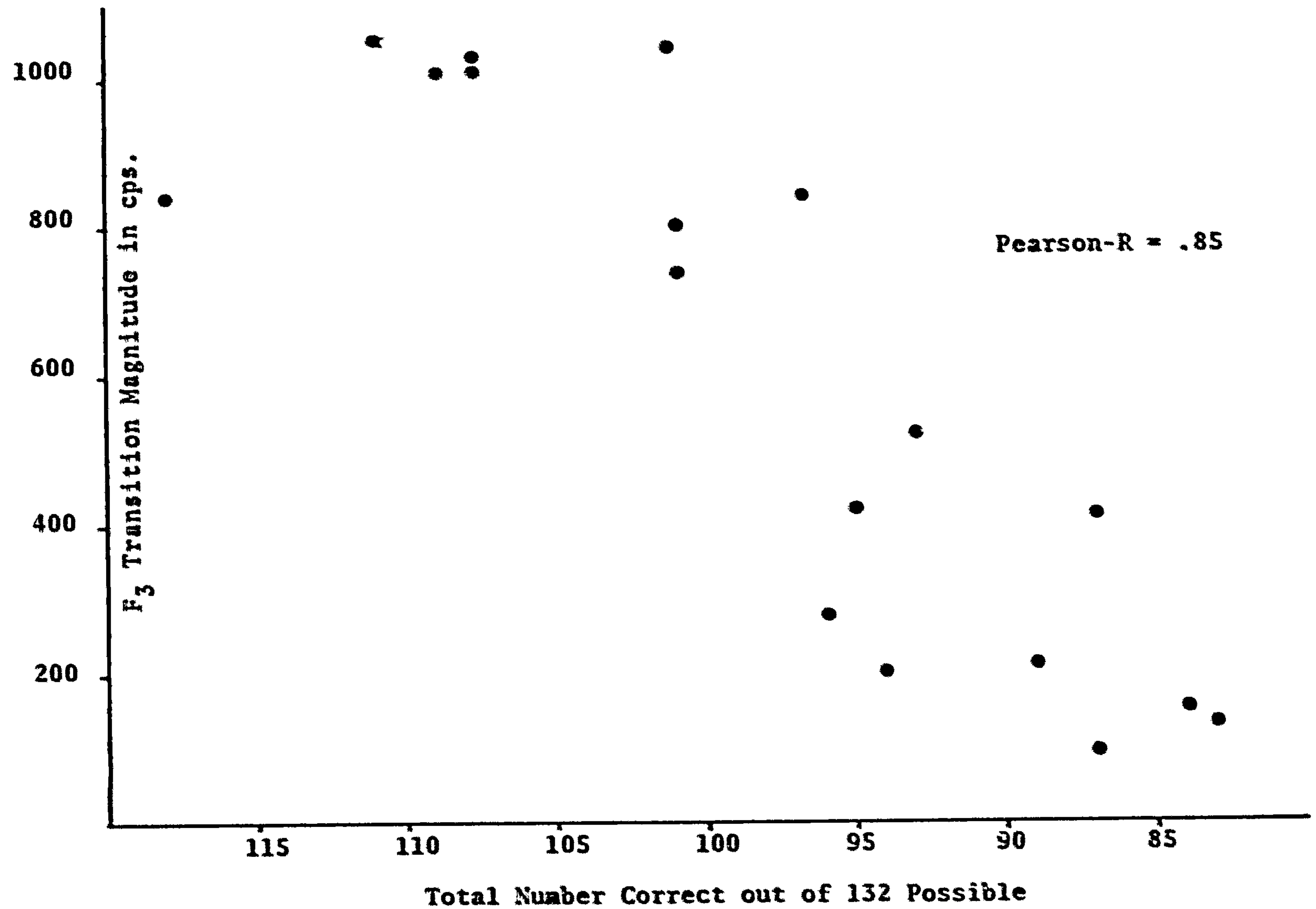


FIGURE 17

Scatter Plot of Magnitude of F_3 Transition as a Function of Corresponding Response Scores for CV Interaction Pairs in Figures 15 and 16



recognition practice, and therefore tended to mark /r/ for syllables containing /ə/ if they were in doubt. Actually a single S with an R bias when the syllable nucleus was /ə/ could have caused the deviation. Yet the deviation of the /rə/ RS from that expected on the basis of the F_3 transition lies within expected limits.

Since there were indications the Ss' exposure to various levels of C correspondence to RS, estimations of exposure were also made for the 28 CV interaction pairs. Their estimations were then graphed in Figures 18, 19, and 20. Figure 18 shows the Rinsland total-count defined frequencies for the various interaction pairs; Figure 19 shows the number of different stimulus syllables found in the stress position of English words, and Figure 20 the number of stimulus syllables which constitute English monosyllabic words. The data are presented in order to decreasing RS difference between /l/ and the /r/ syllable for each syllable nucleus.

There appears to be no (significant) correlation between the various estimations of exposure and the RS for the various CV pairs. While the frequencies associated with /l/ and /r/ syllables containing /V_{1-6}/ have a higher mean frequency (cf., Figures 18 and 19) and this corresponds to the magnitude of the response difference score, it is directly counter to the RS scores themselves. Generally, the frequencies of occurrence for /l/ + /V_{1-6}/ are as high

FIGURE 18

Sum of Rinsland Total-Count Defined Frequencies for Stimulus Syllables that Occur in English as a Monosyllabic Word as a Function of CV Interaction Pair

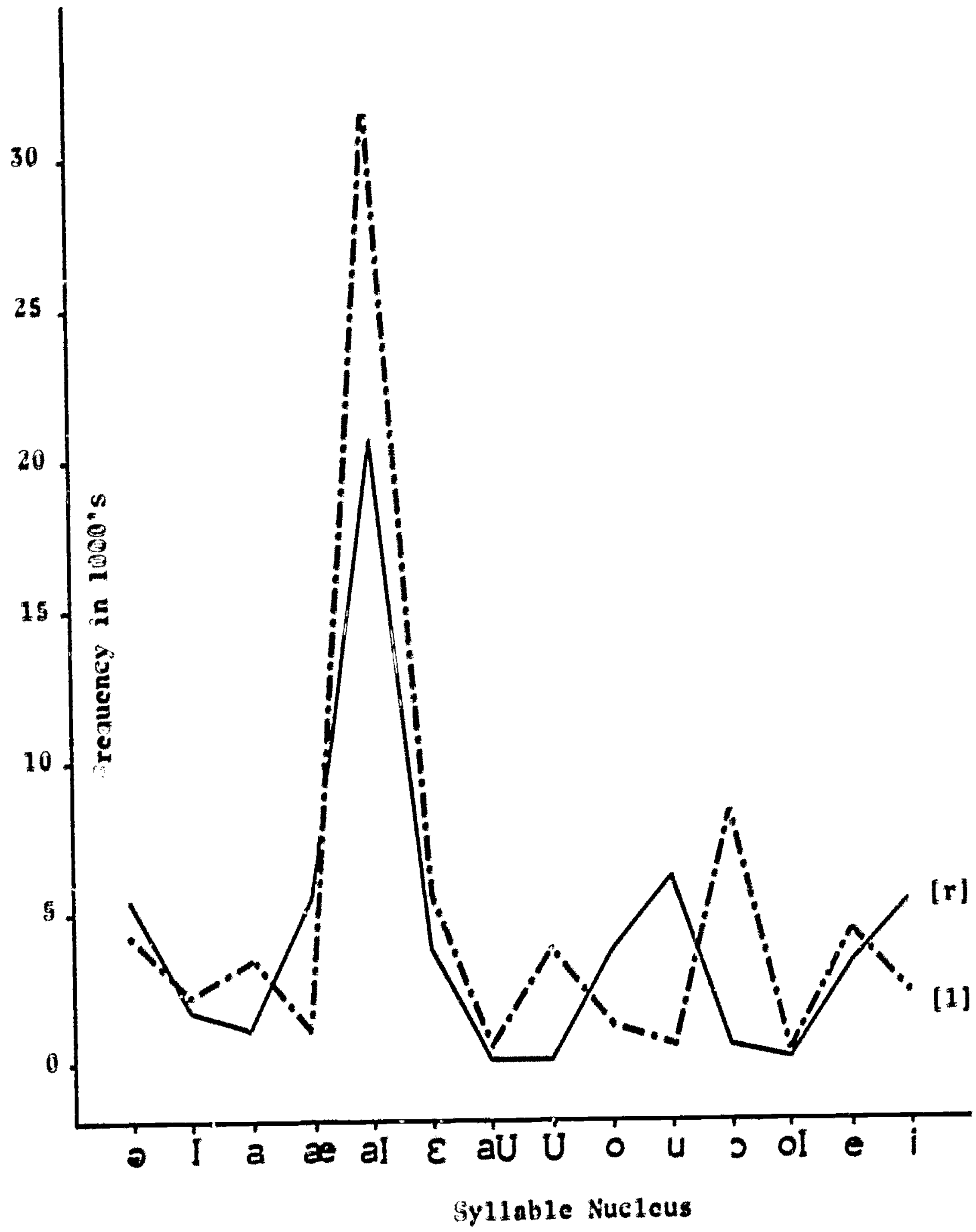


FIGURE 19

Number of Stimulus Syllables that are Found in Stress Position of English Words as a Function of CV Interaction Pair

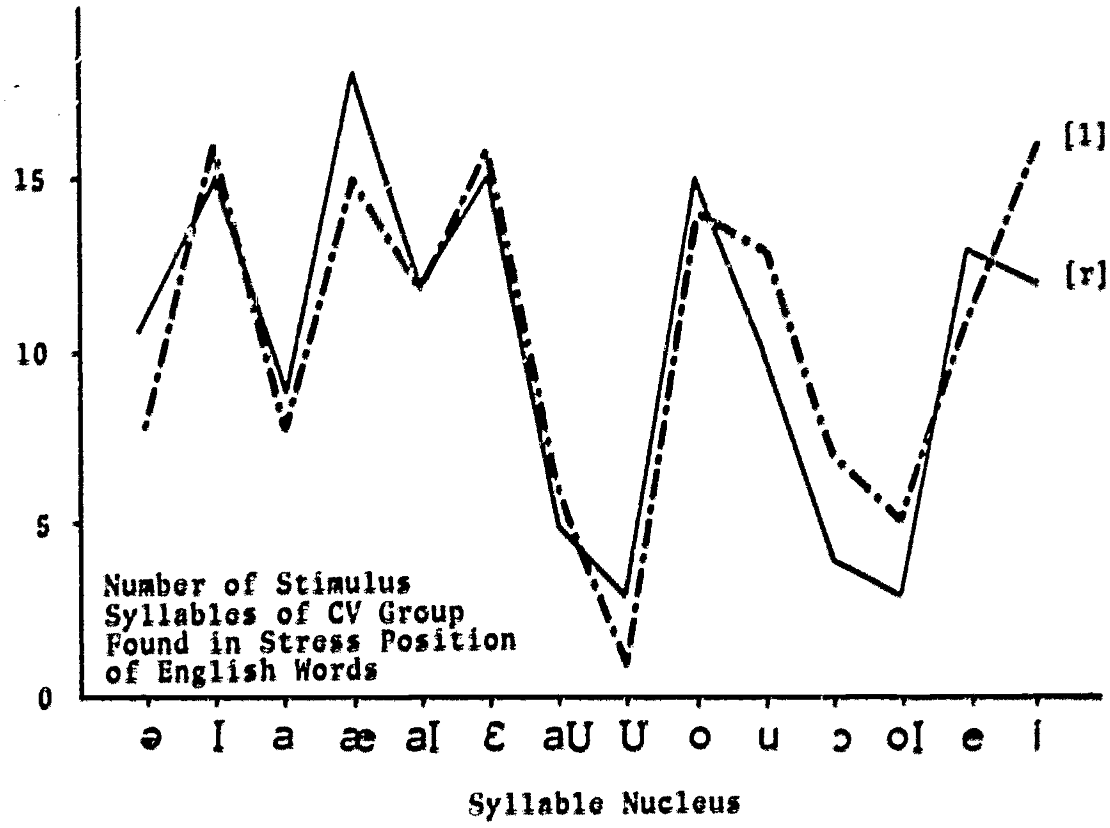
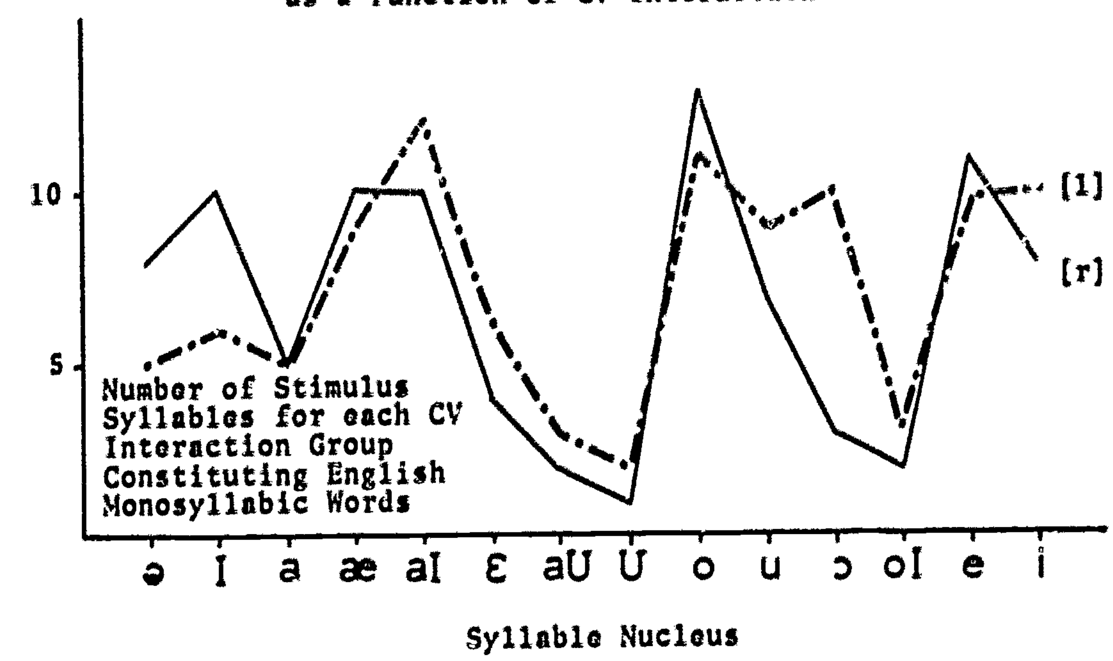


FIGURE 20

Number of Stimulus Syllables Constituting English Monosyllabic Words as a Function of CV Interaction Pair



as the frequencies for /r/ + /V_{1-6}/, yet the RS for /l/ + /V_{1-6}/ are significantly lower than those for /r/ + /V_{1-6}/ at the .01 level according to the Scheffé procedure.

There is no observable effect on the RS for /l/ and /r/ syllables containing vowels /ε, o/ which could be related to the free variation of [l] and [r] before those vowels in some speakers of Japanese (cf., Figure 21). The distribution of the Japanese allophone [ɾ] before the Japanese /i, e/ parallels the apparent difference reversal between the RS of /l/ and /r/ syllables containing those vowels. It seems less likely that the RS were caused by that allophonic distribution than that both the Japanese [ɾ] before the Japanese /i, e/ and RS for /l/ and /r/ syllables containing /i, e/ correspond in some way to the high F_2 and F_3 positions of these Japanese and English vowels.

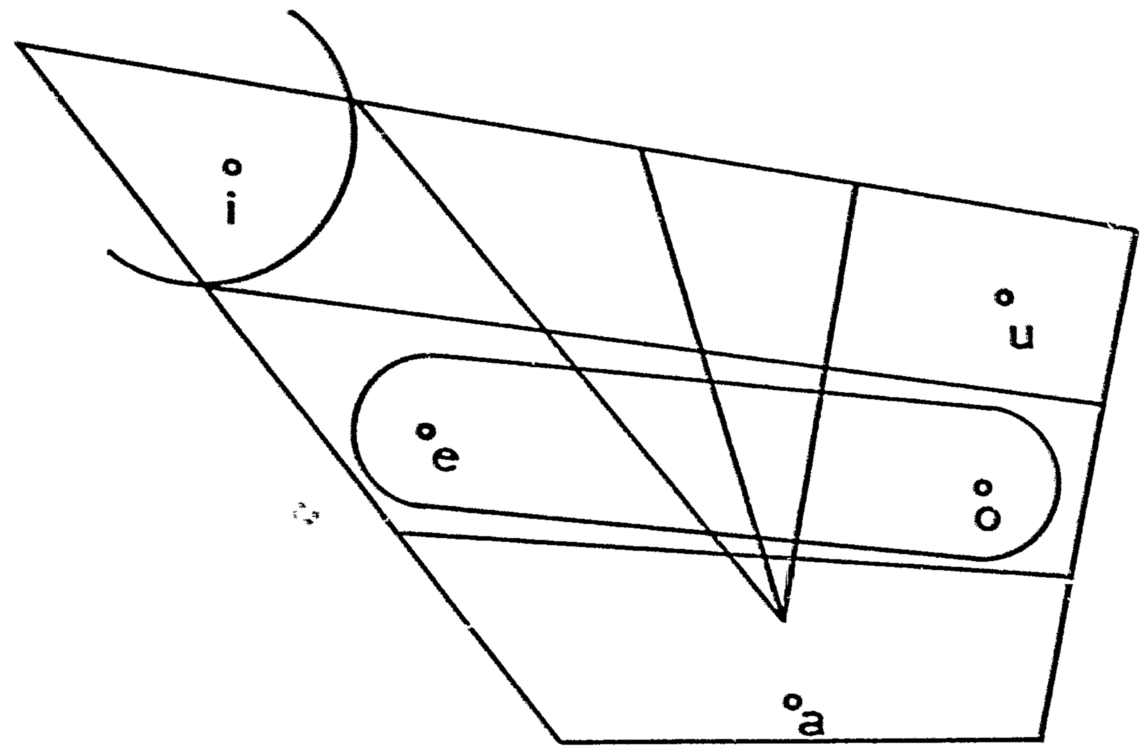
Summary of the Results

An analysis of variance showed the final consonant context in CVC syllables to have a significant main effect on the identification of initial /l/ and /r/. The final consonants /θ, ð, ǰ/ corresponded to significantly lower /l/ and /r/ identification scores than final consonants /l, Ø/ at the .01 level according to the Duncan procedure. (Ø indicates the absence of a final consonant.)

As expected, the presence of a final consonant

FIGURE 21

Allophonic Distribution of Japanese /r/
According to Bloch (1950)



[r] before /i,y/

[l] in free variation with [r] before /e,o/ for some speakers, otherwise [r]

probably hinders the identification of initial /l/ and /r/. Thus, a relatively high identification score was observed for syllables with no final consonant.

Since the Ss were given the final consonant context on the score sheet, the task was changed from one of identification to one of comparison for final /l/ and /r/. Then the task was to determine if the initial consonant was the same as or different from the final consonant. Apparently the initial and final varieties of /l/ were more comparable for the Ss than those of /r/, although the mean number correct for CV/r/ was not significantly different from that of CV/l/ at the .10 level according to the Duncan procedure.

There was no evidence of any significant difference in responses to "real" and "non-real" syllables in general. It was thought that identification of /l/ and /r/ might have been extra low for syllables ending in /θ, ð, ǯ/ because of their extreme rarity. A threshold effect requiring a certain amount of contact with that syllable type in order to identify the initial /l/ and /r/ nominally might yield such results.

At best the effect for /C/ was weak. As the lowest scores occurred for the rarest syllable types, no major emphasis on training these syllables would seem to be warranted.

There was no main effect for the post /l/ and /r/ V (vowel) context. In other words, there was no group of one or more vowels for which the identification scores of both /l/ and /r/ were significantly higher or lower than any others. However, there was a very strong interaction effect between the initial C and the V.

/r/ before /ə, I, a, æ, aI, ε/ was identified significantly better than /l/ before /ə, I, a, æ, aI, ε/. That is, /r/ seemed to be identified significantly better before unrounded vowels than did /l/. Yet, conspicuously absent from that generalization were the unrounded high front /i, e/.

Examination of the third formant transitions from /C-/ to /-V-/ as reported in Lehiste (1964) revealed the third formant transitions from /l/ and /r/ to /i, e/ differed in magnitude, but not in direction. Third formant transitions from /l/ and /r/ to /ə, I, a, æ, ε/ differed in both. The identification of initial /l/ and /r/ bears close relation to the magnitude of the third formant transition from /C-/ to /-V-/ providing the transitions from /l/ and /r/ for that context are opposite in direction. There was no relation between the identification scores and exposure measures for the CV interaction pairs.

Apparently the Ss are acting in much the way an acoustic detector would if it were to utilize only third formant magnitude and direction cues for identification. The great difference in third formant starting positions for /l/ and

/r/ as well as abruptness and duration characteristics seemed not to have been perceived or at least seemed not to have been effectively utilized.

Apparently there is some low level means whereby the listener can infer the less perceivable and less variable characteristics. Otherwise the listener would be left with almost unending complexity and the perception for native speakers would be far more context dependent than the pilot tests indicated.

The Japanese student listeners in this experiment seem to have not yet found themselves in such control of:

1) basic cue detection, such as the detection of less context-variable "dynamic" cues; or
2) low level automatic processing of context-variable cues into stable cues, so that perception would be independent of context.

Deficiency in (1) would point toward the need for more training in detecting and utilizing those cues.

Deficiency in (2) would point to the need for improved identification of preceding and postceding phonemes as well as practice in relating various context-variable cues to one phoneme.

Which of the two would be the goal of instruction to improve the identification of /l/ and /r/ cannot be determined here. Either, if possible, would be sufficient.

REFERENCES

Anderson, T. R. A case for contrastive phonology. *International Review of Applied Linguistics*, 1964, 2, 219-230.

Björkhagen, I. *Modern Swedish Grammar.* Stockholm, Sweden: Svenska Bokförlaget, 1956.

Bloch, B. Studies in Colloquial Japanese IV. *Language*, 1950, 26, 86-125. (Also in Joos, M. *Readings in Linguistics I*, 4th ed.)

Calearo, E., & Lazzaroni, A. Speech intelligibility in relation to the speed of the message. *Laryngoscope*, 1957, 67, 410-419.

Edwards, A. L. *Experimental design in psychological research.* New York: Holt, Rinehart, and Winston, 1960. (Revised Edition)

Francis, W. N. *The structure of American English.* New York: Ronald Press Co., 1958.

French, N. U., & Steinberg, J. C. Factors governing the intelligibility of speech sounds. *Journal of Acoustic Society of America*, 1947, 19, 90-119.

Hanson, G. Phoneme perception. *Uppsala Universitets Årsskrift*, 1960, 11, 109-147.

Hardy, W. G. Problems of audition, perception and understanding. *Volta Review*, 1956, 58, 289-300, 309.

Joos, M. Acoustic phonetics. *Language* Monograph No. 23, 1948.

Kimizuka, S. Problems in teaching English based upon a contrastive analysis of Japanese and English. Unpublished master's thesis, University of California, Los Angeles, 1962.

Nakajima, F. Comparison of Japanese and English. Paper presented at the Specialists' Conference, September 1956. Tokyo, Japan: Kenkyusha Ltd., 1957.

Lehiste, I. Acoustical characteristics of selected English consonants. *International Journal of American Linguistics*, 1964, 30, part IV.

Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs,* 1954, 68, 1-13.

Lindquist, E. J. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin Co., 1953.

Lisker, L. Minimal cues for separating /w, r, l, y/ in intervocalic position. *Word,* 1957, 13, 256-267.

Luescher, E., & Zwislocki, J. Adaptation of the ear to sound stimuli. *Journal of the Acoustic Society of America,* 1949, 21, 135-139.

Moser, H. M., & Dreher, J. J. Effects of training on listeners in intelligibility studies. *Journal of the Acoustic Society of America,* 1955, 27, 1213-1219.

O'Connor, J. D., Gerstman, L. J., Liberman, A. M., Delattre, P. C., & Cooper, F. S. Acoustic cues for the perception of initial /w, j, r, l/ in English. *Word,* 1957, 13, 24-43.

Prator, C. H., Jr. *Manual of American English pronunciation.* New York: Holt, Rinehart, and Winston, 1957. (Revised Edition)

Toyoda, M. Toward a comparative phonetics of Japanese and English. Paper presented at the Specialists' Conference, September 1956. Tokyo, Japan: Kenkyusha Ltd., 1957.

Rinsland, H. D. *A basic vocabulary of elementary school children.* New York: MacMillan Co., 1945.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

Appendix A

You will hear one word for each number. The word will be one of the two words beside the number.

Make a circle around the word you hear. If you are not sure make a circle around the word which is nearer to the one you hear. BE SURE you make a circle around one word and only one word.

After five words you will hear a number. Be sure you are ready for that number after you hear it.

Let's do some examples.

1. rane lane
2. roge loge
3. lan ran
4. raws laws
5. lace race

6. lear rear

Are there any questions?

APPENDIX B

l-l	Real Eng. Words	*	Rins. Freq.	**	r-l	Real Eng. Words	*	Rins. Freq.	**
i					i	real + reel[28]	M	1209	2
I	lift	S			I	rill	M	5	
e					e	rail	M	89	
ε					ε	relegate	S		
æ					æ	rally	S		
a	loll	M	0	2	a	rolick	S		2
ɔ					ɔ				
o					o	role[7] + roll	M	497	
U	lull	M			U				
u					u	rule	M	157	
ə					ə				
aI	lilac	S			aI	rile	M	0	
aU					aU				
oI	loyal	M*	45		oI	royal	M*	42	
			45					1999	

l-∅					r-∅				
i	lea	M	3		i				
I					I				
e	lay	M	914		e	ray	M	0	
ε					ε				
æ					æ				
a					a	rah	M		
ɔ	law	M	544	2	ɔ	raw	M	174	
o	low	M	563		o	row (split 300)	M	150	
U					U				
u	lieu	M	0		u				

ə					ə				
aI	lie + lye[14]	M	234		aI	rye	M	71	
aU	allow	S			aU	row (split 300)	M	150	
oI					oI				
			2258					545	

<u>l-g</u> <u>r-g</u>

i	league	M	114		i	regal	S		
I	ligament	S			I	rig	M	20	
e					e				
ɛ	leg	M	788	2	ɛ	regulate	S		
æ	lag	M	24		æ	rag	M	186	
a					a				
ɔ	log	M	530		ɔ				
o					o	rogue	M	0	
U					U				
u					u				
ə	lug	M	0		ə	rug	M	254	
aI					aI				
aU					aU				
oI					oI				
			1456					460	

<u>l-n</u> <u>r-n</u>

i	lean	M	79		i				
I	lint	S			I				
e	lane + lain[27]	M	84		e	rain + reign[13]	M	1587	
ɛ	lend	S			ɛ	rent	S		

æ	land	S			æ	ran	M	4180	
a					a				
ɔ	lawn	M	160	2	ɔ				
o	loan	M	24		o	roan	M	0	
U					U				
u	lunatic	S			u	maroon	S		
ə	lunge	S			ə	run	M	4373	
aI	line	M	1265		aI	rind	M	7	
aU					aU	round	S		
oI	loin	M	4		oI				
			1616					10,147	

l-č **r-č**

i	leech	M	0		i	reach	M	420	
I					I	rich	M	707	
e					e				
ɛ	letcher	S			ɛ	wretch	M	0	
æ	latch	M	44		æ	rachet	S		
a					a				
ɔ					ɔ				
o					o	roach	M	0	
U					U				
u					u				
ə					ə				
aI					aI	righteous	S		
aU					aU				
oI					oI				
			44					1127	

<u>l-k</u>

i	leak + look[3]	M	75	
I	lick	M	5	
e	lake	M	[illegible]	
ɛ	electric	S		
æ	lack	M	119	
a	lock	M	188	2
ɔ				
o	local	S		
U	look	M	3810	
u	luke	M	0	
ə	luck	M	491	
aI	like	M	24682	
aU				
oI				
			30,674	

<u>r-k</u>

i	reak	M	0	
I	riccochet	S		
e	rake	M	148	
ɛ	wreck	M	0	
æ	rack	M	56	
a	rock	M	884	2
ɔ				
o	baroque	S		
U	rook	M	0	
u				
ə				
aI				
aU				
oI				
			1088	

<u>l-s</u>

i	lease	M	0	
I	listen	S		
e	lace	M	228	
ɛ	less	M	250	
æ	alas	S		
a				
ɔ	loss	M	48	2
o				
U				
u	loose	M	354	

<u>r-s</u>

i				
I	risk	S		
e	race	M	636	
ɛ	rest	S		
æ	harass	S		
a				
ɔ				
o	roast	S		
U				
u	roost	S		

ə					ə	rust	S		
aI	lice	M	30		aI	rice	M	405	
aU	louse	M	0		aU				
oI					oI				
			910					1041	

l-r r-r

i					i				
I	leer	M	0	2	I	rear	M	55	2
e	layer	M*	43		e				
ɛ	lair	M	3		ɛ	rare	M	45	
æ					æ				
a	large	S			a				
ɔ					ɔ				
o	lore + lower	M	212		o	roar	M	72	
U					U				
u	lure	M	0		u	rural	S		
ə					ə				
oI					oI				
aU					aU				
oI					oI				
			258					172	

l-v r-v

i	leave	M	1254		i				
I	live (split 3598)	M	1799		I	rivit	S		
e					e	rave	M		
ɛ	leverage	S			ɛ	revolution	S		

æ	lavish	S			æ	ravage	S		
a					a				
ɔ					ɔ				
o					o	rove	M	6	
U					U				
u	louvers	S			u				
ə	love	M	3652		ə				
aI	live (split 3598)	M	1799		aI	rival	S		
aU					aU				
oI					oI				
			8504					6	

<u>l-d</u> <u>r-d</u>

i	lead	M	358		i	read	M	3594	
I	lid	M	59		I	rid	M	115	
e	laid	M	462		e	raid	M	16	
ɛ	led	M	350		ɛ	red	M	3868	
æ	lad	M	166		æ	eradicate	S		
a					a	rod	M	111	2
ɔ	laud	M	0	2	ɔ				
o	load	M	236		o	road	M	1539	
U					U				
u	lewd	M	0		u	rude	M	52	
ə					ə	rudder	S		
aI	lied	M	3		aI	ride	M	2768	
aU	loud	M	358		aU	roudy	S		
oI	colloid	S			oI	rubberoid	S		
			1992					12,063	

l-ʒ / r-ʒ

	l-ʒ					r-ʒ			
i	legion	S			i	region	S		
I					I	ridge	M	31	
e					e	rage	M	12	
ɛ	ledge	M	19		ɛ	register	S		
æ					æ				
a	lodge	M	55	2	a				
ɔ					ɔ	erogenous	S		2
o					o				
U					U				
u					u	rouge (split 15)	M		8
ə					ə				
aI					aI				
aU					aU				
oI					oI				
			74					43	

l-z / r-z

	l-z					r-z			
i					i	reason	S		
I	lizard	S			I	risen	S		
e	lays	M	155		e	raise + rays[43] + raze	M	605	
ɛ	lesbian	S			ɛ	resin	S	0	1
æ					æ	razz	M		
a					a	resin	S	0	2
ɔ	laws	M	635		ɔ				
o	lows	M	0		o	rose	M	357	

u	lose	M	182		u				
ə					ə				
aI	lies	M	123		aI	rise	M	118	
aU	lousy	S			aU	arouse	S		
oI					oI				
			1095					1080	

<u>l-m</u> <u>r-m</u>

i	lemur	S			i	ream + reem	M	0	
I	limb	M	158		I	rim	M	65	
e	lame	M	86		e				
ɛ	lemon	S			ɛ	remedv	S		
æ	lamb	M	133		æ	ram	M	30	
a					a				
ɔ					ɔ				
o	loam	M	11		o	roam	M	254	
U					U				
u	loom	M	19		u	room	M	5264	
ə					ə	rum	M	155	
aI	lime	M	83		aI	rhyme	M	38	
aU					aU				
oI					oI				
			490					5786	

<u>l-ŋ</u> <u>r-ŋ</u>

i					i				
I	linger	S		2	I	ring	M	825	2
e					e				

ɛ	length	S		2	ɛ				
æ					æ	rang	M	307	
a					a				
ɔ	long	M	6294	2	ɔ	wrong	M	476	2
o					o				
U					U				
u					u				
ə	lung	M	13		ə	rung	M	31	
aI	lying	M	303		aI				
aU					aU				
oI					oI				
			6610					1639	

l-t | r-t

i	elite	S			i				
I	little	S			I				
e	late	M	944		e	rate	M	58	
ɛ	let	M	4173		ɛ	retina	S		
æ	latter	S			æ	rat	M	686	
a	lot	M	3225	2	a	rot	M	24	
ɔ					ɔ	wrought	M	29	
o	lotus	S			o	wrote	M	1471	
U					U				
u	loot	M	8		u	root^{100} + route$^{23\ 1/2}$	M	215	2
ə					ə	rut	M	17	
aI	light	M	1843		aI	write + right^{4005}	M	16676	

aU	lout	M	0		aU	route	M	116	
oI					oI				
			10,193					19,292	

<u>l-f</u> <u>r-f</u>

i	leaf	M	258		i	reef	M	4	
I	lift	S			I	rift	S		
e					e				
ɛ	left	S			ɛ	reference	S		
æ	laugh	M	413		æ	raft	S		
a	loft	S		2	a				
ɔ					ɔ				
o	loaf	M	46		o				
U					U				
u	aloof	S			u	roof	M	405	2
ə					ə	rough	M	282	
aI	life	M	1562		aI				
aU					aU				
oI					oI				
			2279					691	

<u>l-p</u> <u>r-p</u>

i	leap	M	41		i	reap	M	0	
I	lip	M	57		I	rip	M	16	
e					e	rape	M	0	
ɛ	leopard	S			ɛ	repitition	S		
æ	lap	M	162		æ	wrap^{137} + rap^{27}	M	164	
a	lop-sided	S		2	a				

ɔ					ɔ				
o	elope	S			o	rope	M	599	
U					U				
u	loop	M	43		u				
ə	voluptuous	S			ə	rupture	S		
aI					aI	ripe	M	123	
aU					aU				
oI					oI				
			303					902	

<u>l-b</u> <u>r-b</u>

i					i				
I	liberal	S			I	rib	M	7	
e	labial	S			e				
ɛ					ɛ	rebel	S		
æ	laboratory	S			æ	rabid	S		
a					a	rob	M	41	2
ɔ					ɔ				
o	lobe	M	0		o	robe	M	45	
U					U				
u	lubrication	S			u	ruby	S		
ə	land-lubber	S			ə	rub	M	119	
aI					aI				
aU					aU				
oI					oI				
			0					212	

l-ž | r-ž

i	leasure	S			i				
I					I				
e					e				
ɛ					ɛ				
æ					æ				
a					a				
ɔ					ɔ				
o	loge	M	0		o	erosion	S		
U					U				
u					u	rouge	M	15	
ə					ə				
aI					aI				
aU					aU				
oI					oI				
			0					15	

l-š | r-š

i	leash	M	13						
I					I				
e					e	ration	S		
ɛ					ɛ				
æ	lash	M	5		æ	rash	M	6	
a	galoshes	S		2	a				
ɔ					ɔ				
o	lotion	S			o				
U					U				
u					u				

ə	lush	M	0		ə	rush	M	132	
aI					aI				
aU					aU				
oI					oI				
			18					138	

<u>l-θ</u> <u>r-θ</u>

i					i	wreath	M	93	
I	lithographic	S			I				
e					e				
ɛ					ɛ				
æ	lath	M	14		æ	wrath	M	0	
a					a				
ɔ					ɔ				
o	loath	M	0		o				
U					U				
u					u	ruthless	S		
ə					ə				
aI	lithe	M	0		aI				
aU					aU				
oI					oI				
			14					93	

<u>l-ð</u> <u>r-ð</u>

i					i				
I					I	rythm	S		
e	lathe	M	0		e				
ɛ	leather	S			ɛ				

æ	lather	S			æ	rather	S		
a					a				
ɔ					ɔ				
o	loathe	M	0		o				
U					U				
u					u				
ə					ə				
aI					aI	writhe	M	0	
aU					aU				
oI					oI				
			0					0	

	l	r
i	2195	5320
I	2123	1846
e	4175	3151
ɛ	5583	3913
æ	1080	5615
a	3468	1060
ɔ	8211	679
o	1092	4990
U	3810	0
u	606	6108
ə	4156	5343
aI	31,927	20,206
aU	358	266
oI	49	42

*

M -- stimulus syllable is a real English monosyllabic word without clustering.

S -- stimulus syllable occurs only as a stressed syllable in a real English word. Clustering included here.

M*-- Sometimes pronounced as one syllable.

** Number of stimulus syllables.