Phoneme Repetition in the Structure of Languages

Donald G. MacKay*

University of California

Los Angeles, California (UCLA)

*This work was supported in part by an M.I.T. Intramural Fellowship, a NASA grant (Ns G496) to Dr. H. L. Teuber, UCLA Grant 2428, and funds provided by Dr. F. N. Jones.

The author wishes to thank Drs. Eric Holman, W. Wickelgren, H. L. Teuber, J. Fodor, and M. Garrett for their excellent suggestions, and J. Nosaki and K. Acevski for collecting and analyzing the data.

Abstract

The structure of phoneme repetition in Croatian and Hawaiian was found to be remarkably similar. In both languages, immediate repetition of phonemes as in Aachen was very infrequent, but phoneme repetition after some degree of separation as in PROPER was significantly more frequent than chance expectation.

The degree of separation for maximum probability of repetition was slightly different for vowels and consonants in both languages. This pattern of phoneme repetition was unrelated to syllable length, word length or word frequency in these languages.

The hypothesis was advanced that this pattern of repetition resulted from an evolutionary process, reflected not only in recorded phonological changes in the history of languages, but also in errors in speech, and phonetic changes at rapid rates of speech, all of which frequently involve repeated phonemes.

Introduction

Recent studies have shown the importance of phoneme repetition in speech perception and production (MacKay, forthcoming). The present investigation determined whether phoneme repetition in the structure of two unrelated languages was significantly different from chance.

Some of the questions considered were as follows: Do vowels and consonants have identical patterns of repetition? Is the structure of phoneme repetition in different languages similar, i.e., is the structure of phoneme repetition a linguistic universal (as defined in Greenberg, 1963)? Does the probability of phoneme repetition depend on either word frequency or word length? Do syllabic factors influence the structure of phoneme repetition in a language? Is there evidence of evolutionary changes in the history of languages which mold the structure of phoneme repetition? Do repeated phonemes present a problem in natural speech production?

Study 1: Phoneme Repetition in Croatian

In the first study, the structure of phoneme repetition in Croatian was calculated in a way that allowed us to determine a null hypotheses--the probability of phoneme repetition by chance.

Croatian was chosen in this study, mainly for reasons of convenience, since the 1 to 1 correspondence between letters and phonemes in Croatian enabled us to use the Croatian dictionary as a corpus of phonemically transcribed words.

Sampling Procedure

Two hundred and fifty-eight words, all 10 phonemes long, were selected from Filipovic (1955) using the following sampling procedure:

First, a set of 5 chapters of the dictionary was randomly selected. The chapters were A, B, D, L and S. All words of length 10 in these chapters were tabulated. Repetitions of vowels and consonants in these words were separately marked. The degree of separation of the repeated phonemes was then determined, e.g., the repeated A's in AACHEN being separation 0, and the A's in ANALIZE being separation 1, and so on. The frequency of phoneme repetitions was then calculated for each gap length or degree of separation and is shown in Table 1.

Insert Table 1 about here

The Probability of Repetition

By itself the frquency of phoneme repetition has little meaning since repetitions at some gap lengths or degrees of separation are statistically more likely than at others. For example, in a word 10 phonemes long, phoneme repetitions of separation 9 are impossible; and only 1 repetition per 10 phoneme word with separation 8 is possible, 2 repetitions with separation 7, 3 with separation 6 and so on. In theory the maximum possible number of repetitions per word is (L-s-1) where *L* is the length of the word, and *s* is the degree of separation of the repeated phonemes. Thus the maximum number of repetitions in any corpus is n (L-s-1), where *n* is the number of words in the corpus.

The probability of repetition in our corpus was then calculated as:

$$PR = \frac{\text{Actual Frequency of Repetition}}{\text{Possible Frequency of Repetition}}$$

$$= \frac{\text{Actual Frequency of Repetition}}{n\ (L-s-1)}$$

for each degree of separation, where L was 10 and n was 258.

When the same phoneme occurred more than twice in a word the long range gap lengths were calculated. Thus for the repeated R's in PRORATER three degrees of separation would be tabulated 1, 3 and 5. For consonants these long range repetitions only accounted for about 2% of the data. This procedure was necessitated by the mathematical structure of our null hypothesis, discussed below.

The Null Hypothesis

The Null Hypothesis assumes that repetition of a phoneme in a word is a chance event, dependant on the probability of the phoneme in the language. The mathematical structure of this null hypothesis is extremely simple (after Herdan, 1960). For example the chance probability of vowel repetition is:

$$P_V = \frac{\sum_{i=1}^{n} Pi}{n}$$

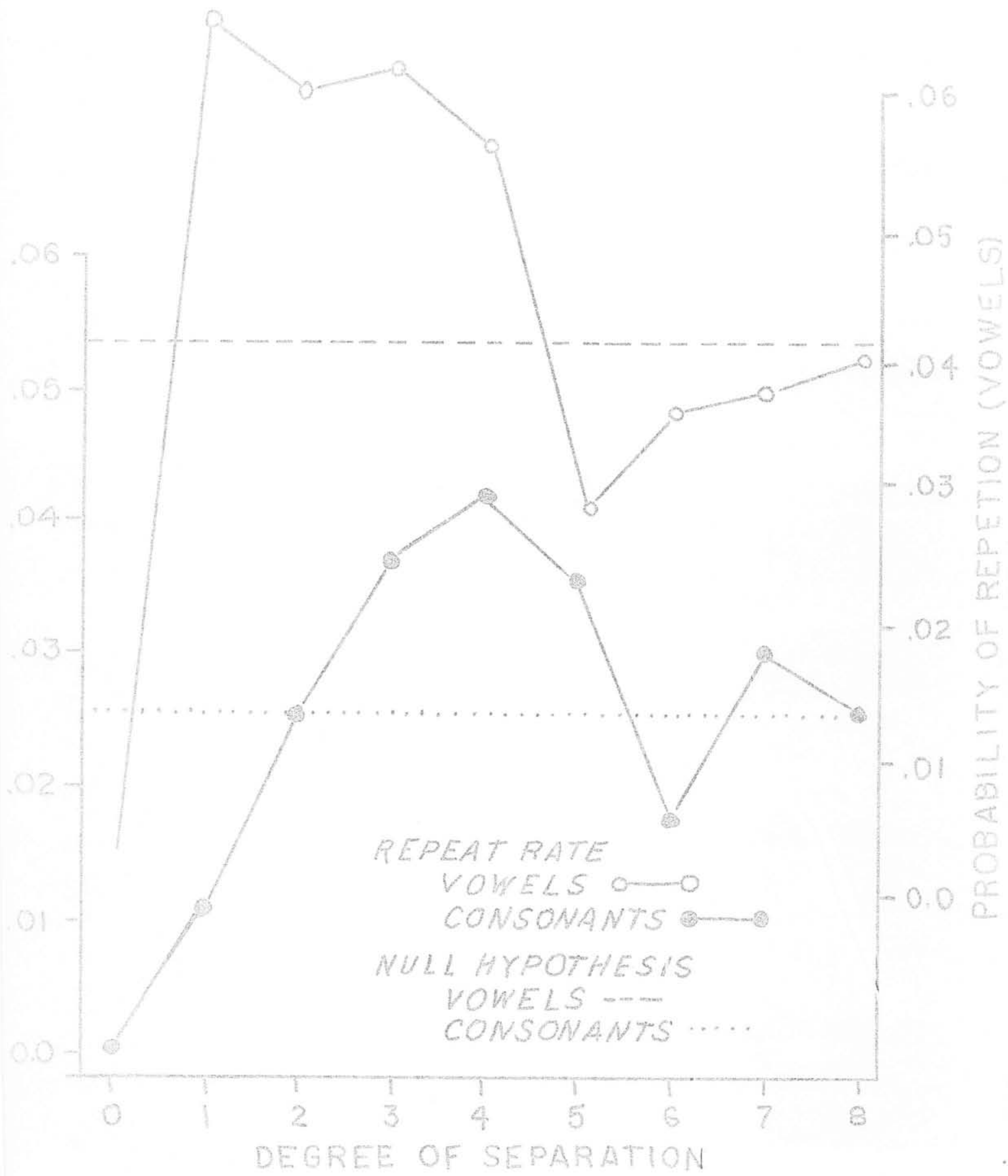
where Pi is the probability of occurrence of the i^{th} vowel, and n is the number of vowels in the language.

The probability of vowels and consonants in the sample of 258 words was determined separately. The average probability of a vowel was .042, and of a consonant .026, so that the probability of repetition of a vowel under the Null Hypothesis was .042 and of a consonant .026, for all degrees of separation as indicated by the broken lines in Figure 1.

Results

The actual probability of repetition of vowels and consonants is shown as a function of degree of separation in Figure 1.

Insert Figure 1 about here



.06
.05
.04
.03
.02
.01
0.0
.06
.05
.04
.03
.02
.01
0.0
PROBABILITY OF REPETION (VOWELS)
REPEAT RATE
VOWELS
CONSONANTS
NULL HYPOTHESIS
VOWELS
CONSONANTS
0
1
2
3
4
5
6
7
8
DEGREE OF SEPARATION

As can be seen in Figure 1, immediate repetition of phonemes was less frequent than would be expected by chance but overshot chance expectation with wider degrees of separation.

The next study was designed to allow appropriate statistical analysis of the data and to test whether the differences between vowels and consonants seen in Figure 1 could be found in other languages.

Study 2: <u>Repetition of Phonemes in Croatian and Hawaiian</u>

Several other questions motivated the second study; namely, whether the pattern of phoneme repetition depends on factors such as word length, word frequency or syllable length and whether the pattern of phoneme repetition in different languages is similar.

Hawaiian was chosen for comparison with Croatian because of the differences in phonological pattern of the two languages: Croatian has 35 phonemes and Hawaiian 13, Croatian has long consonant clusters whereas alternation of vowels and consonants is the rule in Hawaiian.

Sampling Procedure

A sample of 2051 Croatian words and 2028 Hawaiian words was obtained as follows:

Three words were arbitrarily sampled from each page of a Hawaiian and a Croatian dictionary (Pakin and Elbert, 1965); the words were then sorted for length and the actual probabilities of phoneme repetition was separately determined for each word length, as before.

Repetition of Phonemes at the Beginning and Ends of Words

It was noted that the same phoneme rarely came at the beginning and end of Hawaiian words. A statistical analysis verified that significantly fewer repetitions occurred at the beginning and ends of words than would be expected by chance (a difference significant at the .01 level using a two-tailed sign test with word length as the unit of analysis).

However, the same analysis for Croatian consonants showed the opposite tendency, the same phoneme occurring at the beginning and end of a word more frequently than chance (although not significantly so at the .10 level using a two-tailed sign test).

These findings suggest that the beginning and ends of words may impose special language-specific constraints on phoneme repetition. In order to avoid these special effects, repetitions at the beginning and ends of words were excluded from subsequent analyses.

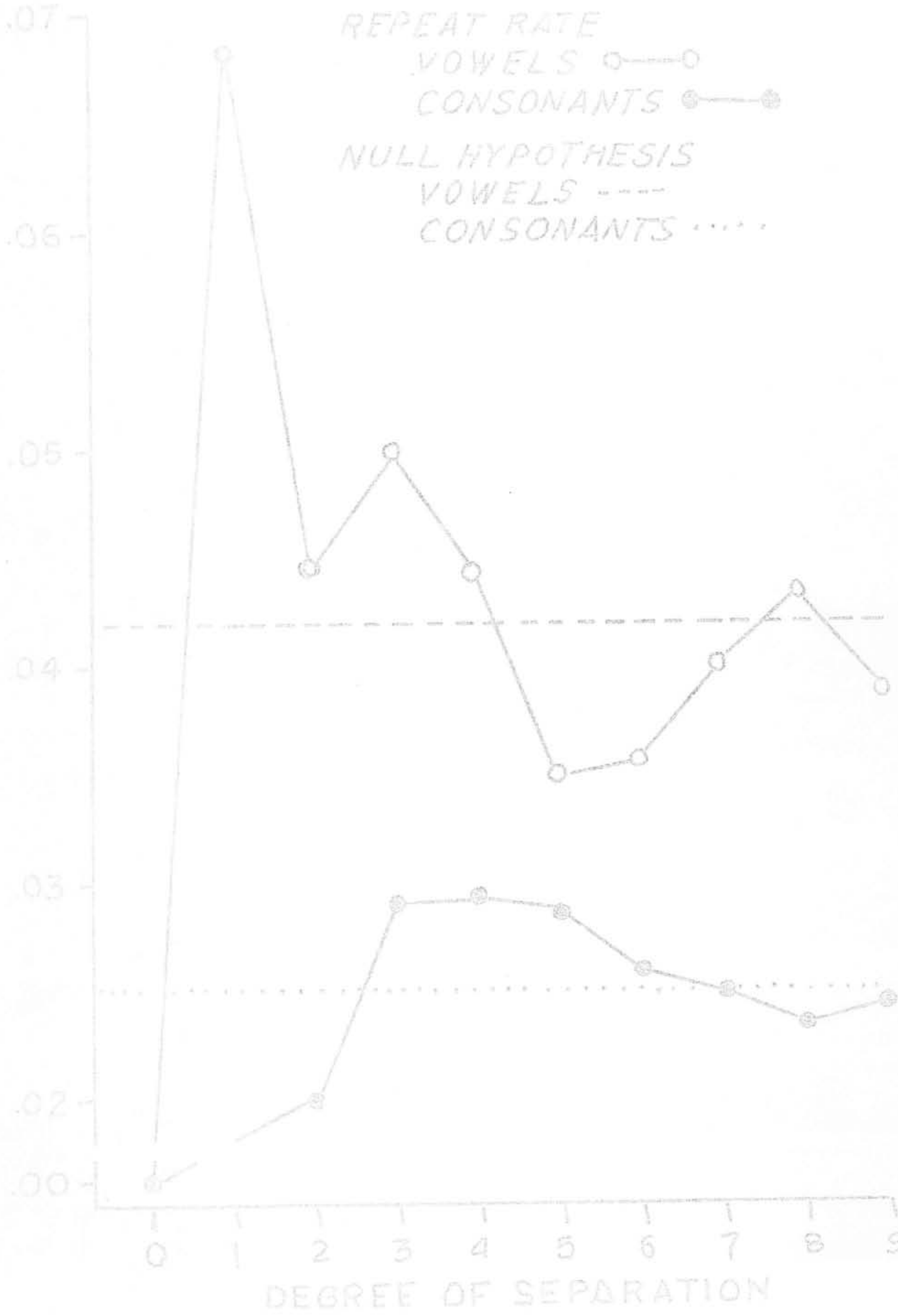
The Average Probability of Repetition

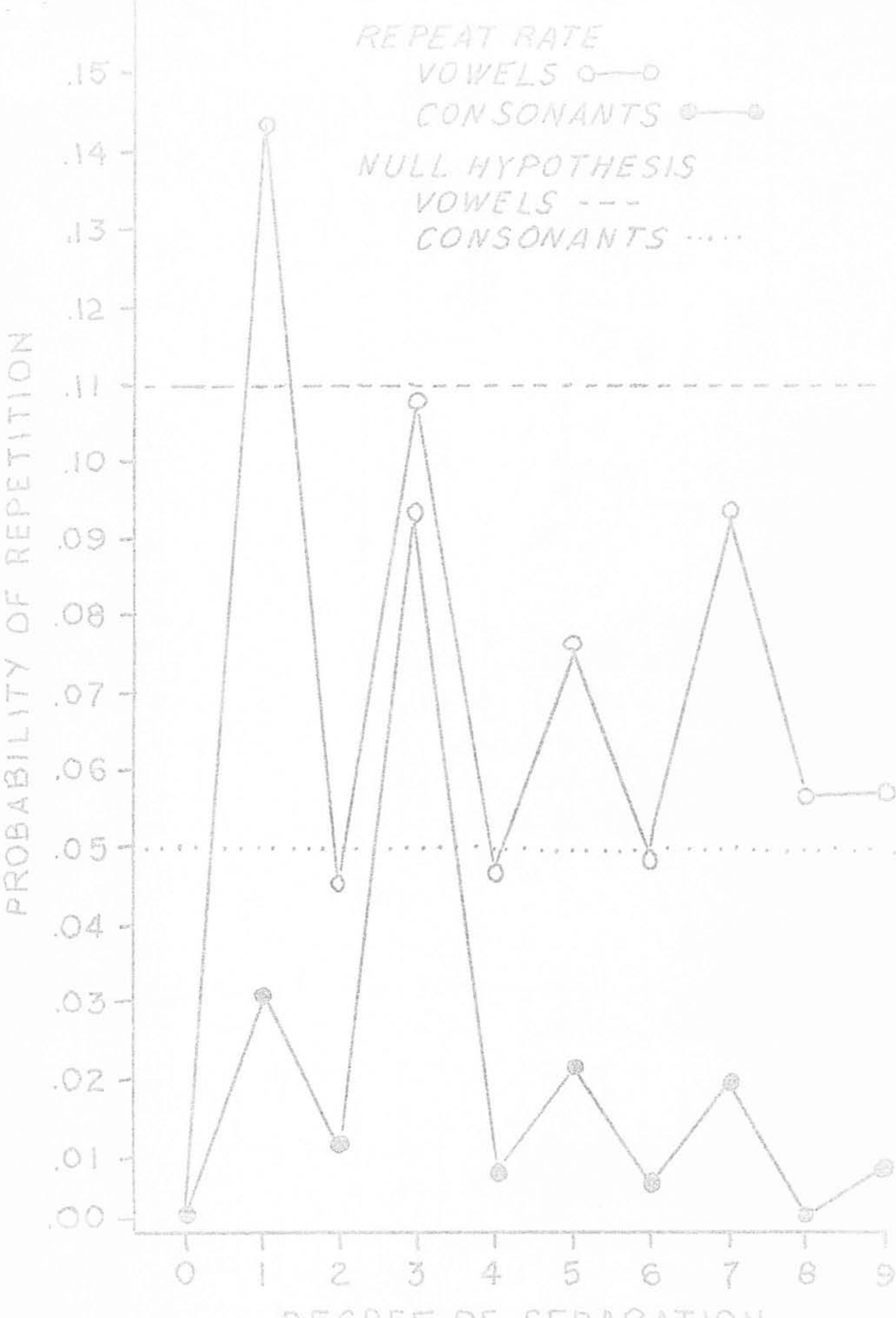
The probability of repetition of vowels and consonants is shown in Tables 2 and 3 for Croatian and in Tables 4 and 5 for Hawaiian. The average probability of phoneme repetition was then calculated, and is shown in Figure 2 for Croatian and in Figure 3 for Hawaiian.

Insert Tables 2, 3, 4, 5, about here

Insert Figures 2 and 3 about here

These phoneme repetition functions were remarkably similar for the two languages: the peak porbability of repeition of vowels came at degree of separation 1 for both languages. For both languages this peak probability of repetition was significantly greater than chance expectation (at the .01


REPEAT RATE
VOWELS
CONSONANTS
NULL HYPOTHESIS
VOWELS
CONSONANTS
.07
.06
.05
.04
.03
.02
.00
0
1
2
3
4
5
6
7
8
9
DEGREE OF SEPARATION


REPEAT RATE
VOWELS
CONSONANTS
NULL HYPOTHESIS
VOWELS
CONSONANTS
.15
.14
.13
.12
.11
.10
.09
.08
.07
.06
.05
.04
.03
.02
.01
.00
PROBABILITY OF REPETITION
0
1
2
3
4
5
6
7
8
9
DEGREE OF SEPARATION

level using a two-tailed Chi-square test. Thus for vowels the Null hypothesis can be rejected for degree of separation 1.

The peak probability of consonant repetition came at separation 3 for both languages. And for both languages this peak probability of repetition was significantly greater than chance expectation (.05 level, two-tailed Chi-square test). Thus, the phoneme repetition pattern was similar for both languages, lower than chance probability of immediate repetition, with greater than chance expectation for separation 1 for vowels and separation 3 for consonants, and chance probability of repetition for longer degrees of separation.

Consonant-Vowel Alternation

The tendency for vowels and consonants to alternate in Hawaiian can be seen in Figure 3. This alternation tendency in Hawaiian can be viewed as superimposed on but independent of the pattern of phoneme repetition. It is of interest that the pattern of phoneme repetition was the same in Hawaiian as in Croatian where the tendency for consonant-vowel alternation was absent.

Differences between Consonants and Vowels

There was no a priori reason to expect the differences in repetition probability for consonants and vowels which we found for both Hawaiian and Croatian. Further research is needed to test the universality of these findings before adding this to the growing list of differences between vowels and consonants (see Shankweiler, 1967).

Word length

No significant correlation between word length and probability of phoneme repetition was found for either language. Because of the close relation between word length and word frequency (Zipf, 1936), it seems likely that phoneme repetition is also unrelated to the frequency of words in a language.

Syllable length

One of the questions in the introduction was whether syllabic factors determined the probability of phoneme repetition in a language. Specifically if the first and last phonemes of a syllable tended to be identical or if syllables tended to start with the same first phoneme, then repetition of phonemes would reflect the average length of syllables in a language.

This hypothesis could be tested by comparing the structure of phoneme repetition in languages where the average syllable length differed. If the peak probability of phoneme repetition in the two languages was the same, we could then conclude that phoneme repetition was unrelated to average syllable length. Similarly if the peak probability of phoneme repetition varied with syllable length in different languages, the structure of phoneme repetition could be attributed to syllabic factors.

The average syllable length was found to be 1.96 phonemes per syllable for Hawaiian and 2.54 phonemes per syllable for Croatian, a difference significant at the .05 level using a two-tailed t-test. Thus even though Croatian and Hawaiian have the same pattern of phoneme repetition, the average length of syllables in the two languages differ. We can therefore conclude that the structure of phoneme repetition must be unrelated to the average length of syllables in these languages.

Discussion

In the hope of gaining some insight into our findings, the discussion will review the role of repeated phonemes shown in studies of the evolution of languages, in studies of errors in natural speech, and of rate dependent changes in speech and in studies of naturally occurring misspellings.

1) Phoneme Repetition in the Evolution of Languages

If we view the pattern of phoneme repetition (shown above) as a universal feature of human language (as defined in Greenberg, 1963), we would expect evidence of evolutionary changes which mold this particular feature of languages.

Some support for this hypothesis is seen in the history of Latin where stipipendium changed to stipendium, dropping the repeated p (Jones, 1962). We find similar changes in the history of German where Heriro became Hero, dropping the repeated r (Heffner, 1964). The frequency of changes such as these strongly suggests that phoneme repetition is a factor in the evolution of languages (Merringer and Mayer, 1895).

2) Errors in Speech

Another question was whether phoneme repetition presents a problem in normal speech production. Merringer and Mayer (1895) compiled a monumental list of errors in speech, many of which involved repeated phonemes. For example, in saying Im Institut native German speakers frequently said Im Stitut (dropping the repeated I) or Im Instut (dropping the repeated T).

There is even some evidence that repeated phonemes play a role in other "higher order" errors in speech. For example, Merringer and Mayer (1895) reported synonymic errors which also frequently involved repeated phonemes. For example, an individual attempting to say either 1 or 2 came out with 3 instead.

1) He is totally responsible.

2) He is solely responsible.

3) He is sotally responsible.

Here sotally is clearly a combination of the synonyms totally and solely, but one cannot help but wonder, as did Lecours (1966) whether the repeated T's in totally, and repeated L's in solely might not have contributed to the error as well.

3) Rate dependent changes in speech

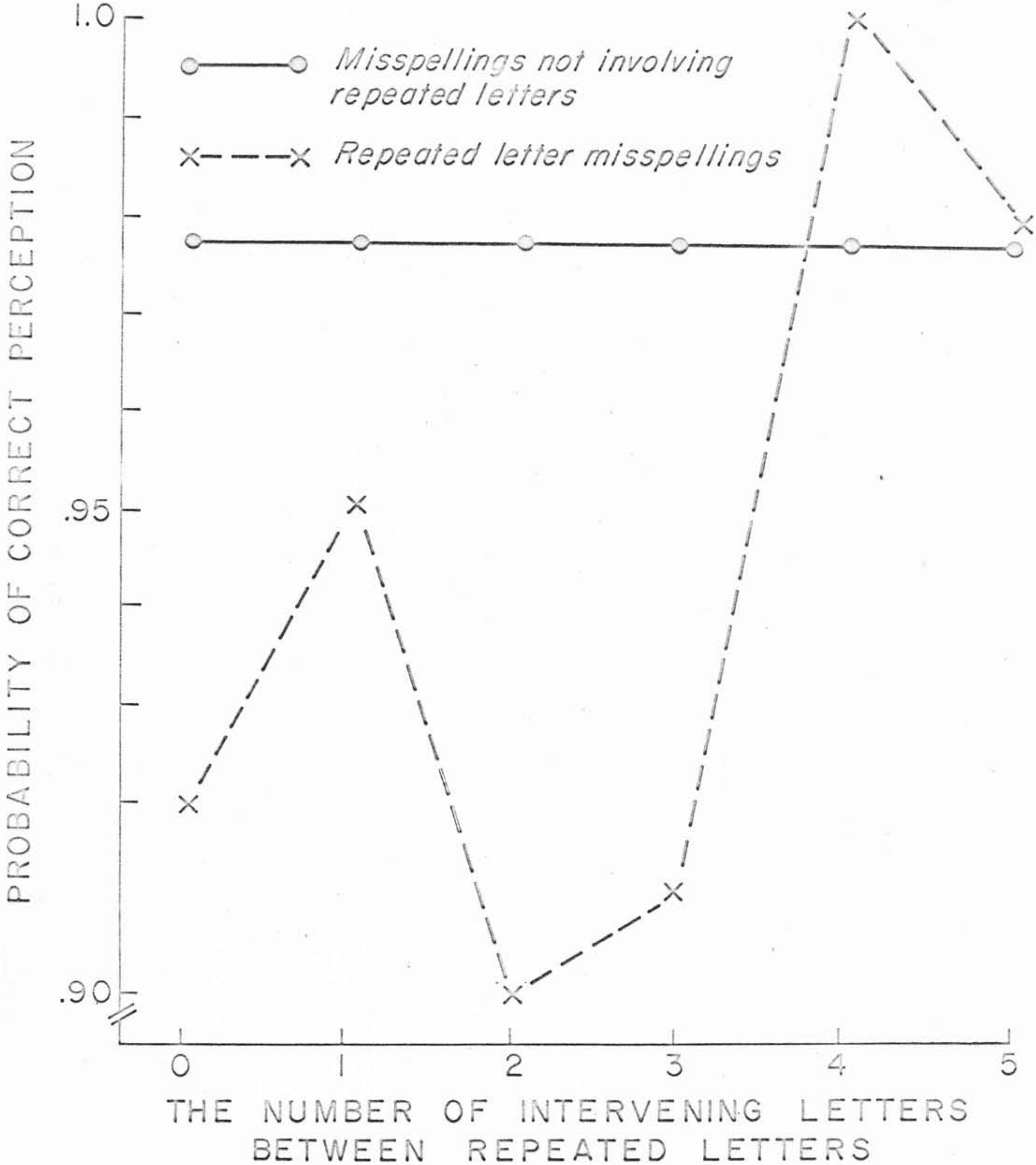
Changes in the phonetic form of words in sentences are so common at rapid rates of speech that it would be misleading to call them errors (Heffner, 1964). For example, the repeated K's in take care are rarely pronounced at rapid rates of speech. Nor are the repeated V's in We have various things (Heffner, 1964). From these examples it is apparent that phoneme repetition may also play a role in rate-dependent phonetic changes. And as suggested by Merringer and Mayer (1895) these errors in speech and changes in phonetics involving repeated phonemes may be directly related to the evolutionary changes in the history of languages which involve repeated phonemes.

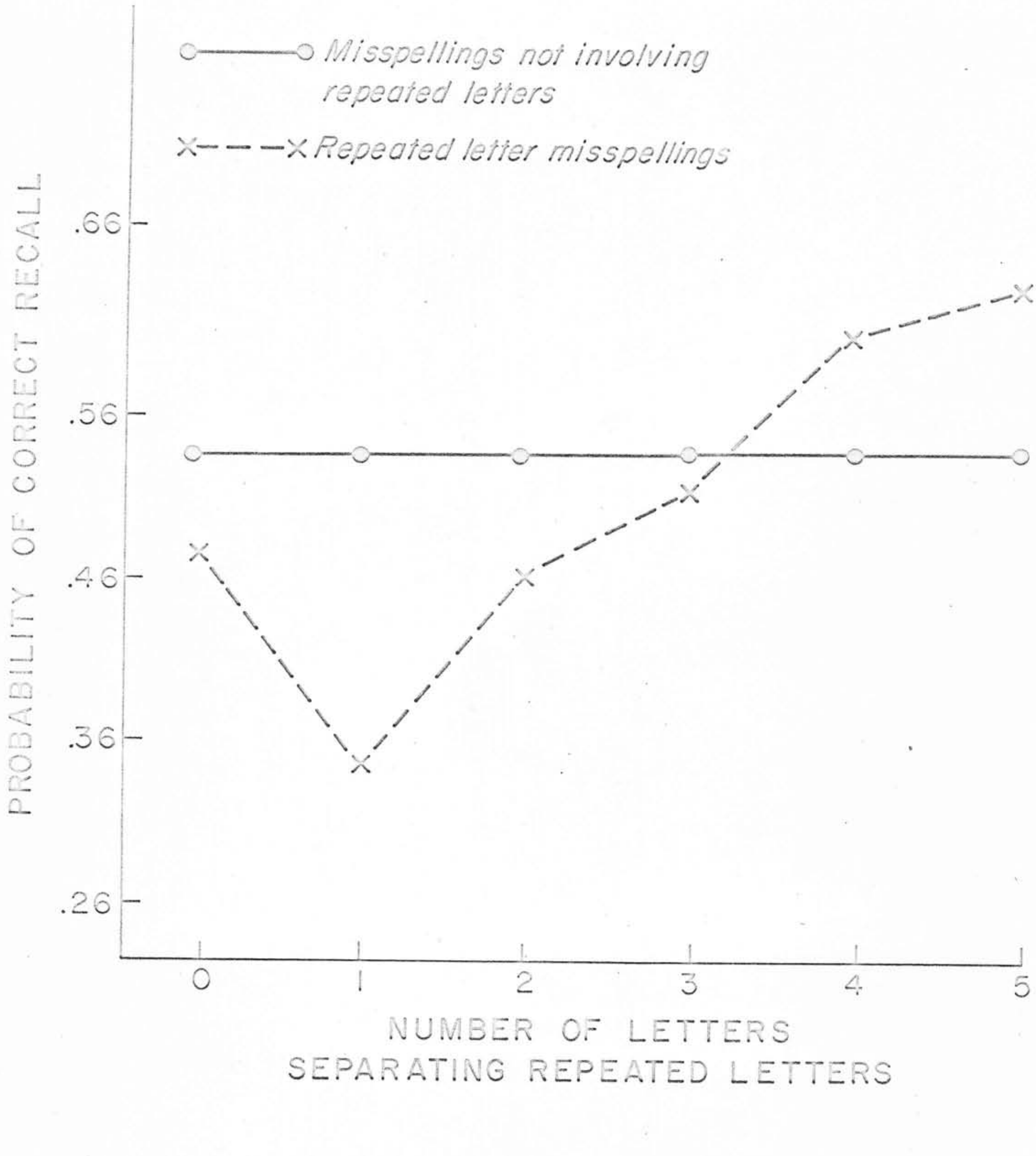
4) Repeated letters in misspellings

MacKay (forthcoming) reported an effect of phoneme repetition on the perception and recall of misspellings. His Ss rapidly read sentences containing spelling errors and later indicated the words in which they perceived a misspelling and attempted to recall how the word was misspelled. Some of the misspellings involved repeated letters and others did not. The probability of perception and recall of repeated letter misspellings involved repeated letters and others did not. The probability of perception and recall of repeated letter misspellings of various gap lengths is shown by the solid line in Figures 4and 5 respectively. The comparable data for misspellings which did not involve repeated letters is

---

Insert Figures 4 and 5 about here


1.0
.95
.90
PROBABILITY OF CORRECT PERCEPTION
Misspellings not involving repeated letters
Repeated letter misspellings
0
1
2
3
4
5
THE NUMBER OF INTERVENING LETTERS BETWEEN REPEATED LETTERS



Misspellings not involving repeated letters
Repeated letter misspellings
PROBABILITY OF CORRECT RECALL
.66
.56
.46
.36
.26
0
1
2
3
4
5
NUMBER OF LETTERS SEPARATING REPEATED LETTERS

shown with the broken straight line in these figures. These functions can be seen to closely resemble the structure of phoneme repetition in languages shown in Figures 1-3.

Naturally occurring misspellings also appear to reflect the difficulty in perception and recall of repeated letter misspellings. Lecours (1966) discovered a "repeated letter effect" in spontaneous misspellings in the diary of Lee Harvey Oswald (1964) whom he classified as a dysgraphic (someone who chronically misspells, due to cerebral lesion or developmental deficit).

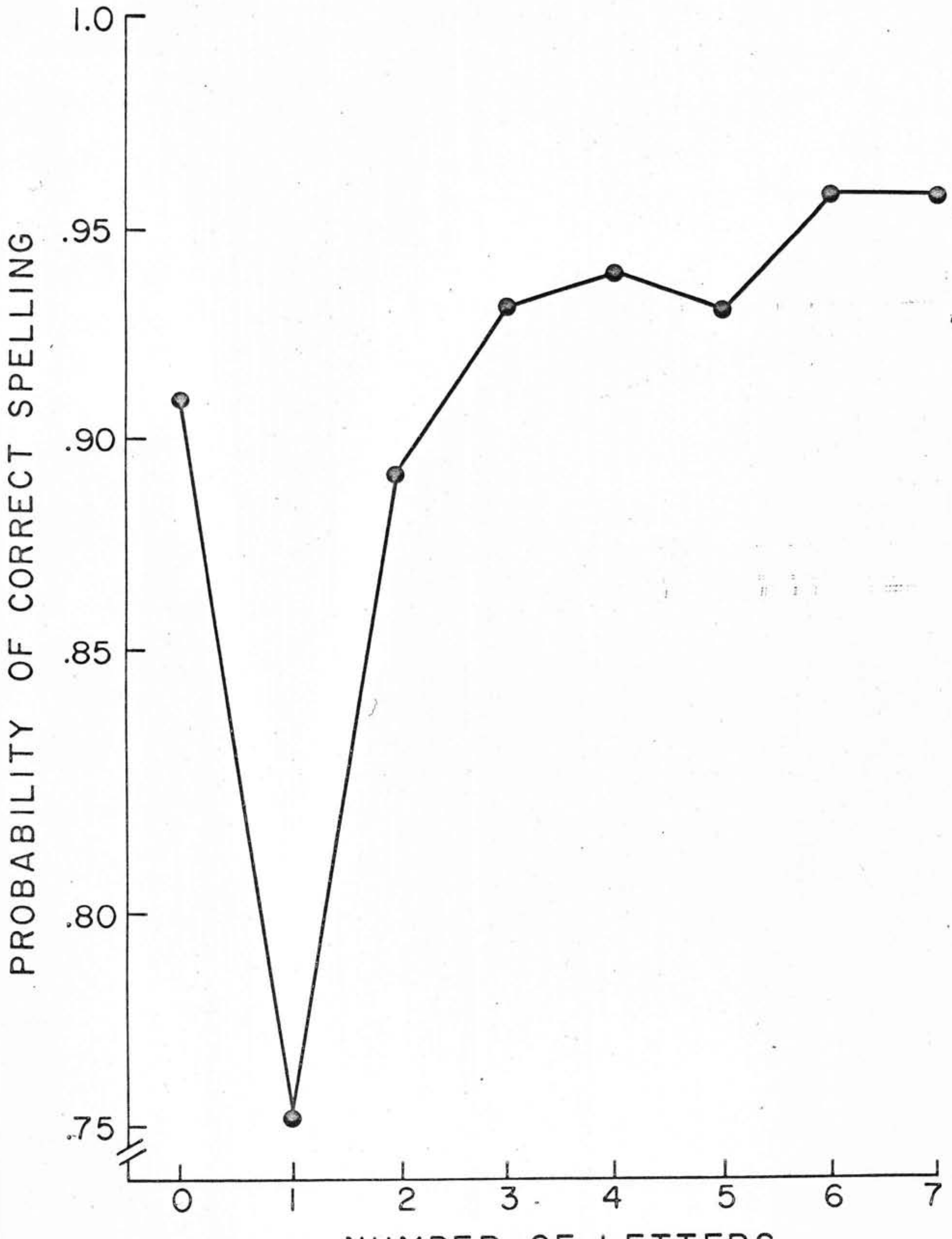
The degree of separation of repeated letter deletions such as ELDERY in Oswald's misspellings was later analyzed (MacKay, forthcoming). For example, the error in ELDERY would be classified as separation 3. As can be seen in Figure 6, the probability that Oswald would misspell a repeated letter sequence varied with the degree of separation of the repeated letters in much the same way as in Figures 4 and 5.

Insert Figure 6 about here

On the basis of these data, MacKay (forthcoming) argued that the problem of repeated elements in speech may reflect a general property of nervous action such as post-excitatory inhibition following the production of a phoneme or letter, followed by post-inhibitory rebound or facilitation (Bullock, 1965). This hypothesis is further strengthened by neurophysiological data cited below.

5) Neurophysiological Evidence

Ohala and Hirano (1967) reported a curious "second peak" of electomyographic activity in the lips of Ss producing the phoneme P. The first


PROBABILITY OF CORRECT SPELLING
1.0
.95
.90
.85
.80
.75
0
1
2
3
4
5
6
7
NUMBER OF LETTERS
SEPARATING REPEATED LETTERS

peak coincided with the actual contraction of the lip muscles, but the second, smaller peak of activity was completely unexpected, and followed about 200 m/sec. after the onset of muscle contraction. This second peak may be interpreted to reflect a general principle of nervous action such as post-inhibitory rebound, which is known to follow the activation of central neuronal aggregates by as much as 200 m/sec. (Tunturi, 1958). Moreover, we would like to suggest that the overshoot in the probability of repetition of phonemes in the structure of languages may reflect this same principle of nervous action.

Summary

A study of phoneme repetition in Croatian revealed that immediate repetition of phonemes as in AACHEN is less common than would be expected by chance. However, with wider degrees of separation as for the repeated R's and P's in PROPER, the probability of repetition significantly overshot chance expectation. This overshoot can be viewed as a law of latent alliteration or phoneme harmony in the structure of languages (after Gleason, 1961). Of course this is by no means a deterministic law, but only a statistical one, applying in general but not in every word.

These findings were replicated in Hawaiian--an unrelated language with a completely different phonological structure from Croatian. This corroboration suggests that the pattern of phoneme repetition is a universal feature of human languages.

Differences were found in the peak probability of repetition for vowels and consonants; which were consistent for Croatian and Hawaiian.

No relation between either syllable length or word length and phoneme repetition was found for these languages.

Phoneme repetition was shown to be a factor in:

1) Changes in the evolution of languages involving repeated phonemes (e.g., in the history of Latin the change from stipipendium to stipendium).

2) Errors in speech involving repeated phonemes were recorded by Merringer and Mayer (1895) in German (Instut for Institut).

3) The frequency of rate-dependent changes in phonetics involving repeated phonemes. . . . . .(e.g., lastime for last time).

4) Misspellings in the spontaneous writing of a dysgraphic (e.g., ELDERY for ELDERLY) and the difficulty in perception and recall of these misspellings by college subjects.

It was suggested that these repeated phoneme phonemena may reflect a general principle of neuromuscular action such as inhibition and post-inhibitory rebound. Nuerophysiological evidence was discussed which directly supported this principle for the production of speech.

References

Bogadele, F.A. (1952) New English-Croatian and Croatian-English dictionary, New York: Hafner.

Bullock, T.H. (1965) Mechanisms of integration, in Bullock, T.H. and Horridge, G.A. Structure and Function in the Nervous Systems of Invertebrates, San Francisco: Freeman.

Comba, P. in Pakin, M.K. and Elbert, S.H., (1965). Hawaiian-English Dictionary, Honolulu: University of Hawaii Press.

Filipovic, P. (ed.) (1955) English-Croatian Dictionary, Zagreb: Zora.

Gleason, H.A. (1961) An Introduction to Descriptive Linguistics, New York: Holt, Rinehart and Winston.

Greenberg, J.H. (ed.) (1963) Universals of Language, Cambridge: M.I.T. Press.

Grujic, B. (1961) Sysskohrvatsko-Engleski Recnik, Beograd: Stampa Crajiko.

Heffner, R-M, S. (1964) General Phonetics, Madison: University of Wisconsin Press.

Herdan, G. (1960) Type-token Mathematics. A textbook of mathematical linguistics, The Hague: Mouton.

Jones, D. (1962) An Outline of English Phonetics, Cambridge, England: Heffner and Sons.

Kiparsky, P. (1967) Phonological change, unpublished Ph.D. dissertation, M.I.T.

References (Cont.)

Lecours, A.R. (1966) Serial order in writing--a study of misspelled words in "dev elopmental dysgraphia," Neuropsychologia, 4, 221-241.

MacKay, D.G. (forthcoming). The repeated letter effect in the perception and recall of misspellings. Neuropsychologia

Merringer, R. and Mayer, K. (1895). Versprechen und Verlesen: Eine Psychologisch-linguistishce Studie, Stuttgart: Göschensche.

Ohala, T. and Hirano, M., (1967). Control mechanisms for the sequencing of neuromuscular events in speech, paper delivered at the AFCRL meeting, M.I.T., Cambridge, Mass.

Oswald, L.H. (1964) Historic Diary, Life, 25.

Pakui, M.K. and Elbert, S.H. (1965) Hawaiian-English Dictionary, Honolulu: University of Hawaii Press.

Shankweiler, D.P. (1966) Hemispheric specialization and the perception of speech, New York: Status Report on Speech Research, Haskins Laboratories.

Tunturi, A.R. (1958) in Chang, H.-L. The Evoked Potentials, in Field, J. et al (eds.) Handbook of Physiology, Washington: 1, 299-313.

Zipf, G.K. (1936) The Psycho-biology of Language, London: Routledge, 20 -49.

Tables

Table 1: The steps in analyzing the probability of repetition or repeat rate of a) consonants and b) vowels in Croatian words 10 phonemes long. The degree of separation or gap length is the number of phonemes separating repeated phonemes, e.g., the L's in ELDERLY are separation 3. The maximum possible frequency is calculated as n(L-s-1) where n is the number of words in the corpus, L the length of the words, and s the degree of separation of the repeated letters.

The repeat rate is the ratio of actual to maximum possible frequency.

The theoretical repeat rate (null hypothesis) is calculated as $\sum_{i=1}^{n} \frac{pi}{n}$ the n consonants (or vowels), Pi being the probability of the i^{th} consonant (or vowel). See text for explanation.

Table 2: The probability of repetition of vowels in Croatian as a function of degree of separation for words from _5_ to _12_ phonemes in length.

Table 3: The probability of repetition of consonants in Croatian as a function of degree of separation for words from _5_ to _12_ phonemes long.

Table 4: The probability of repetition of vowels in Hawaiian as a function of gap length for words from _3_ to _11_ phonemes long.

Table 5: The probability of repetition of consonants in Hawaiian as a function of degree of separation for words from _3_ to _11_ phonemes long.

Legends

Figure 1: The probability of phoneme repetition in 258 Croatian words, 10 phonemes long. Vowels are represented by circles and consonants by dots. The gap length or degree of separation is shown on the abscissa, the repeated L's in ELDERLY being separation 3 for example. The probability of repetition was calculated for each degree of separation as:

$$PR = \frac{FR}{FP}$$

where FR is the actual frequency of repetition, and FP the maximum possible frequency of repetition. The theoretical probability of phoneme repetition under the Null Hypothesis is shown by the broken line for vowels, and the dotted line for consonants, calculated as

$$\sum_{i=1}^{n} \frac{Pi}{n}$$

for the n consonants (or vowels), Pi being the probability of the i^{th} consonant (or vowel). See text for explanation.

Figure 2: The probability of phoneme repetition in Croatian for a random selection of words regardless of length. The actual probability of repetition is represented by circles for vowels and dots for consonants. The Null hypotheses is shown with a broken line for vowels and a dotted line for consonants.

Figure 3: The probability of phoneme repetition in Hawaiian for a random selection of words, regardless of length, the vowels being represented by circles and the consonants by dots. The Null Hypothesis is shown with a broken line for vowels and a dotted line for consonants.

Legends (Cont.)

Figure 4: The probability of detecting repeated letter misspellings such as ELDERY by a group of M.I.T. undergraduates. The gap length or degree of separation is shown on the abscissa, the repeated letters in the above misspelling being degree of separation 3. The probability of perceiving a misspelling not involving repeated letters is shown by the broken line.

Figure 5: The solid line shows the probability of recall of repeated letter misspellings by a group of M.I.T. undergraduates. The degree of separation between the repeated letters is shown on the abscissa. The probability of recalling a misspelling not involving repeated letters is shown by the broken line.

Figure 6: The probability of correct spelling in the diary of L. H. Oswald. This probability was calculated for each degree of separation as:

$$Pc = \frac{Fd}{F}$$

where Fd is the frequency of repeated letter deletions such as ELDERY and F is the frequency of repeated letters in words of the diary, and Pc is the probability of correct spelling.

Table 1

	Degree of Separation	Actual Frequency of Repetition	Maximum Possible Frequency	Actual Repeat Rate	Theoretical Repeat Rate
	0	0	2580	.000	.042
	1	128	2064	.062	.042
	2	105	1806	.058	.042
a) Vowels	3	93	1548	.060	.042
	4	72	1290	.054	.042
	5	80	1130	.057	.042
	6	29	774	.037	.042
	7	18	516	.035	.042
	8	10	258	.039	.042
	0	0	2580	.000	.026
	1	21	2064	.010	.026
	2	43	1806	.024	.026
b) Consonants	3	54	1548	.035	.026
	4	53	1290	.041	.026
	5	39	1130	.034	.026
	6	13	774	.017	.026
	7	15	516	.030	.026
	8	7	258	.025	.026

Table 2

Degree of Separation

Word Length	Number of Words	0	1	2	3	4	5	6	7	8	9
5	576	.000	.056	.045							
6	275	.000	.071	.041	.050						
7	149	.000	.067	.040	.049	.030					
8	151	.000	.074	.032	.057	.043	.037				
9	206	.000	.056	.048	.040	.043	.030	.030			
10	155	.000	.067	.049	.048	.038	.038	.038	.043		
11	145	.000	.072	.043	.057	.048	.036	.044	.042	.043	
12	194	.001	.059	.049	.047	.049	.035	.035	.036	.042	.039
Average	231	.00	.066	.044	.050	.043	.035	.036	.040	.043	.039

Table 3

Degree of Separation

Word Length	Number of Words	0	1	2	3	4	5	6	7	8	9
5	576	.000	.012	.016							
6	275	.000	.020	.026	.032						
7	149	.000	.007	.018	.022	.030					
8	151	.000	.011	.015	.030	.029	.020				
9	206	.000	.016	.019	.039	.028	.022	.028			
10	155	.000	.009	.021	.027	.027	.023	.020	.018		
11	145	.000	.012	.022	.030	.032	.026	.025	.018	.022	
12	194	.000	.008	.022	.020	.027	.040	.015	.038	.023	.024
Average	231	.000	.012	.020	.029	.029	.026	.022	.025	.023	.024

Table 4

Word Length	Number of Words	Degree of Separation 0	1	2	3	4	5	6	7	8	9
3	51	.000									
4	253	.000	.150								
5	223	.000	.081	.080							
6	587	.000	.135	.034	.114						
7	282	.000	.158	.067	.107	.015					
8	355	.000	.175	.045	.113	.016	.099				
9	93	.000	.169	.052	.086	.054	.050	.102			
10	145	.000	.152	.026	.085	.012	.084	.011	.072		
11	22	.000	.136	.051	.155	.068	.055	.068	.121	.114	
12	17	.000	.141	.013	.106	.008	.098	.012	.088	.000	.058
Average	203	.000	.144	.046	.109	.046	.077	.048	.094	.057	.058

Table 5

Word Length	Number of Words	Degree of Separation 0	1	2	3	4	5	6	7	8	9
3	51	.000									
4	253	.000	.040								
5	223	.000	.028	.009							
6	587	.000	.054	.024	.024						
7	282	.000	.031	.007	.017	.023					
8	355	.000	.045	.018	.090	.002	.028				
9	93	.000	.034	.011	.049	.005	.014	.011			
10	145	.000	.028	.012	.195	.004	.029	.000	.031		
11	22	.000	.015	.006	.130	.015	.009	.011	.000	.000	
12	17	.000	.006	.013	.154	.000	.029	.000	.029	.000	.044
Average	203	.000	.031	.011	.094	.008	.022	.006	.020	.000	.094