

AD/A-003 518

LANGUAGE IDENTIFICATION BY STATISTICAL
ANALYSIS

Morton David Rau

Naval Postgraduate School
Monterey, California

September 1974

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD/A003 518

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Language Identification by Statistical Analysis		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis September 1974
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Morton David Rau		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE September 1974
		13. NUMBER OF PAGES 128
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Reproduced by NATIONAL TECHNICAL INFORMATION SERVICE U.S. Department of Commerce Springfield, VA. 22151		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Language Identification, Linguistic Analysis, Linguistics, Statistical Linguistics		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) An analysis was conducted of English and Spanish text. The statistical analysis determined the independent probability of letters and the joint probability of various letter combinations for large samples of each language. Various methods were tested in an attempt to utilize these characteristics to identify the language of a short sample text. By use of the joint		

DD FORM 1473
1 JAN 73
(Page 1)EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

probability of various vowel-consonant relationships and the Kolmogorov-Smirnov Goodness of Fit Test an identification system was defined that provided a significance level of .0077 for a sample of 107 letters (approximately 21 words). Investigation also showed that the space rate or the interword structure in each language contains a measure of intelligence and was useful in identification.

DD Form 1473 (BACK)
1 Jan 73
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Language Identification

by

Statistical Analysis

by

Morton David Rau
Lieutenant Commander, United States Navy
B.S., Tulane University, 1962

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
September 1974

Author

Morton D Rau

Approved by:

R. A. Reitzman

Co-Advisor

V. Michael Powers

Co-Advisor

D. S. Schaefer

Chairman, Department of Operations Research/
Administrative Sciences

J. R. Montgomery

Academic Dean

ABSTRACT

An analysis was conducted of English and Spanish text. The statistical analysis determined the independent probability of letters and the joint probability of various letter combinations for large samples of each language.

Various methods were tested in an attempt to utilize these characteristics to identify the language of a short sample text. By use of the joint probability of various vowel-consonant relationships and the Kolmogorov-Smirnov Goodness of Fit Test an identification system was defined that provided a significance level of .0077 for a sample of 107 letters (approximately 21 words). Investigation also showed that the space rate or the interword structure in each language contains a measure of intelligence and was useful in identification.

TABLE OF CONTENTS

I. INTRODUCTION	8
II. BACKGROUND	10
A. ANALYTICAL CONSIDERATIONS OF THE MODERN LATIN ALPHABET	10
B. YULE'S CHARACTERISTIC (K)	11
C. KOLMOGOROV-SMIRNOV TEST OF GOODNESS OF FIT	13
III. EXPERIMENTS	15
A. GENERAL ANALYSIS CONSIDERATIONS	15
B. GENERAL PROGRAMING CONSIDERATIONS	15
C. POPULATION PARAMETERS	16
D. TESTING PROCEDURES	18
1. Testing by Yule's K	19
2. Testing by Kolmogorov-Smirnov Goodness of Fit Test	20
E. RESULTS	21
IV. CONCLUSIONS	24
APPENDIX A. English Population Values (V-1)	26
APPENDIX B. Spanish Population Values (V-1)	32
APPENDIX C. English Population Values (V-2)	38
APPENDIX D. Spanish Population Values (V-2)	44
APPENDIX E. Significance Levels for Kolmogorov-Smirnov Test	50
APPENDIX F. Final Test Sources	51
APPENDIX G. Population Analysis Program (V-1)	57
APPENDIX H. Population Analysis Program (V-2)	71

APPENDIX I. Language Identification Program (V-1) -----	87
APPENDIX J. Language Identification Program (V-2) -----	106
LIST OF REFERENCES -----	126
INITIAL DISTRIBUTION LIST -----	128

LIST OF TABLES

Table	Title	Page
I.	Proportion of Correct Identifications -----	22
II.	English Joint Probability Table for Vowel-Consonant -----	26
III.	English (V-1) Independent Probability of X_n -----	27
IV.	English (V-1) Joint Probability Table for X_n, X_{n+1} -----	28
V.	English (V-1) Joint Probability Table for $X_n, -, X_{n+2}$ -----	30
VI.	Spanish Joint Probability Table for Vowel-Consonant -----	32
VII.	Spanish (V-1) Independent Probability of X_n -----	33
VIII.	Spanish (V-1) Joint Probability Table for X_n, X_{n+1} -----	34
IX.	Spanish (V-1) Joint Probability Table for $X_n, -, X_{n+2}$ -----	36
X.	English (V-2) Independent Probability of X_n -----	39
XI.	English (V-2) Joint Probability Table for X_n, X_{n+1} -----	40
XII.	English (V-2) Joint Probability Table for $X_n, -, X_{n+2}$ -----	42
XIII.	Spanish (V-2) Independent Probability of X_n -----	45
XIV.	Spanish (V-2) Joint Probability Table for X_n, X_{n+1} -----	46
XV.	Spanish (V-2) Joint Probability Table for $X_n, -, X_{n+2}$ -----	48
XVI.	Significance Levels for Kolmogorov-Smirnov Test -----	50

I. INTRODUCTION

In today's world of ever increasing written collections it requires a certain level of expertise to properly identify and file the material according to its language. In addition to libraries this requirement exists within large government agencies such as the State Department and in various intelligence collection agencies. Although it would be expedient for the staff of these repositories to recognize and understand every language used it does not seem to be a practical requirement. It would at least seem desirable that these persons be provided with a simple method for identifying the language even if they were unable to read the words. An approach to this problem would be to have the clerk utilize a computer terminal to enter a short sample of the text which would be analyzed by the computer and then display the identified language on the terminal. The question that arises is how would the computer determine what language the sample represented. This was the problem approached in this research.

There have been many works published on the subject of mathematical linguistics most of which seem to concern themselves with identifying the author of various literary works or style analysis of individuals or regions. One investigator, Yukio Nakamura, has developed a language identification method for some twenty-five languages based upon the probability of occurrence of various specific characters and words in each language [Ref. 1].

Since language is basically a code used to convey information it would seem that each should have a particular characteristic that would

allow identification based on a statistical pattern. It was the intent of this research to design a set of computer programs which would analyze character frequency and order of occurrence in languages which are written in the Modern Latin Alphabet and then to utilize this data to identify an unknown sample.

II. BACKGROUND

A. ANALYTICAL CONSIDERATIONS OF THE MODERN LATIN ALPHABET

The Modern Latin Alphabet consists of the basic twenty-six characters, additional special characters and characters combined with diacritical signs. Very few languages utilize only the basic twenty-six characters. English is one which relies upon the basic twenty-six, except for loan words and less fashionable use of diacritics, such as 'coördinate'.

Many languages utilize additional characters which can be divided into three classifications:

1. Single characters, e.g., ß in German.
2. Joint characters, e.g., œ in French.

3. Combined characters, e.g., ch in German or Spanish. In some languages these combinations are looked upon as individual letters and have their appropriate place in the alphabet. For the purposes of this analysis they have been considered to be two individual letters.

Diacritical marks are generally used to indicate a difference in pronunciation of a letter but in some languages the marked letter is considered to be a distinct letter, such as the ü in German. In basic linguistics they are classified as modified letters and accordingly for the purposes of this study were treated as forms of the basic letter.

Having defined the criteria for the alphabet to be analyzed the question which arose was to what extent should the interrelationship of letters be considered. It was decided that limiting the interrelationship considerations to three adjacent letters would provide sufficient

indication of trends in letter relationships without broaching too greatly into the influences of word structure. Given any three adjacent characters (X_n, X_{n+1}, X_{n+2}) it was decided to evaluate the frequencies of the following events:

1. Independent occurrence of each character, X_n .
2. Joint occurrence of each pair, X_n, X_{n+1} .
3. $X_n, -, X_{n+2}$: the joint occurrence of a given letter X_{n+2} as the second letter following a given X_n .
4. Vowel-Consonant relationship in the form of the occurrence of:
 - a. vowel following vowel;
 - b. consonant following vowel;
 - c. vowel following consonant;
 - d. consonant following consonant.

Such characteristics as the occurrence of X_n, X_{n+1} generate 676 (26×26) combinations for the English Language. This has to be compared to at least the same number from any given sample. An extensive review of works on statistical linguistic analysis and nonparametric statistics resulted in the selection of two tests as a possible means of comparing a sample with a language population. These were a characteristic 'K' devised by the British statistician G. Udny Yule and the Kolmogorov-Smirnov Test of Goodness of Fit, which are discussed below.

B. YULE'S CHARACTERISTIC (K)

G. Udny Yule was involved in conducting statistical analysis of accident distribution among various people when faced with the problem of measuring the liability of classes of people independent of the period of exposure to risk or the total number of accidents met by the whole

group at risk. As a solution to this problem he developed an equation which yielded what he termed the 'characteristic' which was nothing more than a measure of the whole group's liability. By comparing accident rate or liability to repetitiveness of words in a literary work he proposed the use of this characteristic as a measure of repetitiveness in literature. The characteristic is expressed simply as a numerical value, such as 10.6, or 97.5, or 874.3. Since the analysis of each language and sample was to measure frequency of occurrence it was felt that this measure could be said to be analogous to repetitiveness as defined by Yule.

In general, Yule's approach was to determine for a given literary work how many common nouns occurred (were used) once, how many occurred twice, thrice, etc. He then performed the following calculations to obtain the characteristic (K).

X = occurrence category, i.e., X times occurring

f_x = frequency of X , i.e., the number of words occurring X times

$$S_1 = \sum_{x=1}^n f_x X$$

$$S_2 = \sum_{x=1}^n f_x X^2$$

$$K = 10,000 \frac{S_2 - S_1}{(S_1)^2}$$

The number 10,000 was introduced simply to avoid the inconvenience of handling very small numbers [Ref. 2].

Adaptation of this measurement to the $m \times n$ matrix generated by determination of letter occurrences was as follows:

$$S_1 = \sum_{x=1}^n f_x X = f_1(1) + f_2(2) + \dots + f_n(n)$$

$$S_2 = \sum_{x=1}^n f_x X^2 = f_1(1)^2 + f_2(2)^2 + \dots + f_n(n)^2$$

Instead of X being used to relate the frequency of a word it was equated to the frequency of a specific letter combination.

$\therefore$ Let: $X = Y_{m,n}$ = frequency of occurrence of the mn letter combination

$$f_x = f_{y_{m,n}} = 1$$

$$\therefore S_1 = 1(Y_{1,1}) + 1(Y_{1,2}) + \dots + 1(Y_{m,n})$$

$$S_1 = \sum_{m=1}^m \sum_{n=1}^n Y_{m,n}$$

Similarly:

$$S_2 = \sum_{m=1}^m \sum_{n=1}^n (Y_{m,n})^2$$

This adaptation yields a single value or measure of repetitiveness in a specific category for each language population and sample.

C. KOLMOGOROV-SMIRNOV TEST OF GOODNESS OF FIT

Let X be a random variable with the continuous probability distribution function $f(x) = \text{Prob} [X \leq x]$ and let $X_1, X_2, \dots, X_n$ be a sample of independent variables of size n for X , ordered so that $X_1 \leq X_2 \leq \dots \leq X_n$ which determines the empirical distribution function

$$f_n(x) = \begin{cases} 0 & \text{for } x < X_1 \\ (k/n) & \text{for } X_k \leq x \leq X_{k+1} \quad k = 1, 2, \dots, n-1, \\ 1 & \text{for } X_n \leq x \end{cases}$$

It should be expected that for n large, $f_n(x)$ would very likely approach $f(x)$. Kolmogorov defined the statistic

$$D_n = \text{least upper bound of } |f(x) - f_n(x)|$$

which measures the greatest absolute difference between $f(x)$ and $f_n(x)$.

He further showed that it has the following properties which assists in determining how closely $f_n(x)$ represents the population $f(x)$:

1. the probability distribution of D_n depends on n but is independent of $f(x)$, (the maximum D_n is a random variable)

2. the probability distribution of D_n is given by the relationship

$$\lim_{n \rightarrow \infty} \text{Prob} \left[D_n < \frac{z}{\sqrt{n}} \right] = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 z^2} = L(z)$$

[Refs. 3, 4, 5, and 6].

The function $L(z)$ has been tabulated by N. Smirnov [Ref. 7]. From the equation $D_n \sqrt{n} = z$ a value of z is obtained to enter the tables and thereby determine $\text{Prob} \left[D_n < \frac{z}{\sqrt{n}} \right]$ or find the significance level of the fit as $1 - L(z)$. These values have been computed and are tabulated in Appendix E.

Prior to continuing some comment should be made as to the selection of the Kolmogorov-Smirnov Test over the Chi-Squared Test. The Chi-Squared test requires a large sample size, because n at finite values is not distribution-free and only approaches this condition as n approaches infinity [Refs. 4 and 8].

III. EXPERIMENTS

A. GENERAL ANALYSIS CONSIDERATIONS

In considering the alphabetic structure of a language there is no doubt as to the fact that the letters carry a certain amount of intelligence but it was not certain as to what effect if any was played by the space, other than to separate words. In order to determine the effect of the space as a character in identifying a language two analytical approaches were taken in evaluating the characteristics described in section II. A. The first (version one, Appendix G) was to conduct the analysis considering spaces as separators, not characters. All characters not described as alphabetic, such as the period, comma, etc., were considered to be spaces and signified only the termination of a word. Joint frequencies such as that of X_n, X_{n+1} were counted only when the pair occurred within the same word. This resulted in the measurement of the desired characteristics exclusively within words, without measuring the interword relationships. The second approach (version two, Appendix H) was to consider the space and nonalphabetic characters between words as a space character. This resulted in the measurement of such characteristics as frequency of first letter occurrence, frequency of last letter occurrences, and the interword relationship of last letter to first letter.

B. GENERAL PROGRAMING CONSIDERATIONS

All computer programs were written in American National Standard COBOL.

Input language samples were fed to the programs via standard eighty column punch cards in the following format:

<u>Card Column</u>	<u>Content</u>
1, 2	Blank
3 - 72	Text
73	Blank unless the last word on the previous card could not be completed in column 72, then a '-' was placed in column 73 to signify continuation on the next card.
74 - 80	Sequence Number

Only alphabetic text with normal punctuation was used for input text. Numeric characters appearing in the various text material were converted to their alphabetic equivalent. To allow for the compilation of data from languages which utilize special characters the keypunch characters 1, 2, 3, 4, and 5 were allocated to those classified as vowels while 6 and 7 were for consonants. This conversion, if required, had to be performed by the operator at the time of entry or during data preparation.

In those programs that were written to consider the space as a character of the alphabet, all interword spaces and punctuation marks were compressed to one space.

C. POPULATION PARAMETERS

The actual population in the case of written text is potentially infinite. Due to limited resources it was decided to test the theory using representative samples of two languages as a measure of their respective populations. The source of these population samples was

chosen to be Reader's Digest as it is available in many languages and contains several short articles, by different authors, which have been selected from various areas of interest, thereby suppressing the influence of one person's or region's style on the assumed population.

The English population consisted of text material, excluding advertisements, from Reader's Digest, Vol. 104, No. 623, March 1974, pages 8, 11, 12, 31 - 90, 109 and 110. This text consists of 79,185 characters excluding spaces and 96,100 characters including spaces.

The Spanish population consisted of text material, excluding advertisements, from Selecciones Del Reader's Digest, Vol. 67, No. 402, May 1974, pages 3 - 10, 21 - 30, 37-60, and 75 -80. This text consists of 78,851 characters excluding spaces and 94,992 characters including spaces.

Both populations were analyzed for the characteristics described in section II. A. (with and without space consideration) and the following population means were obtained:

1. Yule's 'K' based on the independent occurrence of each character, X_n ;
2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} ;
3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$;
4. Probability table for the independent occurrence of X_n ;
5. Joint probability table for the occurrence of X_n, X_{n+1} ;
6. Joint probability table for the occurrence of $X_n, -, X_{n+2}$;
7. Joint probability table for vowel-consonant occurrence.

(A tabulation of these values is contained in Appendices A, B, C, and D)

Each probability table was formed into a cumulative relative frequency

curve, the points of which were transferred to punched cards, for later use in testing samples by the Kolmogorov-Smirnov Goodness of Fit Test.

In order to facilitate the use of Yule's 'K' in testing small samples it was necessary to approximate the standard deviation (σ) for each value of Yule's 'K'. The population was divided into groups of ten card samples from which a new 'K' was determined for each sample. The following calculations were then performed:

K = Yule's 'K' obtained from each sample of 10 cards (less than 700 characters).

n = number of 10 card samples contained in the population.

$\bar{K} = \frac{\sum K}{n}$ = mean value of Yule's 'K'.

$$\therefore \sigma_{10} = \sqrt{\frac{(\sum K^2 - n\bar{K})}{(n - 1)}}$$

σ_{10} was then adjusted to a population standard deviation based on one card samples by computing

$$\sigma = \sigma_{10} \sqrt{10}$$

[Ref. 8].

(Values of σ are contained in Appendices A, B, C, and D.) A histogram was plotted for the various values of Yule's 'K' and they appeared to conform to a normal distribution.

D. TESTING PROCEDURES

In constructing testing procedures there were two approaches which were considered. The first was that a small sample would be analyzed and identified as either English or Spanish. The second approach was to analyze a small sample and then report it as either English, Spanish, or undecisive based upon some predefined constraints. Since it was

intended, if the theory worked, that the process could be expanded to cover all languages written in the Modern Latin Alphabet it was felt that the first approach was more realistic (useful).

1. Testing by Yule's K

Since the incremental population samples of ten cards appeared to conform to the normal distribution it was assumed that for all languages the value of K obtained from small samples would conform to the normal distribution. For each sample tested three values of K were obtained, one for each of the three K tests. Each value of K was then compared to its respective population mean, μ , obtained from the representative population analysis, in each language in order to determine the appropriate normal deviate, z, by

$$z = \frac{K - \mu}{\sigma / \sqrt{n}}$$

where n was the number of data cards in that particular sample. For each of the three tests the z for English was compared with the z for that test in Spanish. The smallest absolute value of z identified the language or that language to which the sample had the highest probability of belonging. The smallest z from each test was then used to compute the significance level, α , as a measure of how closely K and μ matched based on the probability of a normal distribution. For absolute values of z less than 4.0, α was computed as follows:

$$\begin{aligned} x &= |z| \\ \alpha &= 1 - 2x(.39894,22804,014) \left[1 - \frac{1}{6}x^2 + \frac{P}{6.66666,66666,67} x^2 \right. \\ &\quad - \frac{P}{8.4} + \frac{P}{10.28571,42857,1} x^2 - \frac{P}{12.22222,22222,2} x^2 \\ &\quad \left. + \frac{P}{14.18181,81818,2} x^2 - \frac{P}{16.15384,61538,5} x^2 \right] \end{aligned}$$

$$\begin{aligned}
& + \frac{P}{18.13333,33333,3} x^2 - \frac{P}{20.11764,70588,2} x^2 \\
& + \frac{P}{22.10526,31578,9} x^2 - \frac{P}{24.09523,80952,4} x^2 \\
& + \frac{P}{26.08695,65217,4} x^2 - \frac{P}{28.08000,0000,0} x^2 \\
& + \frac{P}{30.07407,40740,7} x^2 - \frac{P}{32.06896,55172,4} x^2 \\
& + \frac{P}{34.06451,61290,3} x^2 - \frac{P}{36.06060,60606,1} x^2 \\
& + \frac{P}{38.05714,28571,4} x^2 - \frac{P}{40.05405,40540,5} x^2 \\
& + \frac{P}{42.05128,20512,8} x^2 - \frac{P}{44.04878,04878,0} x^2 \\
& + \frac{P}{46.04651,16279,1} x^2 - \frac{P}{48.04444,44444,4} x^2 \\
& + \frac{P}{50.04255,3191} x^2 - \frac{P}{52.04081,6326} x^2 \\
& + \frac{P}{54.03921,5686} x^2 - \frac{P}{56.03773,5849} x^2 \\
& + \left[\frac{P}{58.03636} x^2 - \frac{P}{60.03509} x^2 + \frac{P}{62.03} x^2 \right]
\end{aligned}$$

where P = the absolute value of the preceding term [Ref. 9]. For absolute values of z greater than or equal to 4.0, α was reported as being less than 0.00006 [Ref. 8].

2. Testing by the Kolmogorov-Smirnov Goodness of Fit Test

For each sample tested four cumulative relative frequency tables were computed, one from each of the probability tables described in section III. C. These tables represented points, $f_n(x)$, on the cumulative relative frequency (CRF) curve for each particular sample. The CRF curve points for the sample were compared against the estimate of the respective curve points, $f(x)$, along the ordinate, for the Spanish and English populations. From each comparison the value of

D_n = least upper bound of $|f(x) - f_n(x)|$ was obtained. For each of the four tests the value of D_n from the comparison with Spanish was compared to that D_n from the comparison with English. The sample was judged to be of that language which resulted in the smallest D_n . The smallest value of D_n from each test was then used to compute the significance level as a measure of how closely the sample and population CRF curves matched. D_n was then multiplied by the square root of the number of data cards contained in the sample to obtain z . For values of z greater than or equal to 0.28 and less than or equal to 1.82, the significance level was obtained from the table contained in Appendix E. For values of z less than 0.28, α was reported as greater than 0.999999. For values of z greater than 1.82, α was reported as less than 0.002645.

E. RESULTS

In an attempt to construct an unbiased set of samples for final testing, an assortment of short stories, novels, poems, reports, and speeches by one hundred and thirty different authors was selected from several anthologies (Appendix F). The samples within this group were equally split between English and Spanish and covered a period from the twelfth century through the present day. Each sample was taken from the beginning of the particular work and was of such length to fill the prescribed text field of ten punch cards.

Each sample was analyzed four times: once utilizing all ten data cards (534 characters); once utilizing the first five data cards (267 characters); once utilizing the first two data cards (107 characters); and once utilizing only the first data card (53 characters).

Furthermore, each of these were processed once under Language-ID-Version-One (Appendix I) which used the space only as a separator and once under Language-ID-Version-Two (Appendix J) which used the space as a character.

In addition to the seven separate tests described in section III. C., three other tests which were evaluated were:

1. a majority vote of all seven tests;
2. a majority vote of the three tests based on Yule's 'K';
3. a majority vote of the four tests based on the Kolmogorov-Smirnov (K-S) test of goodness of fit.

The following tabulation presents the proportion of correct identifications achieved by each of the ten tests under Language-ID-Version-One (V-1) and Language-ID-Version-Two (V-2).

TABLE I
Proportion of Correct Identifications

No. of Cards	K for X_n		K for X_n, X_{n+1}	
	V-1	V-2	V-1	V-2
10	.9538	.6615	.6000	.6231
5	.9000	.6462	.5769	.6384
2	.7615	.5615	.5846	.6385
1	.7077	.5308	.5154	.5923

	K for $X_n, -, X_{n+2}$		K-S for Vowel-Cons.	
	V-1	V-2	V-1	V-2
10	.5308	.5846	1.0000	SAME AS V-1
5	.5077	.5923	1.0000	
2	.4615	.5461	.9923	
1	.5000	.5231	.8770	

	K-S for X_n		K-S for X_n, X_{n+1}	
	V-1	V-2	V-1	V-2
10	1.0000	1.0000	.9923	1.0000
5	.9769	.9769	.9230	.9846
2	.9348	.9348	.8539	.9615
1	.8770	.8693	.7230	.8539

No. of Cards	K-S for $X_n, \dots, X_{n+2}$		Majority Vote	
	V-1	V-2	V-1	V-2
10	.9307	1.0000	.9846	1.0000
5	.8770	.9769	.9615	1.0000
2	.8154	.9308	.9230	.9615
1	.7521	.8153	.8923	.8923

	Majority of K		Majority of K-S	
	V-1	V-2	V-1	V-2
10	.6385	.6154	.9923	1.0000
5	.6615	.6118	.9461	.9769
2	.6308	.5770	.8923	.9461
1	.5846	.5384	.7846	.8385

IV. CONCLUSIONS

The following observations have been made based upon the results presented in section III. E.

1. Yule's characteristic 'K' by itself was not a reliable measure for identifying the language based on a small sample. Two factors could have affected this measure: first was the close proximity of the mean values when viewed in light of the standard deviation; secondly was that with only twenty-six letters being considered there was too great a chance of repetitiveness which in fact was what Yule was measuring.

2. It was observed that a higher proportion of correct identifications occurred when the sample was analyzed using the space as a character. Accordingly it was concluded that the space rate or the interword structure in each language does contain a measure of intelligence and was useful in identification.

3. The use of the Kolmogorov-Smirnov Goodness of Fit Test when applied to vowel-consonant relationships proved to be the most reliable single test with a significance level of 0.0077 for a sample of approximately 107 characters.

4. No majority vote test was as reliable as the vowel-consonant test but the test based on a majority of the seven individual tests did prove to be the second most reliable test. In view of the fact that this test incorporated the results of three low-reliability Yule's 'K' tests it is subject to some doubt as to how reliable it would prove if the program were expanded to include more than two languages.

Although the experiments using only two languages could not produce conclusive evidence, this research tends to support the belief that each language does possess some statistical characteristic which would allow identification using a short sample of text. Expansion of this research to include additional languages would be useful in determining the reliability of this type of identification system when used with short samples of text.

APPENDIX A. ENGLISH POPULATION VALUES (V-1)

1. Yule's 'K' based on the independent occurrence of each character

X_n :

Mean value 645.079

Standard Deviation 81.479

2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} :

Mean value 87.168

Standard Deviation 40.207

3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$:

Mean value 62.083

Standard Deviation 42.821

TABLE II. ENGLISH JOINT PROBABILITY TABLE FOR VOWEL-CONSONANT

		First Letter	
		Vowel	Consonant
Second Letter	Vowel	.0512	.3568
	Consonant	.3549	.2371

TABLE III. ENGLISH (V-1) INDEPENDENT PROBABILITY OF X_n

X_n	Probability
A	.0800
E	.1214
I	.0710
O	.0774
U	.0284
B	.0138
C	.0313
D	.0418
F	.0222
G	.0217
H	.0502
J	.0017
K	.0072
L	.0445
M	.0255
N	.0731
P	.0210
Q	.0008
R	.0631
S	.0640
T	.0880
V	.0112
W	.0188
X	.0020
Y	.0191
Z	.0008

TABLE IV. ENGLISH (V-1) JOINT PROBABILITY TABLE FOR X_n, X_{n+1}

X_{n+1}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0000	.0072	.0016	.0012	.0010	.0022	.0065	.0022	.0015	.0019	.0096	.0001	.0001
E	.0000	.0050	.0030	.0004	.0013	.0046	.0046	.0058	.0019	.0040	.0271	.0005	.0029
I	.0033	.0020	.0000	.0010	.0010	.0012	.0022	.0038	.0040	.0014	.0086	.0000	.0012
O	.0000	.0007	.0062	.0025	.0001	.0021	.0085	.0038	.0058	.0016	.0065	.0006	.0001
U	.0011	.0002	.0000	.0125	.0000	.0023	.0010	.0013	.0010	.0007	.0019	.0009	.0000
B	.0019	.0002	.0011	.0008	.0006	.0001	.0000	.0001	.0000	.0000	.0001	.0000	.0001
C	.0031	.0040	.0066	.0017	.0017	.0000	.0003	.0001	.0000	.0000	.0000	.0000	.0000
D	.0039	.0133	.0035	.0017	.0007	.0000	.0000	.0005	.0000	.0000	.0000	.0000	.0000
F	.0012	.0012	.0019	.0075	.0002	.0000	.0000	.0000	.0017	.0000	.0000	.0000	.0000
G	.0019	.0009	.0038	.0007	.0016	.0000	.0000	.0000	.0000	.0003	.0000	.0000	.0000
H	.0001	.0002	.0000	.0004	.0000	.0000	.0054	.0001	.0000	.0035	.0000	.0000	.0000
J	.0001	.0000	.0000	.0001	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
K	.0012	.0003	.0004	.0010	.0000	.0000	.0026	.0000	.0000	.0000	.0000	.0000	.0000
L	.0106	.0054	.0051	.0052	.0032	.0020	.0013	.0005	.0007	.0006	.0001	.0000	.0000
M	.0031	.0032	.0037	.0052	.0012	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000
N	.0201	.0150	.0238	.0170	.0050	.0000	.0000	.0002	.0000	.0008	.0002	.0000	.0006
P	.0013	.0016	.0011	.0024	.0017	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Q	.0000	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
R	.0120	.0188	.0033	.0126	.0057	.0007	.0022	.0021	.0019	.0020	.0011	.0000	.0000
S	.0086	.0124	.0085	.0026	.0052	.0002	.0002	.0014	.0001	.0007	.0002	.0000	.0006
T	.0135	.0045	.0093	.0039	.0047	.0001	.0030	.0001	.0018	.0002	.0024	.0000	.0000
V	.0022	.0032	.0035	.0019	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000
W	.0009	.0014	.0000	.0038	.0000	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0000
X	.0002	.0014	.0006	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0029	.0014	.0000	.0002	.0001	.0015	.0004	.0007	.0001	.0005	.0003	.0000	.0001
Z	.0002	.0001	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

TABLE IV - Continued

x_{n+1}	x_n													
	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	
A	.0053	.0057	.0034	.0030	.0000	.0062	.0035	.0039	.0007	.0047	.0001	.0001	.0002	
E	.0095	.0081	.0090	.0048	.0000	.0174	.0080	.0128	.0102	.0047	.0001	.0012	.0005	
I	.0070	.0042	.0040	.0012	.0000	.0067	.0048	.0109	.0026	.0034	.0002	.0003	.0001	
O	.0035	.0037	.0037	.0035	.0000	.0077	.0033	.0110	.0007	.0031	.0000	.0022	.0001	
U	.0009	.0010	.0006	.0009	.0010	.0011	.0027	.0019	.0000	.0000	.0000	.0000	.0000	
B	.0000	.0008	.0000	.0000	.0000	.0002	.0002	.0001	.0000	.0000	.0000	.0000	.0000	
C	.0000	.0001	.0030	.0000	.0000	.0013	.0013	.0003	.0000	.0000	.0001	.0001	.0000	
D	.0031	.0000	.0151	.0000	.0000	.0021	.0001	.0000	.0000	.0000	.0000	.0000	.0000	
F	.0006	.0000	.0005	.0000	.0000	.0002	.0001	.0001	.0000	.0000	.0000	.0000	.0000	
G	.0000	.0000	.0114	.0000	.0000	.0019	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
H	.0000	.0000	.0002	.0007	.0000	.0001	.0032	.0300	.0000	.0035	.0001	.0000	.0000	
J	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
K	.0002	.0000	.0008	.0000	.0000	.0011	.0005	.0000	.0000	.0000	.0000	.0000	.0000	
L	.0089	.0001	.0004	.0023	.0000	.0008	.0007	.0011	.0000	.0001	.0000	.0006	.0000	
M	.0002	.0007	.0004	.0001	.0000	.0012	.0011	.0002	.0000	.0000	.0000	.0001	.0000	
N	.0000	.0001	.0008	.0000	.0000	.0018	.0002	.0000	.0000	.0008	.0000	.0001	.0000	
P	.0004	.0025	.0001	.0016	.0000	.0004	.0016	.0001	.0000	.0000	.0005	.0000	.0000	
Q	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	
R	.0002	.0001	.0001	.0046	.0000	.0011	.0000	.0040	.0000	.0002	.0000	.0003	.0000	
S	.0016	.0008	.0039	.0007	.0000	.0060	.0036	.0029	.0000	.0005	.0000	.0006	.0000	
T	.0008	.0000	.0100	.0007	.0000	.0036	.0102	.0017	.0000	.0000	.0007	.0001	.0000	
V	.0003	.0000	.0009	.0000	.0000	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
W	.0001	.0000	.0001	.0000	.0000	.0001	.0004	.0018	.0000	.0000	.0000	.0000	.0000	
X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
Y	.0042	.0008	.0010	.0003	.0000	.0021	.0003	.0042	.0001	.0000	.0001	.0000	.0000	
Z	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	

TABLE V. ENGLISH (V-1) JOINT PROBABILITY TABLE FOR X_n , X_{n+1} , X_{n+2}

X_{n+2}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0021	.0039	.0057	.0034	.0032	.0008	.0045	.0012	.0008	.0011	.0011	.0000	.0001
E	.0175	.0150	.0168	.0151	.0045	.0025	.0042	.0034	.0030	.0023	.0015	.0001	.0005
I	.0093	.0070	.0068	.0046	.0031	.0015	.0035	.0011	.0016	.0011	.0014	.0001	.0001
O	.0031	.0034	.0014	.0012	.0009	.0007	.0020	.0009	.0025	.0013	.0014	.0001	.0007
U	.0011	.0018	.0025	.0012	.0004	.0012	.0032	.0003	.0011	.0001	.0031	.0000	.0000
B	.0003	.0007	.0002	.0004	.0006	.0002	.0004	.0001	.0001	.0002	.0002	.0002	.0000
C	.0023	.0028	.0014	.0011	.0010	.0014	.0000	.0021	.0015	.0002	.0012	.0002	.0000
D	.0168	.0037	.0025	.0038	.0058	.0004	.0007	.0017	.0006	.0005	.0021	.0001	.0011
F	.0006	.0013	.0006	.0015	.0005	.0004	.0002	.0003	.0006	.0001	.0001	.0000	.0000
G	.0023	.0027	.0125	.0037	.0009	.0007	.0005	.0002	.0001	.0000	.0008	.0000	.0000
H	.0021	.0012	.0070	.0019	.0026	.0000	.0030	.0001	.0000	.0001	.0000	.0002	.0000
J	.0000	.0000	.0000	.0001	.0000	.0000	.0090	.0000	.0000	.0000	.0000	.0000	.0000
K	.0024	.0009	.0018	.0028	.0003	.0000	.0030	.0000	.0000	.0000	.0000	.0000	.0000
L	.0067	.0031	.0046	.0086	.0027	.0013	.0038	.0025	.0013	.0005	.0034	.0000	.0000
M	.0009	.0011	.0009	.0021	.0001	.0000	.0034	.0008	.0004	.0002	.0030	.0001	.0001
N	.0026	.0041	.0098	.0051	.0012	.0015	.0035	.0032	.0011	.0018	.0067	.0002	.0015
P	.0017	.0031	.0015	.0023	.0011	.0000	.0037	.0005	.0000	.0001	.0012	.0000	.0001
Q	.0001	.0000	.0000	.0000	.0000	.0000	.0030	.0001	.0000	.0000	.0000	.0000	.0000
R	.0021	.0055	.0005	.0053	.0008	.0016	.0024	.0023	.0078	.0020	.0062	.0004	.0004
S	.0068	.0094	.0049	.0088	.0019	.0008	.0027	.0017	.0008	.0019	.0074	.0009	.0003
T	.0084	.0156	.0070	.0087	.0049	.0031	.0013	.0012	.0004	.0045	.0051	.0000	.0003
V	.0004	.0015	.0010	.0006	.0001	.0002	.0002	.0004	.0016	.0008	.0018	.0000	.0000
W	.0001	.0003	.0002	.0001	.0001	.0001	.0001	.0006	.0003	.0000	.0012	.0001	.0001
X	.0000	.0000	.0000	.0000	.0000	.0002	.0000	.0000	.0001	.0000	.0000	.0000	.0000
Y	.0027	.0018	.0020	.0007	.0002	.0007	.0001	.0015	.0005	.0002	.0013	.0000	.0002
Z	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0001	.0000	.0000

TABLE V - Continued

X_{n+2}	X_n														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	P
A	.0042	.0016	.0023	.0017	.0005	.0047	.0053	.0069	.0001	.0011	.0001	.0014	.0001	.0000	.0000
B	.0038	.0019	.0080	.0045	.0004	.0087	.0061	.0317	.0003	.0019	.0006	.0010	.0001	.0000	.0000
C	.0023	.0012	.0051	.0028	.0003	.0042	.0051	.0054	.0001	.0018	.0002	.0002	.0000	.0000	.0000
D	.0023	.0006	.0028	.0053	.0000	.0022	.0045	.0104	.0001	.0013	.0001	.0001	.0000	.0000	.0000
E	.0007	.0005	.0011	.0004	.0000	.0022	.0024	.0011	.0000	.0011	.0000	.0023	.0000	.0000	.0000
F	.0004	.0001	.0007	.0005	.0000	.0015	.0005	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000
G	.0020	.0011	.0007	.0011	.0000	.0045	.0015	.0025	.0006	.0001	.0001	.0000	.0000	.0000	.0000
H	.0021	.0018	.0013	.0009	.0000	.0049	.0021	.0035	.0010	.0004	.0001	.0003	.0000	.0000	.0000
I	.0008	.0001	.0001	.0000	.0000	.0011	.0005	.0002	.0000	.0002	.0000	.0000	.0000	.0000	.0000
J	.0014	.0007	.0005	.0001	.0000	.0017	.0004	.0006	.0000	.0001	.0000	.0000	.0000	.0000	.0000
K	.0004	.0002	.0010	.0000	.0000	.0018	.0006	.0004	.0000	.0000	.0000	.0003	.0000	.0000	.0000
L	.0000	.0002	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
M	.0005	.0011	.0000	.0000	.0000	.0001	.0000	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000
N	.0000	.0028	.0030	.0022	.0000	.0030	.0029	.0034	.0013	.0020	.0001	.0001	.0000	.0000	.0000
O	.0011	.0002	.0009	.0000	.0000	.0041	.0025	.0026	.0002	.0002	.0001	.0000	.0000	.0000	.0000
P	.0037	.0070	.0032	.0022	.0000	.0040	.0041	.0052	.0028	.0023	.0001	.0008	.0000	.0000	.0000
Q	.0005	.0000	.0002	.0004	.0000	.0020	.0008	.0007	.0000	.0001	.0000	.0000	.0000	.0000	.0000
R	.0000	.0000	.0000	.0000	.0000	.0002	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
S	.0039	.0043	.0057	.0055	.0000	.0005	.0040	.0094	.0047	.0037	.0002	.0004	.0000	.0000	.0000
T	.0042	.0035	.0042	.0016	.0000	.0059	.0015	.0018	.0015	.0031	.0002	.0002	.0000	.0000	.0000
U	.0038	.0028	.0063	.0013	.0000	.0036	.0011	.0016	.0007	.0028	.0000	.0003	.0001	.0000	.0000
V	.0008	.0004	.0007	.0000	.0000	.0017	.0014	.0009	.0001	.0004	.0000	.0000	.0000	.0000	.0000
W	.0013	.0000	.0023	.0002	.0000	.0011	.0002	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0001	.0003	.0000	.0000	.0000	.0008	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0018	.0006	.0022	.0009	.0000	.0025	.0010	.0017	.0000	.0011	.0004	.0000	.0000	.0000	.0000
Z	.0002	.0001	.0004	.0000	.0000	.0001	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000

APPENDIX B. SPANISH POPULATION VALUES (V-1)

1. Yule's 'K' based on the independent occurrence of each character

X_n :

Mean value 751.447

Standard Deviation 82.328

2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} :

Mean value 104.553

Standard Deviation 38.094

3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$:

Mean value 72.680

Standard Deviation 35.210

TABLE VI. SPANISH JOINT PROBABILITY TABLE FOR VOWEL-CONSONANT

		First Letter	
		Vowel	Consonant
Second Letter	Vowel	.0616	.4469
	Consonant	.3713	.1202

TABLE VII. SPANISH (V-1) INDEPENDENT PROBABILITY OF X_n

X_n	Probability
A	.1269
E	.1311
I	.0690
O	.0929
U	.0409
B	.0137
C	.0448
D	.0470
F	.0072
G	.0111
H	.0087
J	.0041
K	.0006
L	.0568
M	.0285
N	.0747
P	.0247
Q	.0087
R	.0619
S	.0752
T	.0464
V	.0114
W	.0000
X	.0014
Y	.0089
Z	.0037

TABLE VIII. SPANISH (V-1) JOINT PROBABILITY TABLE FOR X_n, X_{n+1}

X_{n+1}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0000	.0017	.0115	.0001	.0032	.0055	.0106	.0076	.0009	.0027	.0050	.0012	.0001
E	.0003	.0003	.0108	.0000	.0146	.0013	.0049	.0278	.0015	.0015	.0012	.0013	.0001
I	.0009	.0009	.0000	.0004	.0029	.0033	.0132	.0066	.0026	.0016	.0013	.0001	.0002
O	.0003	.0011	.0104	.0003	.0002	.0013	.0133	.0119	.0015	.0024	.0028	.0019	.0001
U	.0012	.0006	.0002	.0000	.0000	.0008	.0047	.0017	.0014	.0028	.0006	.0006	.0000
B	.0066	.0006	.0017	.0020	.0008	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
C	.0081	.0062	.0070	.0033	.0022	.0000	.0011	.0000	.0000	.0000	.0000	.0000	.0000
D	.0101	.0037	.0053	.0026	.0015	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
F	.0005	.0013	.0008	.0007	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
G	.0008	.0029	.0021	.0011	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
H	.0003	.0000	.0000	.0000	.0000	.0000	.0035	.0000	.0000	.0000	.0000	.0000	.0000
J	.0009	.0014	.0004	.0004	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
K	.0000	.0001	.0000	.0001	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000
L	.0116	.0139	.0048	.0037	.0020	.0018	.0009	.0000	.0004	.0003	.0000	.0000	.0001
M	.0054	.0035	.0041	.0053	.0012	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
N	.0176	.0302	.0095	.0164	.0111	.0000	.0001	.0000	.0000	.0002	.0001	.0000	.0000
P	.0015	.0011	.0005	.0014	.0008	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Q	.0006	.0003	.0002	.0001	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
R	.0173	.0162	.0028	.0102	.0031	.0026	.0009	.0011	.0008	.0023	.0000	.0000	.0000
S	.0157	.0234	.0049	.0230	.0032	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0001
T	.0038	.0037	.0035	.0021	.0011	.0001	.0029	.0000	.0000	.0000	.0000	.0000	.0000
V	.0009	.0015	.0024	.0023	.0003	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0001	.0014	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0009	.0002	.0000	.0004	.0006	.0000	.0000	.0000	.0030	.0000	.0000	.0000	.0000
Z	.0009	.0014	.0008	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

TABLE VIII. Continued

X_{n+1}	X_n													
	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	
A	.0202	.0076	.0109	.0079	.0000	.0156	.0051	.0140	.0029	.0000	.0002	.0009	.0022	
E	.0076	.0086	.0047	.0051	.0000	.0129	.0110	.0146	.0048	.0000	.0000	.0004	.0000	
I	.0059	.0063	.0045	.0020	.0000	.0090	.0068	.0081	.0049	.0000	.0005	.0000	.0000	
O	.0114	.0053	.0094	.0063	.0000	.0109	.0063	.0113	.0015	.0000	.0001	.0009	.0011	
U	.0011	.0023	.0023	.0021	.0110	.0017	.0046	.0024	.0001	.0000	.0000	.0001	.0001	
B	.0001	.0019	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
C	.0002	.0000	.0048	.0000	.0000	.0011	.0018	.0000	.0000	.0000	.0001	.0000	.0001	
D	.0003	.0000	.0047	.0000	.0000	.0014	.0003	.0000	.0000	.0000	.0000	.0000	.0000	
F	.0002	.0000	.0010	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	
G	.0007	.0000	.0012	.0000	.0000	.0014	.0001	.0000	.0000	.0000	.0000	.0000	.0000	
H	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	
J	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
K	.0000	.0000	.0001	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
L	.0043	.0000	.0001	.0022	.0000	.0009	.0002	.0001	.0000	.0000	.0000	.0000	.0000	
M	.0006	.0000	.0001	.0000	.0000	.0027	.0009	.0000	.0000	.0000	.0000	.0000	.0000	
N	.0000	.0000	.0005	.0000	.0000	.0006	.0000	.0000	.0000	.0000	.0000	.0001	.0000	
P	.0002	.0037	.0000	.0000	.0000	.0003	.0025	.0000	.0000	.0000	.0004	.0000	.0000	
Q	.0001	.0000	.0006	.0000	.0000	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	
R	.0001	.0000	.0001	.0049	.0000	.0021	.0000	.0076	.0000	.0000	.0000	.0000	.0000	
S	.0004	.0000	.0026	.0001	.0000	.0020	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
T	.0012	.0000	.0176	.0002	.0000	.0030	.0098	.0001	.0000	.0000	.0003	.0000	.0000	
V	.0006	.0000	.0009	.0000	.0000	.0007	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
Y	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
Z	.0001	.0000	.0008	.0000	.0000	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000	

TABLE IX. SPANISH (V-1) JOINT PROBABILITY TABLE FOR X_n , -, X_{n+2}

X_{n+2}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0196	.0140	.0169	.0095	.0106	.0031	.0094	.0024	.0007	.0023	.0000	.0001	.0000
E	.0128	.0113	.0056	.0084	.0030	.0047	.0052	.0022	.0020	.0020	.0003	.0001	.0000
I	.0171	.0134	.0097	.0091	.0036	.0016	.0043	.0005	.0003	.0015	.0000	.0000	.0000
O	.0158	.0108	.0103	.0084	.0053	.0009	.0086	.0015	.0003	.0016	.0000	.0000	.0001
U	.0034	.0034	.0026	.0016	.0005	.0002	.0015	.0001	.0001	.0003	.0000	.0000	.0000
B	.0011	.0005	.0003	.0012	.0003	.0003	.0010	.0007	.0002	.0001	.0025	.0002	.0000
C	.0024	.0054	.0020	.0012	.0020	.0004	.0008	.0028	.0023	.0004	.0019	.0002	.0000
D	.0035	.0027	.0016	.0021	.0022	.0003	.0025	.0035	.0000	.0005	.0001	.0002	.0000
F	.0002	.0006	.0006	.0004	.0001	.0001	.0001	.0007	.0000	.0000	.0000	.0002	.0000
G	.0021	.0005	.0007	.0011	.0009	.0000	.0002	.0001	.0002	.0000	.0001	.0001	.0000
H	.0008	.0008	.0001	.0008	.0011	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
J	.0000	.0001	.0003	.0003	.0001	.0005	.0000	.0005	.0000	.0001	.0002	.0000	.0000
K	.0002	.0001	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
L	.0041	.0025	.0045	.0008	.0023	.0008	.0037	.0042	.0005	.0006	.0006	.0000	.0002
M	.0021	.0018	.0028	.0010	.0003	.0003	.0057	.0012	.0002	.0002	.0014	.0005	.0000
N	.0014	.0020	.0160	.0003	.0031	.0017	.0112	.0035	.0017	.0027	.0011	.0007	.0000
P	.0004	.0051	.0014	.0019	.0005	.0000	.0013	.0002	.0000	.0000	.0001	.0000	.0000
Q	.0007	.0001	.0002	.0003	.0003	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000
R	.0041	.0040	.0037	.0032	.0031	.0021	.0055	.0043	.0018	.0022	.0013	.0008	.0000
S	.0021	.0028	.0047	.0021	.0028	.0008	.0050	.0086	.0006	.0009	.0016	.0010	.0000
T	.0100	.0248	.0055	.0041	.0036	.0009	.0007	.0012	.0005	.0005	.0001	.0003	.0000
V	.0005	.0013	.0007	.0006	.0008	.0000	.0002	.0008	.0000	.0000	.0001	.0007	.0000
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0000	.0001	.0000	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0005	.0000	.0001
Z	.0009	.0006	.0009	.0000	.0001	.0001	.0002	.0000	.0000	.0001	.0003	.0000	.0000

TABLE IX - Continued

X_{n+2}	L	M	N	P	Q	R	S	T	V	W	X	Y	Z
A	.0046	.0015	.0090	.0021	.0000	.0093	.0078	.0052	.0007	.0000	.0000	.0000	.0001
E	.0029	.0024	.0149	.0057	.0121	.0076	.0048	.0036	.0016	.0000	.0002	.0000	.0000
I	.0016	.0015	.0073	.0023	.0024	.0047	.0026	.0010	.0003	.0000	.0001	.0000	.0000
O	.0029	.0015	.0113	.0032	.0000	.0077	.0050	.0034	.0007	.0000	.0001	.0001	.0003
U	.0009	.0001	.0029	.0003	.0000	.0010	.0024	.0007	.0000	.0000	.0000	.0000	.0000
B	.0017	.0002	.0003	.0004	.0000	.0023	.0014	.0016	.0002	.0000	.0000	.0000	.0002
C	.0032	.0016	.0036	.0024	.0000	.0047	.0010	.0042	.0013	.0000	.0001	.0000	.0001
D	.0016	.0023	.0024	.0017	.0000	.0025	.0013	.0049	.0013	.0000	.0000	.0001	.0003
F	.0005	.0000	.0002	.0000	.0000	.0008	.0002	.0006	.0000	.0000	.0000	.0000	.0000
G	.0019	.0005	.0003	.0001	.0000	.0019	.0015	.0010	.0002	.0000	.0001	.0000	.0000
H	.0000	.0000	.0003	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000
J	.0003	.0006	.0001	.0001	.0000	.0004	.0004	.0001	.0001	.0000	.0000	.0000	.0000
K	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
L	.0002	.0044	.0020	.0018	.0000	.0013	.0035	.0030	.0011	.0000	.0005	.0000	.0001
M	.0023	.0003	.0012	.0001	.0000	.0028	.0013	.0029	.0004	.0000	.0003	.0000	.0000
N	.0030	.0107	.0020	.0022	.0000	.0066	.0050	.0065	.0019	.0000	.0001	.0002	.0006
P	.0002	.0000	.0001	.0001	.0000	.0022	.0009	.0002	.0000	.0000	.0000	.0000	.0000
Q	.0002	.0001	.0000	.0001	.0000	.0001	.0001	.0000	.0002	.0000	.0000	.0000	.0000
R	.0027	.0047	.0058	.0128	.0000	.0015	.0051	.0072	.0027	.0000	.0004	.0005	.0004
S	.0154	.0053	.0079	.0026	.0000	.0102	.0040	.0072	.0022	.0000	.0003	.0002	.0007
T	.0016	.0021	.0015	.0007	.0000	.0024	.0011	.0010	.0003	.0000	.0000	.0000	.0000
V	.0010	.0004	.0010	.0000	.0000	.0008	.0013	.0016	.0003	.0000	.0000	.0000	.0000
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0001	.0000	.0000	.0000	.0001	.0001	.0001	.0000	.0000	.0000	.0000	.0000
Y	.0003	.0009	.0000	.0001	.0000	.0003	.0002	.0002	.0001	.0000	.0000	.0000	.0000
Z	.0006	.0000	.0002	.0002	.0000	.0008	.0000	.0002	.0006	.0000	.0000	.0000	.0000

APPENDIX C. ENGLISH POPULATION VALUES (V-2)

1. Yule's 'K' based on the independent occurrence of each character

X_n :

Mean value 747.770

Standard Deviation 95.973

2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} :

Mean value 87.930

Standard Deviation 31.618

3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$:

Mean value 73.139

Standard Deviation 27.466

- 4.. Vowel-Consonant relationships are displayed in Table II.

TABLE X. ENGLISH (V-2) INDEPENDENT PROBABILITY OF X_n

X_n	Probability
A	.0659
E	.1000
I	.0585
O	.0638
U	.0234
B	.0114
C	.0258
D	.0344
F	.0183
G	.0179
H	.0414
J	.0014
K	.0059
L	.0367
M	.0210
N	.0603
P	.0173
Q	.0006
R	.0520
S	.0528
T	.0725
V	.0093
W	.0155
X	.0016
Y	.0157
Z	.0007
Space	.1760

TABLE XI. ENGLISH (V-2) JOINT PROBABILITY TABLE FOR X_n, X_{n+1}

X_{n+1}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0000	.0047	.0010	.0008	.0007	.0014	.0042	.0014	.0010	.0012	.0062	.0001	.0001
E	.0000	.0033	.0020	.0002	.0008	.0030	.0030	.0037	.0012	.0026	.0176	.0003	.0019
I	.0022	.0013	.0000	.0007	.0006	.0008	.0014	.0025	.0025	.0009	.0056	.0000	.0008
O	.0000	.0005	.0040	.0016	.0000	.0014	.0055	.0025	.0037	.0010	.0042	.0004	.0001
U	.0007	.0001	.0000	.0081	.0000	.0015	.0017	.0008	.0007	.0005	.0012	.0006	.0000
B	.0013	.0002	.0007	.0005	.0004	.0001	.0000	.0000	.0000	.0000	.0001	.0000	.0000
C	.0020	.0026	.0043	.0011	.0011	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000
D	.0025	.0086	.0023	.0011	.0005	.0000	.0000	.0003	.0000	.0000	.0000	.0000	.0000
F	.0008	.0008	.0013	.0049	.0001	.0000	.0000	.0000	.0011	.0000	.0000	.0000	.0000
G	.0012	.0006	.0025	.0005	.0010	.0000	.0000	.0001	.0000	.0002	.0000	.0000	.0000
H	.0000	.0001	.0000	.0002	.0000	.0000	.0035	.0000	.0000	.0023	.0000	.0000	.0000
J	.0001	.0000	.0000	.0001	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
K	.0008	.0002	.0003	.0006	.0000	.0000	.0017	.0000	.0000	.0000	.0000	.0000	.0000
L	.0069	.0035	.0033	.0034	.0020	.0013	.0009	.0003	.0005	.0004	.0001	.0000	.0002
M	.0020	.0020	.0024	.0034	.0008	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000
N	.0130	.0097	.0154	.0110	.0032	.0000	.0000	.0001	.0000	.0005	.0002	.0000	.0004
P	.0008	.0011	.0007	.0016	.0011	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Q	.0000	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
R	.0077	.0122	.0022	.0082	.0037	.0005	.0015	.0014	.0013	.0013	.0007	.0000	.0000
S	.0056	.0050	.0055	.0017	.0034	.0001	.0001	.0009	.0001	.0005	.0001	.0000	.0004
T	.0087	.0029	.0060	.0025	.0030	.0001	.0020	.0000	.0012	.0001	.0015	.0000	.0000
V	.0014	.0021	.0023	.0012	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
W	.0006	.0009	.0000	.0024	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
X	.0002	.0009	.0004	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0019	.0009	.0000	.0001	.0001	.0010	.0003	.0004	.0000	.0003	.0002	.0000	.0000
Z	.0001	.0000	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Space	.0055	.0329	.0016	.0077	.0008	.0002	.0009	.0194	.0051	.0062	.0036	.0000	.0021

TABLE XI - Continued

X_{n+1}	X_n														
	A	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Space
A	.0034	.0037	.0022	.0019	.0000	.0000	.0040	.0023	.0025	.0005	.0030	.0001	.0001	.0001	.0195
E	.0062	.0053	.0058	.0058	.0031	.0000	.0113	.0352	.0083	.0066	.0031	.0001	.0008	.0003	.0045
I	.0045	.0027	.0026	.0026	.0008	.0000	.0043	.0031	.0071	.0017	.0022	.0001	.0002	.0001	.0100
O	.0023	.0024	.0024	.0024	.0022	.0000	.0050	.0021	.0072	.0004	.0020	.0000	.0014	.0000	.0114
U	.0006	.0006	.0004	.0004	.0006	.0006	.0007	.0017	.0012	.0000	.0000	.0000	.0000	.0000	.0020
B	.0000	.0005	.0000	.0000	.0000	.0000	.0001	.0002	.0000	.0000	.0000	.0000	.0001	.0000	.0072
C	.0000	.0001	.0019	.0019	.0000	.0000	.0009	.0008	.0002	.0000	.0000	.0001	.0000	.0000	.0105
D	.0020	.0000	.0098	.0098	.0000	.0000	.0013	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0059
F	.0004	.0000	.0003	.0003	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0084
G	.0000	.0000	.0074	.0074	.0000	.0000	.0012	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0032
H	.0000	.0000	.0001	.0001	.0004	.0000	.0001	.0021	.0194	.0000	.0023	.0001	.0000	.0000	.0107
J	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0012
K	.0002	.0000	.0005	.0005	.0000	.0000	.0007	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0007
L	.0057	.0001	.0003	.0003	.0015	.0000	.0005	.0005	.0007	.0000	.0001	.0000	.0004	.0000	.0043
M	.0001	.0005	.0003	.0003	.0000	.0000	.0008	.0007	.0002	.0000	.0000	.0000	.0000	.0000	.0077
N	.0000	.0000	.0005	.0005	.0000	.0000	.0012	.0001	.0000	.0000	.0005	.0000	.0001	.0000	.0043
P	.0003	.0016	.0001	.0001	.0010	.0000	.0003	.0011	.0000	.0000	.0000	.0003	.0000	.0000	.0074
Q	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0003
R	.0001	.0001	.0000	.0000	.0030	.0000	.0007	.0000	.0026	.0000	.0001	.0000	.0002	.0000	.0046
S	.0010	.0005	.0025	.0025	.0004	.0000	.0039	.0023	.0019	.0000	.0003	.0000	.0004	.0000	.0132
T	.0005	.0000	.0065	.0065	.0004	.0000	.0023	.0066	.0011	.0000	.0000	.0005	.0001	.0000	.0264
V	.0002	.0000	.0006	.0006	.0000	.0000	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0011
W	.0000	.0000	.0001	.0001	.0000	.0000	.0001	.0003	.0011	.0000	.0000	.0000	.0000	.0000	.0098
X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0027	.0005	.0007	.0007	.0002	.0000	.0014	.0002	.0027	.0001	.0000	.0000	.0000	.0000	.0019
Z	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000
Space	.0064	.0026	.0153	.0153	.0016	.0000	.0108	.0230	.0161	.0000	.0019	.0004	.0118	.0000	.0000

TABLE XII. ENGLISH (V-2) JOINT PROBABILITY TABLE FOR X_n , X_{n+1} , X_{n+2}

X_{n+2}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0011	.0048	.0029	.0022	.0016	.0004	.0023	.0029	.0009	.0013	.0009	.0000	.0004
E	.0084	.0079	.0081	.0075	.0021	.0012	.0020	.0021	.0015	.0012	.0008	.0000	.0003
I	.0045	.0054	.0033	.0026	.0015	.0007	.0018	.0017	.0010	.0010	.0009	.0000	.0002
O	.0015	.0037	.0007	.0009	.0005	.0004	.0011	.0015	.0015	.0011	.0009	.0000	.0005
U	.0005	.0014	.0012	.0007	.0002	.0006	.0015	.0004	.0006	.0001	.0015	.0000	.0001
B	.0004	.0016	.0002	.0006	.0003	.0001	.0002	.0013	.0001	.0003	.0002	.0001	.0001
C	.0017	.0036	.0008	.0010	.0006	.0007	.0001	.0018	.0011	.0004	.0007	.0001	.0001
D	.0085	.0028	.0013	.0022	.0028	.0002	.0004	.0016	.0004	.0004	.0011	.0001	.0006
F	.0007	.0020	.0003	.0009	.0002	.0002	.0002	.0011	.0005	.0003	.0002	.0000	.0002
G	.0012	.0021	.0060	.0021	.0005	.0003	.0002	.0003	.0002	.0001	.0004	.0000	.0000
H	.0012	.0028	.0035	.0014	.0013	.0000	.0001	.0014	.0003	.0004	.0003	.0001	.0001
J	.0000	.0002	.0000	.0001	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000
K	.0012	.0006	.0009	.0014	.0002	.0000	.0000	.0001	.0000	.0000	.0001	.0000	.0000
L	.0035	.0024	.0023	.0043	.0013	.0007	.0018	.0017	.0017	.0004	.0017	.0000	.0002
M	.0009	.0022	.0006	.0015	.0001	.0000	.0016	.0011	.0004	.0003	.0017	.0001	.0001
N	.0015	.0029	.0047	.0026	.0006	.0007	.0041	.0021	.0006	.0010	.0033	.0001	.0007
P	.0013	.0032	.0007	.0015	.0005	.0000	.0004	.0008	.0003	.0003	.0007	.0000	.0001
Q	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
R	.0012	.0038	.0003	.0029	.0004	.0008	.0012	.0016	.0038	.0011	.0031	.0002	.0003
S	.0039	.0072	.0024	.0051	.0010	.0004	.0014	.0021	.0007	.0012	.0038	.0004	.0003
T	.0043	.0112	.0035	.0051	.0024	.0015	.0007	.0039	.0016	.0034	.0031	.0000	.0004
V	.0002	.0009	.0005	.0003	.0001	.0001	.0001	.0004	.0008	.0004	.0009	.0000	.0000
W	.0004	.0024	.0003	.0004	.0001	.0001	.0001	.0011	.0003	.0003	.0009	.0000	.0002
X	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0001	.0000	.0000	.0000	.0000
Y	.0014	.0012	.0010	.0005	.0001	.0003	.0001	.0009	.0003	.0002	.0007	.0000	.0002
Z	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Space	.0161	.0239	.0129	.0164	.0049	.0020	.0045	.0024	.0007	.0026	.0136	.0001	.0011

TABLE XII - Continued

X_n														
	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Space
A	.0027	.0011	.0031	.0010	.0003	.0034	.0064	.0050	.0001	.0006	.0001	.0020	.0000	.0186
E	.0021	.0010	.0045	.0022	.0002	.0045	.0034	.0156	.0001	.0010	.0003	.0008	.0000	.0211
I	.0013	.0007	.0031	.0015	.0002	.0025	.0041	.0036	.0001	.0010	.0001	.0009	.0000	.0151
O	.0016	.0005	.0023	.0027	.0000	.0018	.0039	.0065	.0001	.0007	.0001	.0008	.0000	.0288
U	.0004	.0003	.0007	.0002	.0000	.0011	.0013	.0008	.0000	.0006	.0000	.0013	.0000	.0081
B	.0005	.0001	.0006	.0003	.0000	.0011	.0012	.0008	.0000	.0001	.0000	.0005	.0000	.0008
C	.0014	.0007	.0012	.0006	.0000	.0029	.0019	.0021	.0003	.0002	.0001	.0007	.0000	.0012
D	.0012	.0010	.0012	.0005	.0000	.0027	.0016	.0022	.0005	.0003	.0001	.0006	.0001	.0005
F	.0008	.0002	.0007	.0001	.0000	.0010	.0015	.0008	.0000	.0002	.0001	.0009	.0000	.0056
G	.0009	.0004	.0003	.0001	.0000	.0010	.0005	.0006	.0000	.0001	.0000	.0002	.0000	.0003
H	.0005	.0002	.0015	.0002	.0000	.0015	.0014	.0013	.0000	.0002	.0001	.0007	.0000	.0211
J	.0001	.0001	.0001	.0000	.0000	.0002	.0002	.0001	.0000	.0000	.0000	.0001	.0000	.0000
K	.0003	.0005	.0001	.0000	.0000	.0001	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0001
L	.0002	.0014	.0017	.0011	.0000	.0017	.0018	.0021	.0007	.0010	.0001	.0003	.0000	.0038
M	.0007	.0002	.0010	.0001	.0000	.0025	.0019	.0018	.0001	.0002	.0001	.0005	.0000	.0015
N	.0019	.0034	.0019	.0011	.0000	.0022	.0025	.0028	.0013	.0012	.0001	.0008	.0002	.0162
P	.0006	.0001	.0007	.0002	.0000	.0015	.0012	.0008	.0000	.0001	.0000	.0004	.0000	.0018
Q	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002
R	.0021	.0021	.0030	.0027	.0000	.0005	.0024	.0048	.0023	.0018	.0001	.0005	.0001	.0090
S	.0025	.0019	.0029	.0009	.0000	.0036	.0021	.0022	.0007	.0016	.0001	.0012	.0000	.0031
T	.0025	.0019	.0065	.0008	.0000	.0035	.0038	.0034	.0004	.0016	.0001	.0017	.0001	.0054
V	.0004	.0002	.0004	.0000	.0000	.0009	.0008	.0005	.0001	.0002	.0000	.0001	.0000	.0011
W	.0010	.0002	.0016	.0002	.0000	.0011	.0017	.0010	.0000	.0001	.0000	.0006	.0000	.0016
X	.0000	.0000	.0002	.0000	.0000	.0000	.0004	.0002	.0000	.0000	.0000	.0000	.0000	.0007
Y	.0010	.0003	.0013	.0004	.0000	.0013	.0006	.0011	.0000	.0007	.0002	.0001	.0000	.0021
Z	.0010	.0000	.0002	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Space	.0100	.0027	.0197	.0006	.0000	.0094	.0061	.0123	.0027	.0024	.0001	.0003	.0001	.0084
X_{n+2}														

 X_{n+2}

APPENDIX D. SPANISH POPULATION VALUES (V-2)

1. Yule's 'K' based on the independent occurrence of each character

X_n :

Mean value	806.481
Standard Deviation	77.810

2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} :

Mean value	108.784
Standard Deviation	30.665

3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$:

Mean value	83.553
Standard Deviation	22.886

4. Vowel-Consonant relationships are displayed in Table VI.

TABLE XIII. SPANISH (V-2) INDEPENDENT PROBABILITY OF X_n

X_n	Probability
A	.1053
E	.1088
I	.0573
O	.0771
U	.0339
B	.0114
C	.0372
D	.0390
F	.0060
G	.0092
H	.0072
J	.0034
K	.0005
L	.0471
M	.0236
N	.0620
P	.0205
Q	.0073
R	.0514
S	.0625
T	.0385
V	.0095
W	.0000
X	.0011
Y	.0074
Z	.0031
Space	.1699

TABLE XIV. SPANISH (V-2) JOINT PROBABILITY TABLE FOR X_n, X_{n+1}

X_{n+1}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0000	.0011	.0076	.0001	.0021	.0036	.0070	.0050	.0006	.0018	.0033	.0008	.0000
E	.0002	.0002	.0071	.0000	.0096	.0008	.0033	.0183	.0010	.0010	.0008	.0009	.0001
I	.0006	.0006	.0000	.0002	.0019	.0022	.0087	.0044	.0017	.0011	.0009	.0001	.0002
O	.0002	.0007	.0068	.0002	.0002	.0008	.0088	.0079	.0010	.0016	.0018	.0012	.0000
U	.0008	.0004	.0001	.0000	.0000	.0005	.0031	.0011	.0009	.0019	.0004	.0004	.0000
B	.0043	.0004	.0011	.0013	.0005	.0000	.0000	.0000	.0000	.0009	.0000	.0000	.0000
C	.0054	.0041	.0046	.0022	.0014	.0000	.0007	.0000	.0000	.0000	.0000	.0000	.0000
D	.0067	.0025	.0035	.0017	.0010	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
F	.0003	.0009	.0005	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
G	.0005	.0019	.0014	.0007	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
H	.0002	.0000	.0000	.0000	.0000	.0000	.0023	.0000	.0000	.0000	.0000	.0000	.0000
J	.0006	.0009	.0003	.0003	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
K	.0000	.0001	.0000	.0001	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000
L	.0077	.0092	.0032	.0024	.0013	.0012	.0006	.0000	.0003	.0002	.0000	.0000	.0001
M	.0036	.0023	.0027	.0035	.0008	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
N	.0116	.0200	.0063	.0108	.0073	.0000	.0001	.0000	.0000	.0002	.0000	.0000	.0000
P	.0010	.0007	.0003	.0009	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Q	.0004	.0002	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
R	.0114	.0107	.0018	.0067	.0020	.0017	.0006	.0007	.0005	.0015	.0000	.0000	.0000
S	.0104	.0155	.0033	.0152	.0021	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0001
T	.0025	.0025	.0023	.0014	.0007	.0001	.0019	.0000	.0000	.0000	.0000	.0000	.0000
V	.0006	.0010	.0016	.0015	.0002	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0010	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0006	.0002	.0000	.0003	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Z	.0006	.0010	.0005	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Space	.0352	.0311	.0022	.0268	.0013	.0000	.0000	.0014	.0000	.0000	.0000	.0000	.0001

TABLE XIV - Continued

X_{n+1}	X_n														
	A	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Space
A	.0133	.0050	.0072	.0052	.0000	.0103	.0034	.0093	.0019	.0000	.0000	.0001	.0006	.0014	.0146
E	.0050	.0057	.0031	.0034	.0000	.0085	.0073	.0096	.0032	.0000	.0000	.0000	.0003	.0000	.0196
I	.0039	.0041	.0030	.0013	.0000	.0059	.0045	.0054	.0032	.0000	.0000	.0003	.0000	.0000	.0032
O	.0075	.0035	.0062	.0041	.0000	.0072	.0042	.0075	.0010	.0000	.0000	.0000	.0006	.0007	.0034
U	.0007	.0015	.0016	.0014	.0000	.0072	.0030	.0016	.0001	.0000	.0000	.0000	.0001	.0001	.0058
B	.0001	.0012	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0022
C	.0002	.0000	.0032	.0000	.0000	.0000	.0012	.0000	.0000	.0000	.0000	.0001	.0000	.0001	.0134
D	.0002	.0000	.0031	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0193
F	.0001	.0000	.0006	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0028
G	.0005	.0000	.0008	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0021
H	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0045
J	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0010
K	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
L	.0028	.0000	.0000	.0000	.0015	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0159
M	.0004	.0000	.0001	.0000	.0000	.0000	.0018	.0006	.0000	.0000	.0000	.0000	.0000	.0000	.0078
N	.0000	.0000	.0003	.0000	.0000	.0000	.0004	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0049
P	.0001	.0025	.0000	.0000	.0000	.0000	.0002	.0017	.0000	.0000	.0000	.0003	.0000	.0000	.0123
Q	.0000	.0000	.0004	.0000	.0000	.0000	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0057
R	.0001	.0000	.0000	.0000	.0032	.0000	.0014	.0000	.0050	.0000	.0000	.0000	.0000	.0000	.0039
S	.0003	.0000	.0017	.0000	.0000	.0000	.0013	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0124
T	.0008	.0000	.0116	.0001	.0001	.0000	.0020	.0065	.0000	.0000	.0000	.0002	.0000	.0000	.0061
V	.0004	.0000	.0006	.0000	.0000	.0000	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0031
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Z	.0000	.0000	.0005	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0058
Space	.0106	.0000	.0178	.0001	.0000	.0000	.0071	.0295	.0001	.0001	.0000	.0000	.0058	.0007	n/c

TABLE XV. SPANISH (V-2) JOINT PROBABILITY TABLE FOR X_n , -, X_{n+2}

X_{n+2}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0123	.0095	.0086	.0074	.0054	.0016	.0047	.0013	.0004	.0012	.0000	.0001	.0000
E	.0102	.0090	.0031	.0078	.0016	.0024	.0026	.0013	.0010	.0010	.0001	.0001	.0000
I	.0092	.0074	.0049	.0051	.0019	.0008	.0022	.0003	.0002	.0007	.0000	.0000	.0000
O	.0088	.0059	.0052	.0045	.0027	.0004	.0043	.0008	.0001	.0008	.0000	.0000	.0000
U	.0032	.0029	.0014	.0018	.0003	.0001	.0007	.0000	.0001	.0001	.0000	.0000	.0000
B	.0011	.0007	.0002	.0008	.0001	.0001	.0005	.0004	.0001	.0000	.0013	.0001	.0000
C	.0042	.0051	.0011	.0021	.0011	.0002	.0004	.0015	.0012	.0002	.0010	.0001	.0000
D	.0056	.0037	.0009	.0051	.0012	.0002	.0012	.0020	.0000	.0003	.0001	.0001	.0000
F	.0008	.0007	.0003	.0005	.0001	.0000	.0001	.0004	.0000	.0000	.0000	.0001	.0000
G	.0016	.0007	.0004	.0008	.0005	.0000	.0001	.0000	.0001	.0000	.0001	.0000	.0000
H	.0011	.0014	.0001	.0014	.0006	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
J	.0002	.0002	.0002	.0002	.0000	.0003	.0000	.0003	.0000	.0000	.0001	.0000	.0000
K	.0001	.0001	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
L	.0049	.0059	.0025	.0022	.0012	.0004	.0019	.0021	.0002	.0003	.0003	.0000	.0001
M	.0079	.0024	.0016	.0018	.0003	.0001	.0028	.0007	.0001	.0001	.0007	.0002	.0000
N	.0015	.0020	.0082	.0010	.0016	.0008	.0056	.0018	.0008	.0013	.0005	.0003	.0000
P	.0030	.0045	.0009	.0025	.0004	.0000	.0007	.0002	.0000	.0000	.0000	.0000	.0000
Q	.0015	.0010	.0002	.0015	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
R	.0029	.0028	.0019	.0020	.0016	.0010	.0027	.0022	.0009	.0011	.0006	.0004	.0000
S	.0033	.0039	.0026	.0029	.0015	.0004	.0025	.0044	.0003	.0005	.0008	.0005	.0000
T	.0061	.0133	.0028	.0029	.0019	.0004	.0004	.0006	.0002	.0003	.0000	.0001	.0000
V	.0012	.0013	.0004	.0007	.0005	.0000	.0001	.0004	.0000	.0000	.0000	.0004	.0000
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0015	.0008	.0001	.0010	.0000	.0000	.0001	.0001	.0000	.0000	.0002	.0000	.0001
Z	.0004	.0003	.0004	.0000	.0001	.0000	.0001	.0000	.0000	.0000	.0002	.0000	.0000
Space	.0176	.0235	.0095	.0210	.0094	.0019	.0036	.0181	.0002	.0013	.0011	.0008	.0001

TABLE XV - Continued

X_{n+2}	X_n													
	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Space
A	.0032	.0008	.0061	.0011	.0000	.0055	.0064	.0026	.0004	.0000	.0000	.0005	.0001	.0263
E	.0023	.0012	.0100	.0028	.0060	.0048	.0054	.0018	.0008	.0000	.0001	.0006	.0001	.0326
I	.0010	.0008	.0040	.0012	.0012	.0024	.0018	.0005	.0002	.0000	.0000	.0000	.0000	.0116
O	.0017	.0007	.0060	.0016	.0000	.0039	.0032	.0017	.0004	.0000	.0000	.0002	.0002	.0239
U	.0005	.0001	.0024	.0002	.0000	.0009	.0018	.0003	.0000	.0000	.0000	.0002	.0000	.0171
B	.0011	.0001	.0004	.0002	.0000	.0012	.0009	.0008	.0001	.0000	.0000	.0001	.0001	.0009
C	.0029	.0008	.0032	.0012	.0000	.0028	.0029	.0021	.0007	.0000	.0000	.0005	.0001	.0018
D	.0020	.0012	.0026	.0009	.0000	.0019	.0054	.0025	.0006	.0000	.0000	.0006	.0002	.0010
F	.0045	.0000	.0004	.0000	.0000	.0006	.0006	.0003	.0000	.0000	.0000	.0001	.0000	.0006
G	.0011	.0002	.0004	.0000	.0000	.0010	.0010	.0005	.0001	.0000	.0000	.0001	.0000	.0004
H	.0003	.0000	.0005	.0000	.0000	.0002	.0008	.0000	.0000	.0000	.0000	.0002	.0000	.0007
J	.0003	.0003	.0002	.0000	.0000	.0003	.0004	.0001	.0000	.0000	.0000	.0000	.0000	.0003
K	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
L	.0005	.0022	.0032	.0009	.0000	.0017	.0035	.0015	.0006	.0000	.0002	.0007	.0001	.0100
M	.0016	.0002	.0014	.0000	.0000	.0016	.0016	.0015	.0002	.0000	.0002	.0004	.0001	.0013
N	.0020	.0053	.0014	.0011	.0000	.0034	.0033	.0033	.0010	.0000	.0001	.0004	.0003	.0152
P	.0014	.0000	.0010	.0000	.0000	.0014	.0029	.0001	.0000	.0000	.0000	.0003	.0001	.0008
Q	.0003	.0001	.0005	.0000	.0000	.0004	.0010	.0000	.0001	.0000	.0000	.0001	.0001	.0004
R	.0018	.0023	.0033	.0064	.0000	.0009	.0032	.0036	.0013	.0000	.0002	.0004	.0002	.0076
S	.0083	.0027	.0052	.0013	.0000	.0056	.0042	.0036	.0011	.0000	.0001	.0008	.0004	.0056
T	.0014	.0010	.0016	.0004	.0000	.0015	.0016	.0005	.0001	.0000	.0000	.0002	.0000	.0011
V	.0007	.0002	.0008	.0000	.0000	.0005	.0011	.0008	.0001	.0000	.0000	.0000	.0000	.0003
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0009
Y	.0003	.0005	.0004	.0001	.0000	.0004	.0015	.0001	.0000	.0000	.0000	.0001	.0002	.0001
Z	.0003	.0000	.0001	.0001	.0000	.0004	.0000	.0001	.0003	.0000	.0000	.0000	.0000	.0001
Space	.0117	.0031	.0071	.0010	.0000	.0082	.0076	.0103	.0014	.0000	.0000	.0010	.0010	.0095

APPENDIX E. TABLE XVI - Significance Levels for Kolmogorov-Smirnov Test

z	α	z	α	z	α	z	α	z	α
0.28	.999999	0.59	.877240	0.90	.392730	1.21	.106970	1.52	.019690
0.29	.999996	0.60	.864282	0.91	.379072	1.22	.101896	1.53	.018524
0.30	.999991	0.61	.850771	0.92	.365714	1.23	.097028	1.54	.017422
0.31	.999979	0.62	.836775	0.93	.352662	1.24	.092352	1.55	.016378
0.32	.999954	0.63	.822247	0.94	.339918	1.25	.087868	1.56	.015390
0.33	.999909	0.64	.807333	0.95	.327484	1.26	.083568	1.57	.014456
0.34	.999829	0.65	.792013	0.96	.315364	1.27	.079444	1.58	.013574
0.35	.999697	0.66	.776363	0.97	.303556	1.28	.075495	1.59	.012740
0.36	.999489	0.67	.760418	0.98	.292060	1.29	.071712	1.60	.011952
0.37	.999174	0.68	.744220	0.99	.280874	1.30	.068092	1.61	.011209
0.38	.998715	0.69	.727811	1.00	.270000	1.31	.064630	1.62	.010508
0.39	.998071	0.70	.711235	1.01	.259434	1.32	.061318	1.63	.009846
0.40	.997192	0.71	.694529	1.02	.249174	1.33	.058152	1.64	.009223
0.41	.996028	0.72	.677735	1.03	.239220	1.34	.055128	1.65	.008636
0.42	.994524	0.73	.660887	1.04	.229566	1.35	.052244	1.66	.008083
0.43	.992623	0.74	.644019	1.05	.220206	1.36	.049488	1.67	.007562
0.44	.990270	0.75	.627167	1.06	.211140	1.37	.046858	1.68	.007072
0.45	.987410	0.76	.610360	1.07	.202364	1.38	.044350	1.69	.006611
0.46	.983995	0.77	.593628	1.08	.193872	1.39	.041960	1.70	.006177
0.47	.979978	0.78	.576998	1.09	.185658	1.40	.039682	1.71	.005770
0.48	.975318	0.79	.560495	1.10	.177718	1.41	.037514	1.72	.005388
0.49	.969983	0.80	.544143	1.11	.170050	1.42	.035448	1.73	.005028
0.50	.963645	0.81	.527959	1.12	.162644	1.43	.033484	1.74	.004691
0.51	.957186	0.82	.511970	1.13	.155498	1.44	.031618	1.75	.004375
0.52	.949694	0.83	.496192	1.14	.148606	1.45	.029842	1.76	.004078
0.53	.941466	0.84	.480634	1.15	.141962	1.46	.028154	1.77	.003800
0.54	.932503	0.85	.465318	1.16	.135558	1.47	.026552	1.78	.003540
0.55	.922817	0.86	.450256	1.17	.129388	1.48	.025030	1.79	.003296
0.56	.912423	0.87	.435454	1.18	.123452	1.49	.023588	1.80	.003068
0.57	.901344	0.88	.420930	1.19	.117742	1.50	.022218	1.81	.002845
0.58	.889605	0.89	.406684	1.20	.112250	1.51	.020920	1.82	.002645

APPENDIX F. FINAL TEST SOURCES

From Introduction to Literature: Stories [Ref. 10]:

Anderson, Sherwood, Death in the Woods.
Benét, Stephen Vincent, By The Waters of Babylon.
Crane, Stephen, The Open Boat.
Curley, Daniel, A Story of Love, Etc.
Edmonds, Walter D., Death of Red Peril.
Fitzgerald, F. Scott, Absolution.
Forster, E. M., The Other Side of the Hedge.
Galsworthy, John, Quality.
Hardy, Thomas, The Three Strangers.
Porter, Katherine Anne, The Grave.
Schwartz, Delmore, In Dreams Begin Responsibilities.
Steele, Wilbur Daniel, Footfalls.
Steinbeck, John, from The Red Pony: The Gift.
Stevenson, Robert Louis, Markheim.
White, E. B., The Dove.

From Five Centuries of Spanish Literature [Ref. 11]:

Anonymous, Lazarillo de Tormes.
Cervantes Saavedra, Miguel De, El ingenioso hidalgo, don Quijote de la Mancha.
Manuel, Don Juan, El Conde Lucanor.

From Antología Venezolana [Ref. 12]:

Alfonzo, Alfredo Armas, La hora y punto.
Arráiz, Antonio, El mar es como un potro vigoroso.

Coll, Pedro Emilio, El diente roto.
Croce, Arturo, Angustia apagada.
Diaz Rodríguez, Manuel, Mi Alma era una mina abandonada.
Diaz Sanchez, Ramon, Poderes de la mas alta valia.
Diaz Solis, Gustavo, La efigie.
Fombona, Rufino Blanco, El catire.
Gallegos, Romulo, Canaima.
Garmendia, Julio, Eladia.
Guaramato, Oscar, Vecindad.
La Parra, Teresa De, Vicente Cochocho.
Marquez Salas, Antonio, Como Dios!.
Meneses, Guillermo, El compañero Juan Ruiz.
Mujica, Hector, Las tres ventanas.
Núñez, Enrique Bernardo, Vocchi.
Palacios, Lucila, Cap. XX (Cubil).
Picon Salas, Mariano, Josefita.
Pocaterra, Jose Rafael, Patria, la mestiza.
Pardo, Isaac J., El domador de potros.
Pardon, Julian, II(Primavera nocturna).
Rivas Mijares, Humberto, El murado.
Rosales, Julio, El mejor rabula.
Ruiz, Jose Fabbiani, A orillas del viejo rio.
Silva, Miguel Otero, Fiebre.
Stolk, Gloria, Cap. I (Amargo fondo).
Urbaneja Achelpohl, Luis M., Ovejon...!.
Uslar Pietri, Arturo, X (Las lanzas coloradas).

From An Anthology of Spanish American Literature [Ref. 13]:

Altamirano, Ignacio Manuel, El dia de muertos.

Bastamente Carlos Inga, Calixto (Concolorcorvo), El lazarillo de ciegos caminantes.

Bello, Andres, Silva a la Argicultura de la Zona Torrida.

Blanco-Fombona, Rufino, El conquistador español del siglo XVI.

Bolivar, Simon, Carta a un caballero que tomaba gran interes en la causa republicana en la america del sur.

Chocano, Jose Santos, La Epopeya Del Pacifico.

Cortes, Hernan, The Death of Moctezuma and the Tragic Retreat of the Spaniards from the Aztec Capital the Night of June 30, 1520.

Dario, Ruben, Azul (El rey burgues).

Diaz del Castillo, Bernal, The Imprisonment of Moctezuma.

Fernandez de Lizardi, Jose Joaquin, El periquillo sarniento.

Garcia Calderon, Francisco, La creacion de un continente.

Garcilaso de la Vega, El Inca, Comentarios reales de los incas.

Gonzalez Prada, Manuel, Discurso en el politeama.

Gutierrez Najera, Manuel, La Novela Del Tranvia.

Hernandez, Jose, Martin Fierro.

Juana Ines de la Cruz, Sor, Respuesta a Sor Filotea de la Cruz.

Latorre, Mariano, Chilenos del mar (El piloto Oyarzo).

Lillo, Baldomero, El chiflon del diablo.

Lopez Albuja, Enrique, Cuentos andinos (Como habla la coca).

Marti, Jose, Nuestra America.

Mitre, Bartolome, La historia de San Martin.

Montalvo, Juan, Washington y Bolivar.

Nervo, Amado, Una esperanza.

Palma, Ricardo, Las orejas del Alcalde.

Quiroga, Horacio, El desierto.

Reyes, Alfonso, Vision de Anahuac.

Rodo, Jose Enrique, Ariel.

Rojas, Manuel, Hombres del sur (El machorro).

Rojas, Ricardo, La argentinidad.

Sarmiento, Domingo Faustino, Facundo.

Siguenza y Gongora, Carlos de, Infortunios de Alonso Rameriz.

Ugarte, Manuel, El destino de un continente (La neuva Roma).

Vasconcelos, Jose, La raza cosmica.

Viana, Javier de, Leña seca (El domador).

From The Creative Reader [Ref. 14]:

Algren, Nelson, The Moon of the Arfy Darfy.

Amster, L. J., Center of Gravity.

Auchincloss, Louis, Power in Trust.

Berriault, Gina, The Birthday Party.

Boyle, Kay, The Ballet of Central Park.

Buechler, James, John Sobieski Runs.

Calisher, Hortense, Little Did I Know.

Connell, Evan S., Jr., The Suicide.

Cozzens, James Gould, One Hundred Ladies.

Elkin, Stanley, Perlmutter at the East Pole.

Gary, Romain, A Humanist.

Gold, Hebert, Dance of the Divorced.

Goyen, William, Figure over the Town.

Greene, Graham, The Root of All Evil.

Malamud, Bernard, The Refugee.

McCullers, Carson, Sucker.

Miller, Warren, Chaos, Disorder and the Late Show.

Moravia, Alberto, A Tough Nut.
O'Connor, Frank, A Life of Your Own.
O'Faolain, Sean, One Man, One Boat, One Girl.
Pynchon, Thomas, The Secret Integration.
Saroyan, William, Boys and Girls Together.
Seager, Allan, No More Roses.
Shaw, Irwin, The Inhabitants of Venus.
Singer, Isaac Bashevis, Esther Kreindel The Second.
Swados, Harvey, A Story For Teddy.
Updike, John, The Lucid Eye in Silver Town.
White, Robin, Walker's Peak.
Williams, Thomas, The Old Dancers.
Wood, Malcolm, The Appraiser.

From Best Modern Short Stories [Ref. 15]:

Aiken, Conrad, Mr. Arcularis.
Audry, Colette, The Gloves.
Bierce, Ambrose, An Occurrence at Owl Creek Bridge.
Bowen, Elizabeth, The Demon Lover.
Conrad, Hoseph, Amy Foster.
Faulkner, William, A Rose For Emily.
Granberry, Edwin, A Trip to Czardis.
Hawthorne, Nathaniel, My Kinsman, Major Molineux.
Hemingway, Ernest, A Clean, Well-Lighted Place.
James, Henry, Paste.
Joyce, James, A Little Cloud.
Lawrence, D. H., The Horse Dealer's Daughter.
Lowry, Malcolm, Strange Comfort Afforded by the Profession.

Mansfield, Katherine, The Fly.

Maupassant, Guy de, Two Friends.

Poe, Edgar Allan, Ligeia.

Taylor, Elizabeth, The First Death of Her Life.

Thurber, James, The Catbird Seat.

Welty, A Still Moment.

West, Jessamyn, Sixteen.

[illegible]

57

007050	05	FILLER PIC X(19) VALUE SPACES.	ANAL-ONE
007100	05	PROB-ONE PIC 9.99999.	ANAL-ONE
007150	05	FILLER PIC X(37) VALUE SPACES.	ANAL-ONE
007200	01	OUTPT-ONE-F.	ANAL-ONE
007250	05	FILLER PIC X(38) VALUE SPACES.	ANAL-ONE
007300	05	FILLER PIC X(22) VALUE TOTAL.	ANAL-ONE
007350	05	ONE-F-ILF-TLT PIC ZZ,ZZ,ZZ9.	ANAL-ONE
007400	05	FILLER PIC X(63) VALUE SPACES.	ANAL-ONE
007450	01	OUTPT-TWO-A.	ANAL-ONE
007500	05	FILLER PIC X(46) VALUE SPACES.	ANAL-ONE
007550	05	FILLER PIC X(87) VALUE FREQUENCY AND PROBABILITY OF (X(2	ANAL-ONE
007600	01	)/X(1)).	ANAL-ONE
007650	05	THREE-B.	ANAL-ONE
007700	05	FILLER PIC X(46) VALUE SPACES.	ANAL-ONE
007750	05	FILLER PIC X(40) VALUE ALL.	ANAL-ONE
007800	05	FILLER PIC X(47) VALUE SPACES.	ANAL-ONE
007850	01	FIRST-OUTPT.	ANAL-ONE
007900	05	FILLER PIC X(31) VALUE SPACES.	ANAL-ONE
007950	05	FILLER PIC X(102) VALUE FIRST LETTER, X(1)*****	ANAL-ONE
008000	01	*****	ANAL-ONE
008050	05	FIRST-LTR-TITLES.	ANAL-ONE
008100	05	FILLER PIC X(14) VALUE SPACES.	ANAL-ONE
008150	05	FIRST-LTR PIC X(9); OCCURS 13 TIMES.	ANAL-ONE
008200	01	FILLER-LTR PIC XX VALUE SPACES.	ANAL-ONE
008250	05	DEPN-FREQ-LINE.	ANAL-ONE
008300	05	FILLER PIC X(5) VALUE SPACES.	ANAL-ONE
008350	05	FREQ-RCH-LTR PIC X.	ANAL-ONE
008400	05	FILLER PIC XXX VALUE (F).	ANAL-ONE
008450	05	DEPN-FREQ PIC B8ZZ,ZZ9 OCCURS 13 TIMES.	ANAL-ONE
008500	05	FILLER PIC X(7) VALUE SPACES.	ANAL-ONE
008550	01	DEPN-PROB-LINE.	ANAL-ONE
008600	05	FILLER PIC X(5) VALUE SPACES.	ANAL-ONE
008650	05	PROB-RCH-LTR PIC X.	ANAL-ONE
008700	05	FILLER PIC XXX VALUE (P).	ANAL-ONE
008750	05	DEPN-PROB PIC B89.99999 OCCURS 13 TIMES.	ANAL-ONE
008800	05	FILLER PIC X(7) VALUE SPACES.	ANAL-ONE
008850	05	FILLER PIC X(46) VALUE SPACES.	ANAL-ONE
008900	05	FILLER PIC X(87) VALUE FREQUENCY AND PROBABILITY OF (X(3	ANAL-ONE
009000	01	)/X(1)).	ANAL-ONE
009050	05	THREE-A.	ANAL-ONE
009100	05	FILLER PIC X(47) VALUE SPACES.	ANAL-ONE
009150	05	FILLER PIC X(33) VALUE P(VOWEL/VOWEL).	ANAL-ONE
009200	05	P-VOW-VOW PIC 9.99999.	ANAL-ONE
009250	05	FILLER PIC X(46) VALUE SPACES.	ANAL-ONE
009300	05	FILLER PIC X(47) VALUE SPACES.	ANAL-ONE
009350	05	FILLER PIC X(33) VALUE P(CONSONANT/VOWEL).	ANAL-ONE
009400	05	FILLER PIC X(47) VALUE SPACES.	ANAL-ONE

009450	05	P-CONS-VCW	PIC	9.99999	SPACES.	ANAL--ONE
009500	05	FILLER-PIC	X(46)	VALUE	SPACES.	ANAL--ONE
009550	01	OUTPT-FOUR-C				ANAL--ONE
009600	05	FILLER-PIC	X(47)	VALUE	SPACES.	ANAL--ONE
009650	05	FILLER-PIC	X(33)	VALUE	*P(VCWEL/CONSONANT)*.	ANAL--ONE
009700	05	P-VOW-CONS	PIC	9.99999		ANAL--ONE
009750	05	FILLER-PIC	X(46)	VALUE	SPACES.	ANAL--ONE
009800	01	OUTPT-FCUP-D				ANAL--ONE
009850	05	FILLER-PIC	X(47)	VALUE	SPACES.	ANAL--ONE
009900	05	FILLER-PIC	X(33)	VALUE	*P(CONSONANT/CONSONANT)*.	ANAL--ONE
009950	05	P-CONS-CUNS	PIC	9.99999		ANAL--ONE
010000	05	FILLER-PIC	X(46)	VALUE	SPACES.	ANAL--ONE
010050	01	OUTPT-FOUR-E				ANAL--ONE
010100	05	FILLER-PIC	X(11)	VALUE	SPACES.	ANAL--ONE
010150	05	FILLER-PIC	X(32)	VALUE	*YULE-S	ANAL--ONE
010200	05	OUTPT-FOUR-LANG	PIC	X(30)		ANAL--ONE
010250	05	YULES-K	PIC	Z(12)9.999		ANAL--ONE
010300	05	FILLER-PIC	X(40)	VALUE	SPACES.	ANAL--ONE
010350	01	OUTPT-FOUR-F				ANAL--ONE
010400	05	FILLER-PIC	X			ANAL--ONE
010450	05	FILLER-PIC	X(94)	VALUE	*KOLMGOROV-SMIRNOV CUMULATED REL	ANAL--ONE
010500	05	ATIVE FREQUENCY	TABLE	FOR	INDEPENDENT LETTER OCCURRENCES IN:	ANAL--ONE
010550	05	OUTPT-FOUR-KS-LANG	PIC	X(30)		ANAL--ONE
010600	05	FILLER-PIC	X(9)	VALUE	SPACES.	ANAL--ONE
010650	01	OUTPT-FOUR-G				ANAL--ONE
010700	05	FILLER-PIC	X(37)	VALUE	SPACES.	ANAL--ONE
010750	05	FILLER-PIC	X(25)	VALUE	*LETTER*	ANAL--ONE
010800	05	FILLER-PIC	X(71)	VALUE	*CUM. REL. FREQ.*.	ANAL--ONE
010850	01	KS-LINE				ANAL--ONE
010900	05	FILLER-PIC	X(40)	VALUE	SPACES.	ANAL--ONE
010950	05	KS-LTR	PIC	X(25)	VALUE	SPACES.
011000	05	FILLER-PIC	9.99999			ANAL--ONE
011050	05	KS-CRF	PIC	X(60)	VALUE	SPACES.
011100	05	FILLER-PIC				ANAL--ONE
011150	01	ERQF-LINE				ANAL--ONE
011200	05	FILLER-PIC	X			ANAL--ONE
011250	05	FILLER-PIC	X(28)	VALUE	*OVERFLOW ERROR AT STATEMENT *	ANAL--ONE
011300	05	ER-STEP	PIC	9(6)		ANAL--ONE
011350	05	FILLER-PIC	X(21)	VALUE	* CARD IN PROCESS WAS *	ANAL--ONE
011400	05	ER-CARD	PIC	9(6)		ANAL--ONE
011450	05	FILLER-PIC	X(71)	VALUE	ALL '*'. *	ANAL--ONE
011500	01	DATA-FORM				ANAL--ONE
011550	05	CC	PIC	X		ANAL--ONE
011600	05	FILLER-PIC	X(7)			ANAL--ONE
011650	05	OUTPT-FIVE-A				ANAL--ONE
011700	05	FILLER-PIC	X(49)	VALUE	SPACES.	ANAL--ONE
011750	05	FILLER-PIC	X(37)	VALUE	*JOINT PROBABILITY TABLE FOR X(1)-ANAL--ONE	ANAL--ONE
011800	05					


```

019050 PERFORM COL-TITLE-B VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-ONE
019100 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
019150 PERFORM OUTPUT-TWO-DUMP-B VARYING B FROM 34 BY 1 UNTIL B > 56. ANAL-ONE
019200 PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
019250 PERFORM COL-TITLE-C VARYING A FROM 27 BY 1 UNTIL A > 33. ANAL-ONE
019300 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
019350 PERFORM OUTPUT-TWO-DUMP-C VARYING B FROM 34 BY 1 UNTIL B > 56. ANAL-ONE
019400 PERFORM COL-TITLE-THREE-PREP. ANAL-ONE
019450 PERFORM COL-TITLE-A VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
019500 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
019550 PERFORM OUTPUT-TWO-DUMP-A VARYING B FROM 57 BY 1 UNTIL B > 66. ANAL-ONE
019600 PERFORM COL-TITLE-THREE-PREP. ANAL-ONE
019650 PERFORM COL-TITLE-B VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-ONE
019700 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
019750 PERFORM OUTPUT-TWO-DUMP-B VARYING B FROM 57 BY 1 UNTIL B > 66. ANAL-ONE
019800 PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
019850 PERFORM COL-TITLE-C VARYING A FROM 27 BY 1 UNTIL A > 33. ANAL-ONE
019900 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
020000 PERFORM OUTPUT-TWO-DUMP-C VARYING B FROM 57 BY 1 UNTIL B > 66. ANAL-ONE
020050 PERFORM VC-DUMP. ANAL-ONE
020100 PERFORM YULES-K * 2). ANAL-ONE
020150 COMPUTE ILF-TOTAL = (10000 * (S - ILF-TOTAL)) / ANAL-ONE
020200 WRITE PRINT-LINE FROM OUTPUT-FOUR-E BEFORE 4. ANAL-ONE
020250 WRITE PRINT-LINE FROM OUTPUT-FOUR-F BEFORE 2. ANAL-ONE
020300 WRITE PRINT-LINE FROM OUTPUT-FOUR-G BEFORE 2. ANAL-ONE
020350 PERFORM KS-CALC VARYING A FROM 1 BY 1 UNTIL A > 33. ANAL-ONE
020400 PERFORM JOINT-PROB. ANAL-ONE
020450 INITIALIZE-RT. ANAL-ONE
020500 MOVE LANG-KEY TO LANG-ID. ANAL-ONE
020550 MOVE ZEROS TO VC-TOTAL, A, B, N, F-CONS-CONS, F-VOW-CONS, ANAL-ONE
020600 MOVE F-CONS-VOW, F-VOW-VOW, FOL-FREQ-TOTAL, S, ILF-TOTAL, ANAL-ONE
020650 MOVE CRF-TOTAL, CARD-NR, L, M. ANAL-ONE
020700 PERFORM ZEROING-ONE VARYING A FROM 1 BY 1 UNTIL A > 33. ANAL-ONE
020750 PERFORM ZEROING-TWO VARYING A FROM 1 BY 1 UNTIL A > 33. ANAL-ONE
020800 PERFORM AFTER B FROM 1 BY 1 UNTIL B > 66. ANAL-ONE
020850 MOVE LANGUAGE TO LANG-DESIG, OUTPT-FOUR-LANG, ANAL-ONE
020900 MOVE OUTPT-FOUR-KS-LANG. ANAL-ONE
020950 READ POPULATION AT END GO TO DUMP. ANAL-ONE
021000 IND-TOTAL-CALC. ANAL-ONE
021050 COMPUTE ILF-TOTAL = ILF-TOTAL + ILF (A). ANAL-ONE
021100 IND-PROB-CALC. ANAL-ONE
021150 IF ILF (A) > 0 COMPUTE ILP (A) ROUNDED = ANAL-ONE
021200 IF ILF (A) / ILF-TOTAL, ELSE MOVE ZEROS TO ILP (A). ANAL-ONE
021250 CUTPT-ONE-PREP. ANAL-ONE
021300 MOVE SPACES TO PRINT-LINE. ANAL-ONE
021350 ANAL-ONE
021400 ANAL-ONE

```


023850	023900	023950	024000	024050	024100	024150	024200	024250	024300	024350	024400	024450	024500	024550	024600	024650	024700	024750	024800	024850	024900	024950	025000	025050	025100	025150	025200	025250	025300	025350	025400	025450	025500	025550	025600	025650	025700	025750	025800	025850	025900	025950	026000	026050	026100	026150	026200
--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

```

CCL-TITLES-C.
  COMPUTE L = A - 20.
  MOVE LTR (A) TO FIRST-LTR (L).
DEPN-THREE-CALC.
  MOVE ZEROS TO FOL-FREQ-TOTAL.
  PERFORM DEPN-TWO-SUM VARYING B FROM 34 BY 1 UNTIL B > 66.
  PERFORM DEPN-TWO-PRCB VARYING B FROM 34 BY 1 UNTIL B > 66.
CUTPT-THREE-PRCP.
  MOVE SPACES TO PRINT-LINE.
  WRITE PRINT-LINE FROM CUTPT-THREE-A AFTER TOP-OF-PAGE.
  WRITE PRINT-LINE FROM CUTPT-TWO-THREE-B BEFORE 1.
  WRITE PRINT-LINE FROM LANG-LINE BEFORE 0.
  WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2.
  WRITE PRINT-LINE FROM FIPST-LTR-OUTPT BEFORE 2.
VC-DUMP.
  COMPUTE VC-TOTAL = F-CCNS-CONS + F-VOW-CONS + F-CONS-VOW
  + F-VOW-VOW.
  IF F-VOW-VOW > 0 COMPUTE P-VOW-VOW ROUNDED = F-VOW-VOW /
  VC-TOTAL ELSE MOVE ZEROS TO P-VOW-VOW.
  IF F-CONS-VOW > 0 COMPUTE P-CONS-VOW ROUNDED = F-CONS-VOW /
  VC-TOTAL ELSE MOVE ZEROS TO P-CONS-VOW.
  IF F-VOW-CONS > 0 COMPUTE P-VOW-CONS ROUNDED = F-VOW-CONS /
  VC-TOTAL ELSE MOVE ZEROS TO P-VOW-CONS.
  IF F-CONS-CONS > 0 COMPUTE P-CONS-CONS ROUNDED =
  F-CONS-CONS / VC-TOTAL ELSE MOVE ZEROS TO P-CONS-CONS.
  WRITE PRINT-LINE FROM LANG-LINE AFTER TOP-OF-PAGE.
  WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2.
  WRITE PRINT-LINE FROM CUTPT-FOUR-A BEFORE 2.
  WRITE PRINT-LINE FROM CUTPT-FOUR-B BEFORE 2.
  WRITE PRINT-LINE FROM CUTPT-FOUR-C BEFORE 2.
  WRITE PRINT-LINE FROM CUTPT-FOUR-D BEFORE 3.
YULE-PRCP.
  COMPUTE S = S + (ILF (A) ** 2).
KKS-CALC.
  COMPUTE CRF-TOTAL = CRF-TOTAL + ILP (A).
  MOVE LTR (A) TO KS-LTR.
  MOVE CRF-TOTAL TO KS-CRF.
  WRITE PRINT-LINE FROM KS-LINE BEFORE 1.
E-1.
  MOVE 014300 TO ER-STEP.
  GO TO DUMP.
E-2.
  MOVE 014901 TO ER-STEP.
  GO TO DUMP.
E-3.
  MOVE 015400 TO ER-STEP.
  GO TO DUMP.
E-4.

```


007050	05	FILLER	PIC X(37)	VALUE SPACES:	ANAL-TWO
007100	05	FILLER	PIC X(61)	VALUE ALL:	ANAL-TWO
007150	05	FILLER-E.	PIC X(35)	VALUE SPACES:	ANAL-TWO
007200	01	OUTPUT-TWO	PIC X(40)	VALUE SPACES.	ANAL-TWO
007250	05	FILLER	PIC X(22)	VALUE SPACES.	ANAL-TWO
007300	05	FILLER-ONE	PIC ZZ,ZZ	VALUE SPACES.	ANAL-TWO
007350	05	FILLER-ONE	PIC X(19)	VALUE SPACES.	ANAL-TWO
007400	05	FILLER-ONE	PIC 9.9999	VALUE SPACES.	ANAL-TWO
007450	05	FILLER-ONE	PIC X(37)	VALUE SPACES.	ANAL-TWO
007500	01	OUTPUT-TWO	PIC X(38)	VALUE SPACES:	ANAL-TWO
007550	05	FILLER	PIC X(22)	VALUE SPACES:	ANAL-TWO
007600	05	FILLER	PIC ZZ,ZZ	VALUE TOTAL:	ANAL-TWO
007650	05	FILLER	PIC X(63)	VALUE SPACES.	ANAL-TWO
007700	05	FILLER	PIC X(46)	VALUE SPACES:	ANAL-TWO
007750	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
007800	01	OUTPUT-TWO	PIC X(46)	VALUE SPACES:	ANAL-TWO
007850	05	FILLER	PIC X(46)	VALUE SPACES:	ANAL-TWO
007900	05	FILLER	PIC X(40)	VALUE ALL:	ANAL-TWO
007950	05	FILLER	PIC X(47)	VALUE SPACES:	ANAL-TWO
008000	01	OUTPUT-TWO	PIC X(31)	VALUE SPACES:	ANAL-TWO
008050	05	FILLER	PIC X(102)	VALUE SPACES:	ANAL-TWO
008100	05	FILLER	PIC X(14)	VALUE SPACES:	ANAL-TWO
008150	05	FILLER	PIC X(9)	VALUE SPACES:	ANAL-TWO
008200	05	FILLER	PIC XX	VALUE SPACES:	ANAL-TWO
008250	01	OUTPUT-TWO	PIC X(5)	VALUE SPACES.	ANAL-TWO
008300	05	FILLER	PIC XX	VALUE SPACES:	ANAL-TWO
008350	05	FILLER	PIC XX	VALUE SPACES:	ANAL-TWO
008400	05	FILLER	PIC BBZZ	VALUE SPACES:	ANAL-TWO
008450	05	FILLER	PIC X(7)	VALUE SPACES:	ANAL-TWO
008500	01	OUTPUT-TWO	PIC X(5)	VALUE SPACES.	ANAL-TWO
008550	05	FILLER	PIC XX	VALUE SPACES:	ANAL-TWO
008600	05	FILLER	PIC BBZZ	VALUE SPACES:	ANAL-TWO
008650	05	FILLER	PIC X(7)	VALUE SPACES:	ANAL-TWO
008700	01	OUTPUT-TWO	PIC X(5)	VALUE SPACES.	ANAL-TWO
008750	05	FILLER	PIC XX	VALUE SPACES:	ANAL-TWO
008800	05	FILLER	PIC BBZZ	VALUE SPACES:	ANAL-TWO
008850	05	FILLER	PIC X(7)	VALUE SPACES:	ANAL-TWO
008900	01	OUTPUT-TWO	PIC X(46)	VALUE SPACES:	ANAL-TWO
008950	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
009000	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
009050	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
009100	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
009150	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
009200	01	OUTPUT-TWO	PIC X(87)	VALUE SPACES:	ANAL-TWO
009250	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
009300	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
009350	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO
009400	05	FILLER	PIC X(87)	VALUE SPACES:	ANAL-TWO

009450	01	OUTPT-FOUR-A.	PIC X(47) VALUE SPACES.	ANAL-TWO
009500		05 FILLER	PIC X(33) VALUE 'P(VOWEL/VOWEL)'.	ANAL-TWO
009550		05 FILLER	PIC 9.99999S.	ANAL-TWO
009600		05 P-VOW-VOW	PIC X(46) VALUE SPACES.	ANAL-TWO
009650		05 FILLER		ANAL-TWO
009700	C1	OUTPT-FOUR-B.	PIC X(47) VALUE SPACES.	ANAL-TWO
009750		05 FILLER	PIC X(33) VALUE 'P(CONSONANT/VOWEL)'.	ANAL-TWO
009800		05 FILLER		ANAL-TWO
009850		05 L).	PIC 9.99999S.	ANAL-TWO
009900		05 P-CONS-VOW	PIC X(46) VALUE SPACES.	ANAL-TWO
009950		05 FILLER		ANAL-TWO
010000	C1	OUTPT-FOUR-C.	PIC X(47) VALUE SPACES.	ANAL-TWO
010050		05 FILLER	PIC X(33) VALUE 'P(VOWEL/CONSONANT)'.	ANAL-TWO
010100		05 FILLER		ANAL-TWO
010150		05 T).	PIC 9.99999S.	ANAL-TWO
010200		05 P-VOW-CONS	PIC X(46) VALUE SPACES.	ANAL-TWO
010250		05 FILLER		ANAL-TWO
010300	01	OUTPT-FOUR-D.	PIC X(47) VALUE SPACES.	ANAL-TWO
010350		05 FILLER	PIC X(33) VALUE 'P(CONSONANT/CONS)	ANAL-TWO
010400		05 FILLER		ANAL-TWO
010450		05 ANANT).	PIC 9.99999S.	ANAL-TWO
010500		05 P-CONS-CONS	PIC X(46) VALUE SPACES.	ANAL-TWO
010550		05 FILLER		ANAL-TWO
010600	01	OUTPT-FOUR-E.	PIC X(11) VALUE SPACES.	ANAL-TWO
010650		05 FILLER	PIC X(32) VALUE 'YULE-S CHARACTER	ANAL-TWO
010700		05 FILLER		ANAL-TWO
010750		05 STIC, K, FOR:	PIC X(30) VALUE SPACES.	ANAL-TWO
010800		05 OUTPT-FOUR-LANG	PIC Z(12) 9.999.	ANAL-TWO
010850		05 YULES-K	PIC X(40) VALUE SPACES.	ANAL-TWO
010900		05 FILLER		ANAL-TWO
010950	01	OUTPT-FOUR-F.	PIC X.	ANAL-TWO
011000		05 FILLER	PIC X(94) VALUE 'KOLMOGOROV-SMIRN	ANAL-TWO
011050		05 FILLER	TABLE FOR INDEPENDENT LETTER	ANAL-TWO
011100		05 CUMULATED RELATIVE FREQUENCY		ANAL-TWO
011150		05 IN:.		ANAL-TWO
011200		05 CC OUTPUT-FOUR-KS-LANG		ANAL-TWO
011250		05 FILLER	PIC X(30) VALUE SPACES.	ANAL-TWO
011300	C1	OUTPT-FOUR-G.		ANAL-TWO
011350		05 FILLER	PIC X(37) VALUE SPACES.	ANAL-TWO
011400		05 FILLER	PIC X(25) VALUE 'LETTER'.	ANAL-TWO
011450		05 FILLER	PIC X(71) VALUE 'CUM. REL. FREQ'.	ANAL-TWO
011500	C1	XS-LINE.		ANAL-TWO
011550		05 FILLER	PIC X(40) VALUE SPACES.	ANAL-TWO
011600		05 FILLER		ANAL-TWO
011650		05 FILLER	PIC X(25) VALUE SPACES.	ANAL-TWO
011700		05 FILLER	PIC 9.99999S.	ANAL-TWO
011750		05 FILLER		ANAL-TWO
011800	01	ERROR-LINE.	PIC X(60) VALUE SPACES.	ANAL-TWO

011850	05	FILLER	PIC X(28)	VALUE	'OVERFLOW ERROR	ANAL-TWO
011900	05	FILLER	PIC X(6)	VALUE	' CARD IN PROCESS	ANAL-TWO
011950	05	STATEMENT	PIC X(21)	VALUE	'	ANAL-TWO
012000	05	STEP	PIC X(6)	VALUE	'	ANAL-TWO
012050	05	FILLER	PIC X(71)	VALUE	ALL '*'.	ANAL-TWO
012100	05	WAS	PIC X,	OCCURS 73 TIMES.		ANAL-TWO
012150	05	ER-CARD	PIC X(49)	VALUE SPACES.		ANAL-TWO
012200	05	FILLER	PIC X(37)	VALUE 'JOINT' PROBABILITY		ANAL-TWO
012250	01	DATA-FORM.	PIC X(47)	VALUE SPACES.		ANAL-TWO
012300	01	OUTPT-FIVE-A.	PIC X(11)	VALUE SPACES.		ANAL-TWO
012350	05	FILLER	PIC X(25)	VALUE 'YULES' K FOR X(1)		ANAL-TWO
012400	05	FILLER	PIC Z(12)9.999.	VALUE SPACES.		ANAL-TWO
012450	05	TABLE FOR X(1)-X(2)'	PIC X(80)	VALUE SPACES.		ANAL-TWO
012500	05	FILLER	PIC X(49)	VALUE SPACES.		ANAL-TWO
012550	05	OUTPT-FIVE-B.	PIC X(37)	VALUE 'JOINT' PROBABILITY		ANAL-TWO
012600	05	FILLER	PIC X(47)	VALUE SPACES.		ANAL-TWO
012650	05	FILLER	PIC X(11)	VALUE SPACES.		ANAL-TWO
012700	05	FILLER	PIC X(25)	VALUE 'YULES' K FOR X(1)		ANAL-TWO
012750	05	X(2)';	PIC Z(12)9.999.	VALUE SPACES.		ANAL-TWO
012800	05	YULES-K-2	PIC X(80)	VALUE SPACES.		ANAL-TWO
012850	05	FILLER	PIC X(49)	VALUE SPACES.		ANAL-TWO
012900	05	OUTPT-SIX-A.	PIC X(37)	VALUE 'JOINT' PROBABILITY		ANAL-TWO
012950	05	FILLER	PIC X(47)	VALUE SPACES.		ANAL-TWO
013000	05	FILLER	PIC X(11)	VALUE SPACES.		ANAL-TWO
013050	05	TABLE FOR X(1)-X(3)'	PIC X(25)	VALUE 'YULES' K FOR X(1)		ANAL-TWO
013100	05	FILLER	PIC Z(12)9.999.	VALUE SPACES.		ANAL-TWO
013150	05	OUTPT-SIX-B.	PIC X(80)	VALUE SPACES.		ANAL-TWO
013200	05	FILLER	PIC X,			ANAL-TWO
013250	05	FILLER	PIC X(10).			ANAL-TWO
013300	05	X(3)';	PIC X(9)	OCCURS 11 TIMES.		ANAL-TWO
013350	05	YULES-K-3	PIC X(10).			ANAL-TWO
013400	05	FILLER	PIC X(10).			ANAL-TWO
013450	05	TEST	PIC X(9)	OCCURS 11 TIMES.		ANAL-TWO
013500	05	SEQ-NR	PIC X(10).			ANAL-TWO
013550	05	LANG-ID	PIC X(10).			ANAL-TWO
013600	05	LANG-ID	PIC X(10).			ANAL-TWO
013650	05	LANG-ID	PIC X(10).			ANAL-TWO
013700	05	LANG-ID	PIC X(10).			ANAL-TWO
013750	05	LANG-ID	PIC X(10).			ANAL-TWO
013800	05	LANG-ID	PIC X(10).			ANAL-TWO
013850	05	LANG-ID	PIC X(10).			ANAL-TWO
013900	05	LANG-ID	PIC X(10).			ANAL-TWO
013950	05	LANG-ID	PIC X(10).			ANAL-TWO
014000	05	LANG-ID	PIC X(10).			ANAL-TWO
014050	05	LANG-ID	PIC X(10).			ANAL-TWO
014100	05	LANG-ID	PIC X(10).			ANAL-TWO
014150	05	LANG-ID	PIC X(10).			ANAL-TWO
014200	05	LANG-ID	PIC X(10).			ANAL-TWO


```

016650 IF MATCH = 1 GO TO DELTA, ELSE ADD 1 TO N.
016700 CK-TWO-BRAVO.
016750 ADD 1 TO N.
016800 GO TO CK-TWO-ALPHA.
016850 CK-THREE.
016900 TO FOL-FREQ (A, 68) ON SIZE ERROR GO TO E-6.
016950 GO TO ECHO.
017000 SHIFT-CNE.
017050 PERFORM READ-A-CARD.
017100 MOVE SPACES TO CL (1), CL (2).
017150 MOVE 3 TO N.
017200 SHIFT-TWO.
017250 MOVE CL (72) TO CL (2).
017300 PERFORM READ-A-CARD.
017350 MOVE 2 TO N.
017400 MOVE 3 TO M.
017450 SHIFT-THREE.
017500 MOVE CL (72) TO CL (2).
017550 PERFORM READ-A-CARD.
017600 MOVE 1 TO N.
017650 MOVE 3 TO L.
017700 ALT-THREE.
017750 MOVE CL (72) TO CL (1).
017800 MOVE SPACES TO CL (2).
017850 PERFORM READ-A-CARD.
017900 MOVE 0 TO N.
017950 MOVE 2 TO L.
018000 READ-A-CARD.
018050 READ POPULATION AT END GO TO EOJ.
018100 IF NEW-LANG GO TO FLUSH.
018150 ADD 1 TO CARD-NR ON SIZE ERROR GO TO E-7.
018200 PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.
018250 DAT-SHIFT.
018300 COMPUTE Y = X + 2.
018350 MOVE CC (X) TO CL (Y).
018400 FLUSH.
018450 IF CARD-NR NOT = ZERO PERFORM CLOSE-OUT-RT.
018500 PERFORM INITIALIZE-RT.
018550 GO TO ALPHA.
018600 EOJ.
018650 PERFORM CLOSE-OUT-RT.
018700 PERFORM LOOP-RT 5 TIMES.
018750 MOVE SPACES TO PRINT-LINE.
018800 WRITE PRINT-LINE AFTER TOP-OF-PAGE.
018850 CLOSE POPULATION, LANG-ANAL, CARD-FILE.
018900 STOP RUN.
018950 LCOP-RT.
019000 MOVE SPACES TO PRINT-LINE.
019050

```


[illegible]


```

028650 MOVE ZEROS TO FOL-FREQ-TOTAL.
028670 PERFORM DEPNTWO-SUM VARYING B FROM 35 BY 1 UNTIL B > 68.
028700 PERFORM DEPNTWO-PROB VARYING B FROM 35 BY 1 UNTIL B > 68.
028750 DEPNTWO-SUM.
028800 COMPUTE FOL-FREQ-TOTAL = FCL-FREQ-TOTAL + FOL-FREQ (A, B).
028830 DEPNTWO-PROB.
028890 IF FOL-FREQ (A, B) > 0 COMPUTE FOL-PROB (A, B) ROUNDED =
028920 (FOL-FREQ (A, B) / FOL-FREQ-TOTAL), ELSE MOVE ZEROS
028950 TO FOL-PROB (A, B).
028980 OUTPT-CNE-SPACES TO PRINT-LINE.
029010 WRITE PRINT-LINE FROM OUTPT-CNE-A AFTER TOP-OF-PAGE.
029040 WRITE PRINT-LINE FROM OUTPT-CNE-B BEFORE 1.
029070 WRITE PRINT-LINE FROM LANG-LINE AFTER 1.
029100 WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 1.
029130 WRITE PRINT-LINE FROM OUTPT-CNE-C AFTER 2.
029160 WRITE PRINT-LINE FROM OUTPT-CNE-D BEFORE 2.
029190 OUTPT-CNE-PRINT.
029220 MOVE LTR (A) TO LTR-CNE.
029250 MOVE ILF (A) TO FREQ-CNE.
029280 MOVE ILF (A) TO PROB-CNE.
029310 WRITE PRINT-LINE FROM OUTPT-CNE-E BEFORE 2.
029340 OUTPT-TWO-SPACES TO PRINT-LINE.
029370 WRITE PRINT-LINE FROM OUTPT-TWO-A AFTER TOP-OF-PAGE.
029400 WRITE PRINT-LINE FROM OUTPT-TWO-B BEFORE 1.
029430 WRITE PRINT-LINE FROM LANG-LINE BEFORE 0.
029460 WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2.
029490 WRITE PRINT-LINE FROM FIRST-LTR-OUTPT BEFORE 2.
029520 CCL-TITLE-STA.
029550 LTR (A) TO FIRST-LTR (A).
029580 OUTPT-TWO-DUMP-A.
029610 PERFORM DUMP-A VARYING A FROM 1 BY 1 UNTIL A > 13.
029640 WRITE PRINT-LINE FROM DEPNTWO-FREQ-LINE BEFORE 1.
029670 PERFORM DUMP-B VARYING A FROM 1 BY 1 UNTIL A > 13.
029700 WRITE PRINT-LINE FROM DEPNTWO-PROB-LINE BEFORE 2.
029730 OUTPT-TWO-DUMP-B.
029760 PERFORM DUMP-A VARYING A FROM 14 BY 1 UNTIL A > 26.
029790 WRITE PRINT-LINE FROM DEPNTWO-FREQ-LINE BEFORE 1.
029820 PERFORM DUMP-B VARYING A FROM 14 BY 1 UNTIL A > 26.
029850 WRITE PRINT-LINE FROM DEPNTWO-PROB-LINE BEFORE 2.
029880 OUTPT-TWO-DUMP-C.
029910 PERFORM DUMP-A VARYING A FROM 27 BY 1 UNTIL A > 34.
029940 WRITE PRINT-LINE FROM DEPNTWO-FREQ-LINE BEFORE 1.
029970 PERFORM DUMP-B VARYING A FROM 27 BY 1 UNTIL A > 34.
030000 WRITE PRINT-LINE FROM DEPNTWO-PROB-LINE BEFORE 2.
030030 CCL-TITLE-S-B.
030060 L = A - 13.

```



```

033450      TO DEPN-PRCB (M).
033500      IF A > 26 COMPUTE M = A - 26, MOVE FOL-PRCB (A, B)
033550      ZEROING-ONE.
033600      MOVE ZEROS TO ILF (A).
033650      MOVE ZEROS TO ILP (A).
033700      FLUSHING.
033750      MOVE SPACES TO FIRST-LTR (A).
033800      MOVE ZEROS TO DEPN-FREQ (A), DEPN-PRCB (A).
033850      ZEROING-TWO.
033900      MOVE ZEROS TO FOL-FREQ (A, B).
033950      MOVE ZEROS TO FOL-PRCB (A, B).
034000      CUTPT-FIVE-PRCP.
034050      WRITE PRINT-LINE FROM CUTPT-FIVE-A AFTER TOP-OF-PAGE.
034100      WRITE PRINT-LINE FROM CUTPT-TWO-THREE-B BEFORE 1.
034150      WRITE PRINT-LINE FROM LANG-LINE BEFORE 1.
034200      WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2.
034250      WRITE PRINT-LINE FROM FIRST-LTR-OUTPT BEFORE 2.
034300      CUTPT-DUMP-D.
034350      PERFORM DUMP-B VARYING A FROM 1 BY 1 UNTIL A > 13.
034400      WRITE PRINT-LINE FROM DEPN-PRCB-LINE BEFORE 2.
034450      CUTPT-DUMP-E.
034500      PERFORM DUMP-B VARYING A FROM 14 BY 1 UNTIL A > 26.
034550      WRITE PRINT-LINE FROM DEPN-PRCB-LINE BEFORE 2.
034600      CUTPT-DUMP-F.
034650      PERFORM DUMP-B VARYING A FROM 27 BY 1 UNTIL A > 34.
034700      WRITE PRINT-LINE FROM DEPN-PRCB-LINE BEFORE 2.
034750      CUTPT-SIX-PRCP.
034800      WRITE PRINT-LINE FROM CUTPT-SIX-A AFTER TOP-OF-PAGE.
034850      WRITE PRINT-LINE FROM CUTPT-TWO-THREE-B BEFORE 1.
034900      WRITE PRINT-LINE FROM LANG-LINE BEFORE 1.
034950      WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2.
035000      WRITE PRINT-LINE FROM FIRST-LTR-OUTPT BEFORE 2.
035050      YULES-K-PRCP.
035100      COMPUTE S = S + (FOL-FREQ (A, B) ** 2).
035150      CARD-DUMP.
035200      COMPUTE CRF-TOTAL = CRF-TOTAL + FOL-PRCB (A, B).
035250      ADD 1 TO L.
035300      MOVE CRF-TOTAL TO CRF-OUT (L).
035350      IF WS-SEQ-NR = 105, ADD 1 TO WS-SEQ-NR.
035400      MOVE WS-SEQ-NR TO WS-SEQ-NR.
035450      MOVE CRF-TOTAL TO L-SEQ-NR.
035500      WRITE CARD-OUT FROM LAST-CARD-FILL.
035550      IF L = 11 ADD 1 TO WS-SEQ-NR.
035600      MOVE WS-SEQ-NR TO SEQ-NR.
035650      MOVE ZEROS TO L.
035700      WRITE CARD-OUT FROM CARD-FILL.
035750      E-1.
035800

```

```

035850      MOVE 016450  TO ER-STEP.
035900      GO TO DUMP.
E-2.      MOVE 016850  TO ER-STEP.
035950      GO TO DUMP.
E-3.      MOVE 017300  TO ER-STEP.
036000      GO TO DUMP.
E-4.      MOVE 017550  TO ER-STEP.
036100      GO TO DUMP.
E-5.      MOVE 018200  TO ER-STEP.
036150      GO TO DUMP.
E-6.      MOVE 018850  TO ER-STEP.
036200      GO TO DUMP.
E-7.      MOVE 019800  TO ER-STEP.
036250      GO TO DUMP.
E-8.      MOVE 022450  TO ER-STEP.
036300      GO TO DUMP.
E-9.      MOVE 022500  TO ER-STEP.
036350      GO TO DUMP.
E-10.     MOVE 022600  TO ER-STEP.
036400      GO TO DUMP.
E-11.     MOVE 022700  TO ER-STEP.
036450      GO TO DUMP.
DUMP.     MOVE CARD-NR TO ER-CARD E
           WRITE PRINT-LINE FROM E
           GO TO EQJ.
037600

```

APPENDIX I. LANGUAGE IDENTIFICATION PROGRAM (V-1)

```

000050 IDENTIFICATION DIVISION.
000100 PROGRAM-ID. LANGUAGE-ID-VERSION-ONE.
000150 AUTHOR. LCDR. MORTON D. RAU, USN.
000200 REMARKS. THIS PROGRAM IS PART OF MY MASTER OF SCIENCE IN SAMPLE
000250 MANAGEMENT THESIS AND IS DESIGNED TO IDENTIFY SHORT THE
000300 TEXT OF A LANGUAGE AS EITHER ENGLISH OR SPANISH. THE
000350 DECISION IS BASED ON STATISTICAL CHARACTERISTICS OF LETTERS
000400 WITHIN INDEPENDENT WORDS WITH NO CONSIDERATION GIVEN TO
000450 SPACES.
000500 INPUT DATA MUST APPEAR IN THE FOLLOWING ORDER:
000550 3 ENGLISH A CARDS - ICRF TABLE VALUES
000600 99 ENGLISH B CARDS - DCRF TABLE VALUES X(2)/X(1)
000650 3 ENGLISH C CARDS - DCRF TABLE VALUES X(3)/X(1)
000700 99 SPANISH A CARDS - ICRF TABLE VALUES
000750 99 SPANISH B CARDS - DCRF TABLE VALUES X(2)/X(1)
000800 99 SPANISH C CARDS - DCRF TABLE VALUES X(3)/X(1)
000850 1 HEAD CARD PRECEDING EACH GROUP OF SAMPLE CARDS.
000900 MAXIMUM CHARACTERS, NOT COUNTING SPACES.
000950 INDIVIDUAL DIVISION.
001000 ENVIRONMENT SECTION.
001050 CONFIGURATION. IBM-360-67.
001100 OBJECT-COMPUTER. IBM-360-67.
001150 SPECIAL-NAMES.
001200 COI IS PAGE-TCP.
001250 INPUT-OUTPUT SECTION.
001300 FILE-CONTROL. SAMPLE ASSIGN TO UR-S-INI.
001350 FILE-SELECT. RESULT ASSIGN TO UR-S-OUT1.
001400 DATA SECTION.
001450 FILE SELECTION.
001500 FD SAMPLE RECORDS ARE OMITTED
001550 LABEL CONTAINS 5 RECORDS
001600 BLOCK CONTAINS 5 RECORDS
001650 DATA RECORD IS SAMPLE-CARD, HEAD-CARD, DATA-CARD.
001700 SAMPLE FILLER.
001750 01 05 FILLER-FLD.
001800 01 05 FILLER
001850 01 05 FILLER
001900 01 05 FILLER
001950 01 05 FILLER
002000 01 05 FILLER
002050 01 05 FILLER
002100 01 05 FILLER
002150 01 05 FILLER
002200 01 05 FILLER

```

002250	05	TEST-NAME	PIC	X(20):	E
002300	05	FILLER	PIC	X(58):	E
002350	05	DATA-CARD:	PIC	A	E
002400	05	DC-ID	PIC	999	E
002450	05	DC-SEQ	PIC	X(10):	E
002500	05	DC-LANG	PIC	9V9(5), OCCURS 11 TIMES.	E
002550	05	CRF-IN			E
002600	FD	RESULT RECORDS ARE OMITTED			E
002650		LABEL CONTAINS 5 RECORDS			E
002700		DATA RECORD IS PRINT-LINE			E
002750		PRINT-LINE	PIC	X(133):	E
002800	01	WORKING-STORAGE SECTION.			E
002850	77	A	PIC	99, COMP, SYNC:	E
002900	77	B	PIC	99, COMP, SYNC:	E
002950	77	C	PIC	99, COMP, SYNC:	E
003000	77	D	PIC	9(12), COMP, SYNC:	E
003050	77	E	PIC	99, COMP, SYNC:	E
003100	77	F	PIC	99, COMP, SYNC:	E
003150	77	G	PIC	99, COMP, SYNC:	E
003200	77	H	PIC	99, COMP, SYNC:	E
003250	77	I	PIC	9(4):	E
003300	77	J	PIC	9(4), COMP, SYNC:	E
003350	77	K	PIC	9(4), COMP, SYNC:	E
003400	77	L	PIC	9(4), COMP, SYNC:	E
003450	77	M	PIC	9(4), COMP, SYNC:	E
003500	77	N	PIC	9(4), COMP, SYNC:	E
003550	77	O	PIC	9(4), COMP, SYNC:	E
003600	77	P	PIC	9(4), COMP, SYNC:	E
003650	77	Q	PIC	9(3):	E
003700	77	R	PIC	9(13)V999	E
003750	77	S	PIC	9(15)V999 COMP, SYNC:	E
003800	77	T	PIC	9(15)V999 COMP, SYNC:	E
003850	77	U	PIC	99V999 COMP, SYNC:	E
003900	77	V	PIC	99V99(12) COMP, SYNC:	E
003950	77	W	PIC	99V99(12) COMP, SYNC:	E
004000	77	X	PIC	99V99(6) COMP, SYNC:	E
004050	77	Y	PIC	9V9(5), COMP, SYNC:	E
004100	77	Z	PIC	9V9(5), COMP, SYNC:	E
004150	77	ENG-D	PIC	9V9(5), COMP, SYNC:	E
004200	77	SPAN-D	PIC	9V9(5), COMP, SYNC:	E
004250	77	SAMPLE-SIZE	PIC	9(4), COMP, SYNC:	E
004300	77	ET	PIC	99V99 COMP, SYNC:	E
004350	77	FFI2	PIC	9(4):	E
004400	77	LTR-TABLE-VALUES.			E
004450	01	FILLER	PIC	X VALUE	E
004500	05	FILLER	PIC	X VALUE	E
004550	05	FILLER	PIC	X VALUE	E
004600					E
004650					E
004700					E
004750					E
004800					E
004850					E
004900					E
004950					E
005000					E
005050					E
005100					E
005150					E
005200					E
005250					E
005300					E
005350					E
005400					E
005450					E
005500					E
005550					E
005600					E
005650					E
005700					E
005750					E
005800					E
005850					E
005900					E
005950					E
006000					E
006050					E
006100					E
006150					E
006200					E
006250					E
006300					E
006350					E
006400					E
006450					E
006500					E
006550					E
006600					E
006650					E
006700					E
006750					E
006800					E
006850					E
006900					E
006950					E
007000					E

007050	01	05	ENG-VC-CRF	PIC	9V9(5)	OCCURS 4 TIMES.	ONE
007100		05	SPAN-VC-CRF-TABLE-VALUES.	PIC	9V9(5)	VALUE 0.06163.	ONE
007150		05	FILLER	PIC	9V9(5)	VALUE 0.43293.	ONE
007200		05	FILLER	PIC	9V9(5)	VALUE 0.87982.	ONE
007250		05	FILLER	PIC	9V9(5)	VALUE 1.00000.	ONE
007300	01	05	SPAN-VC-CRF-TABLE REDEFINES.	PIC	SPAN-VC-CRF-TABLE-VALUES.		ONE
007350		05	SPAN-VC-CRF	PIC	9V9(5)	OCCURS 4 TIMES.	ONE
007400	01	05	IND-FREQ-TABLE.	PIC	9(4)	COMP SYNC OCCURS 33 TIMES.	ONE
007450		05	ILF	OCCURS 23 TIMES.			ONE
007500	01	05	DEPENDENCY-FREQ-TABLE.	PIC	9(4)	COMP SYNC OCCURS 66 TIMES.	ONE
007550		05	FST-LTR	OCCURS 33 TIMES.			ONE
007600		05	IO FOL-FREQ	PIC	9V9(5)	COMP SYNC OCCURS 66 TIMES.	ONE
007650	01	05	DEPENDENCY-PROB-TABLE.	OCCURS 33 TIMES.			ONE
007700		05	GIVEN-LTR	PIC	9V9(5)	COMP SYNC OCCURS 66 TIMES.	ONE
007750		05	IO FOL-PRCB	PIC	9(6)	VALUE 999999.	ONE
007800	01	05	ALFA-TABLE-VALUES.	PIC	9(6)	VALUE 999996.	ONE
007850		05	FILLER	PIC	9(6)	VALUE 999991.	ONE
007900		05	FILLER	PIC	9(6)	VALUE 999979.	ONE
007950		05	FILLER	PIC	9(6)	VALUE 999954.	ONE
008000		05	FILLER	PIC	9(6)	VALUE 999909.	ONE
008050		05	FILLER	PIC	9(6)	VALUE 999829.	ONE
008100		05	FILLER	PIC	9(6)	VALUE 999697.	ONE
008150		05	FILLER	PIC	9(6)	VALUE 999548.	ONE
008200		05	FILLER	PIC	9(6)	VALUE 999174.	ONE
008250		05	FILLER	PIC	9(6)	VALUE 998715.	ONE
008300		05	FILLER	PIC	9(6)	VALUE 998071.	ONE
008350		05	FILLER	PIC	9(6)	VALUE 997192.	ONE
008400		05	FILLER	PIC	9(6)	VALUE 996028.	ONE
008450		05	FILLER	PIC	9(6)	VALUE 994524.	ONE
008500		05	FILLER	PIC	9(6)	VALUE 992623.	ONE
008550		05	FILLER	PIC	9(6)	VALUE 990270.	ONE
008600		05	FILLER	PIC	9(6)	VALUE 987410.	ONE
008650		05	FILLER	PIC	9(6)	VALUE 975995.	ONE
008700		05	FILLER	PIC	9(6)	VALUE 975978.	ONE
008750		05	FILLER	PIC	9(6)	VALUE 975318.	ONE
008800		05	FILLER	PIC	9(6)	VALUE 969983.	ONE
008850		05	FILLER	PIC	9(6)	VALUE 963945.	ONE
008900		05	FILLER	PIC	9(6)	VALUE 957186.	ONE
008950		05	FILLER	PIC	9(6)	VALUE 949694.	ONE
009000		05	FILLER	PIC	9(6)	VALUE 941466.	ONE
009050		05	FILLER	PIC	9(6)	VALUE 932503.	ONE
009100		05	FILLER	PIC	9(6)	VALUE 922517.	ONE
009150		05	FILLER	PIC	9(6)	VALUE 912423.	ONE
009200		05	FILLER	PIC	9(6)	VALUE 901344.	ONE
009250		05	FILLER	PIC	9(6)	VALUE 889605.	ONE
009300		05	FILLER				ONE
009350		05	FILLER				ONE
009400		05	FILLER				ONE
009450		05	FILLER				ONE
009500		05	FILLER				ONE
009550		05	FILLER				ONE
009600		05	FILLER				ONE
009650		05	FILLER				ONE
009700		05	FILLER				ONE
009750		05	FILLER				ONE
009800		05	FILLER				ONE
009850		05	FILLER				ONE
009900		05	FILLER				ONE
009950		05	FILLER				ONE
010000		05	FILLER				ONE

016650	01	05	FILLER	PIC X(18)	VALUE ALL ' '.	IDEN-ONE
016700		05	FILLER	PIC X(26)	VALUE SPACES.	IDEN-ONE
016750		05	LINE-ONE.	PIC X(24)	VALUE SPACES.	IDEN-ONE
016800		05	FILLER	PIC X(18)	'IDENTIFIED BY YULE'.	IDEN-ONE
016850		05	FILLER	PIC X	VALUE QUOTE.	IDEN-ONE
016900		05	FILLER	PIC X(30)	'S K FOR SINGLE LETTER	IDEN-ONE
016950		05	AS:.	PIC X(30)	VALUE SPACES.	IDEN-ONE
017000		05	BANK-ONE	PIC X(30)	VALUE SPACES.	IDEN-ONE
017050		05	FILLER	PIC X(24)	VALUE SPACES.	IDEN-ONE
017100	01	05	LINE-TWO.	PIC X(18)	'IDENTIFIED BY YULE'.	IDEN-ONE
017150		05	FILLER	PIC X	VALUE QUOTE.	IDEN-ONE
017200		05	FILLER	PIC X(30)	'S K FOR X(2)/X(1) AS	IDEN-ONE
017250		05	FILLER	PIC X(30)	VALUE SPACES.	IDEN-ONE
017300		05	FILLER	PIC X(24)	VALUE SPACES.	IDEN-ONE
017350		05	FILLER	PIC X(18)	'IDENTIFIED BY YULE'.	IDEN-ONE
017400		05	FILLER	PIC X	VALUE QUOTE.	IDEN-ONE
017450		05	BANK-TWO	PIC X(30)	'S K FOR X(3)/X(1) AS	IDEN-ONE
017500		05	LINE-THREE.	PIC X(30)	VALUE SPACES.	IDEN-ONE
017550	01	05	FILLER	PIC X(24)	VALUE SPACES.	IDEN-ONE
017600		05	FILLER	PIC X(18)	'IDENTIFIED BY YULE'.	IDEN-ONE
017650		05	FILLER	PIC X	VALUE QUOTE.	IDEN-ONE
017700		05	FILLER	PIC X(30)	'S K FOR X(3)/X(1) AS	IDEN-ONE
017750		05	FILLER	PIC X(30)	VALUE SPACES.	IDEN-ONE
017800		05	BANK-THREE	PIC X(30)	VALUE SPACES.	IDEN-ONE
017850		05	LINE-FOUR.	PIC X(24)	VALUE SPACES.	IDEN-ONE
017900	01	05	FILLER	PIC X(18)	'IDENTIFIED BY YULE'.	IDEN-ONE
017950		05	FILLER	PIC X	VALUE QUOTE.	IDEN-ONE
018000		05	FILLER	PIC X(30)	'S K FOR X(3)/X(1) AS	IDEN-ONE
018050		05	FILLER	PIC X(30)	VALUE SPACES.	IDEN-ONE
018100		05	BANK-FOUR	PIC X(30)	VALUE SPACES.	IDEN-ONE
018150		05	LINE-FIVE.	PIC X(24)	VALUE SPACES.	IDEN-ONE
018200	01	05	FILLER	PIC X(18)	'IDENTIFIED BY K-S TE	IDEN-ONE
018250		05	FILLER	PIC X(49)	VALUE SPACES.	IDEN-ONE
018300		05	FILLER	PIC X	VALUE SPACES.	IDEN-ONE
018350		05	FILLER	PIC X(30)	VALUE SPACES.	IDEN-ONE
018400		05	BANK-FIVE	PIC X(24)	VALUE SPACES.	IDEN-ONE
018450	01	05	LINE-SIX.	PIC X(49)	VALUE SPACES.	IDEN-ONE
018500		05	FILLER	PIC X(30)	VALUE SPACES.	IDEN-ONE
018550		05	FILLER	PIC X(24)	VALUE SPACES.	IDEN-ONE
018600		05	FILLER	PIC X(49)	VALUE SPACES.	IDEN-ONE
018650		05	FILLER	PIC X	VALUE SPACES.	IDEN-ONE
018700		05	FILLER	PIC X(30)	VALUE SPACES.	IDEN-ONE
018750	01	05	BANK-SIX	PIC X(30)	VALUE SPACES.	IDEN-ONE
018800		05	LINE-SEVEN.	PIC X(24)	VALUE SPACES.	IDEN-ONE
018850		05	FILLER	PIC X(49)	VALUE SPACES.	IDEN-ONE
018900		05	FILLER	PIC X	VALUE SPACES.	IDEN-ONE
018950		05	FILLER	PIC X(30)	VALUE SPACES.	IDEN-ONE
019000		05	FILLER	PIC X(24)	VALUE SPACES.	IDEN-ONE
		05	ST FOR X(3)/X(1) AS:.	PIC X(49)	VALUE SPACES.	IDEN-ONE

```

019050 05 BANK-SEVEN
019100 05 FILLER
019150 05 BANK-FORM-LANG
019200 05 FILLER
019250 05 BANK-EQUAL
019300 05 BANK-ALFA
019350 05 ERROR-LINE.
019400 05 FILLER
019450 05 FILLER
019500 05 ERROR-ANS
019550 05 FILLER
019600 05 DATA-FORM.
019650 05 DATA-CC
019700 05 IND-PRGB-TABLE.
019750 05 VC-CRF-TABLE.
019800 05 VC-CRF
019850 05 LINE-EIGHT.
019900 05 FILLER
020000 05 FILLER
020050 05 FILLER
020100 05 BANK-EIGHT
020150 05 FILLER
020200 05 AINED-CN.
020250 05 O-CARD-NR
020300 05 FILLER
020350 05 FILE DIVISION.
020400 05 PRCCEDURE
020450 05 START-UP.
020500 05 INPUT SAMPLE, OUTPUT RESULT.
020550 05 MOVE 1 TO A. CARD-NR.
020600 05 PERFORM ENG-FILL-1 33 TIMES.
020650 05 MOVE 1 TO A. B. CARD-NR.
020700 05 PERFORM ENG-FILL-2 1089 TIMES.
020750 05 MOVE 1 TO A.
020800 05 MOVE 34 TO B. CARD-NR.
020850 05 MOVE 0 TO L. CARD-NR.
020900 05 PERFORM ENG-FILL-3 1089 TIMES.
020950 05 MOVE 1 TO A.
021000 05 MOVE 0 TO L. CARD-NR.
021050 05 PERFORM SPAN-FILL-1 33 TIMES.
021100 05 MOVE 1 TO A. B. CARD-NR.
021150 05 MOVE 0 TO L.
021200 05 PERFORM SPAN-FILL-2 1089 TIMES.
021250 05 MOVE 1 TO A.
021300 05 MOVE 34 TO B.
021350 05 MOVE 0 TO L. CARD-NR.
021400

PIC X(30). VALUE SPACES.
PIC X(30). VALUE SPACES.
PIC X(10). VALUE SPACES.
PIC X(10). VALUE SPACES.
PIC X(10). VALUE SPACES.
PIC X(10). VALUE SPACES.
PIC X(21) VALUE ALL '*'.
PIC X(20) VALUE 'OVERFLOW ERROR ON:'.
PIC X(20). VALUE ALL '*'.
PIC X(72) VALUE ALL '*'.
PIC X, OCCURS 73 TIMES.
PIC 9V9(5) COMP SYNC OCCURS 33 TIMES.
PIC 9V9(5) COMP SYNC OCCURS 4 TIMES.
PIC X.
PIC X(14) VALUE 'SAMPLE SIZE = '.
PIC Z129.
PIC X(29) VALUE ' SINGLE LETTERS CONT
PIC Z19.
PIC X(62) VALUE ' CARDS.'.

```

```

021450 PERFORM SPAN-FILL-3 1089 TIMES.
021500 MOVE ZEROS TO CARD-NR.
021550 BEGIN.
021600 READ SAMPLE AT END GO TO ECJ.
021650 IF NEXT-SAMPLE PERFORM CLOSE-CUT-RT THPU INITIALIZE-RT.
021700 PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.
021750 ADD 1 TO CARD-NR CN SIZE ERROR GO TO ALERT.
021800 MOVE 3 TO N.
021850 DATA-COLL.
021900 IF CC (N) = SPACE GO TO BLANK-COL-STEP.
021950 MOVE 1 TO A.
022000 CK-ONE.
022050 IF CC (N) = LTR (A) GO TO PROC-TWO.
022100 IF A = 33 GO TO COL-STEP.
022150 ADD 1 TO A.
022200 GO TO CK-ONE.
022250 PROC-TWO.
022300 ADD 1 TO I IF (A) ON SIZE ERROR GO TO ALERT.
022350 COMPUTE M = N + 1.
022400 IF CC (M) = SPACE GO TO COL-STEP.
022450 IF CC (M) = '-' AND M = 73 PERFORM CARD-STEP-ONE.
022500 MOVE 1 TO B.
022550 CK-TWO.
022600 IF CC (M) = LTR (B) GO TO PROC-THREE.
022650 IF B = 33 GO TO COL-STEP.
022700 ADD 1 TO B.
022750 GO TO CK-TWO.
022800 PROC-THREE.
022850 ADD 1 TO I.
022900 IF A < 11 AND B < 11 ADD 1 TO F-VOW-VOW CN SIZE.
022950 IF A ERROR GO TO ALERT.
023000 IF A < 11 AND B NOT < 11 ADD 1 TO F-CONS-VOW ON SIZE.
023050 IF A ERROR GO TO ALERT.
023100 IF A NOT < 11 AND B < 11 ADD 1 TO F-VOW-CONS ON SIZE.
023150 IF A NOT < 11 AND B NOT < 11 ADD 1 TO F-CONS-CONS ON SIZE.
023200 IF A ERROR GO TO ALERT.
023250 IF A NOT < 11 AND B NOT < 11 ADD 1 TO F-CONS-CONS ON SIZE.
023300 COMPUTE L = N + 2.
023350 IF CC (L) = SPACE GO TO COL-STEP.
023400 IF CC (L) = '-' AND L = 73 PERFORM CARD-STEP-TWO.
023450 MOVE 1 TO B.
023500 CK-THREE.
023550 IF CC (L) = LTR (B) PERFORM ADEE.
023600 IF CC (L) = LTR (B) GO TO COL-STEP.
023650 IF B = 33 GO TO COL-STEP.
023700 ADD 1 TO B.
023750 GO TO CK-THREE.
023800 COL-STEP.

```

```

023850 IF N NOT < 72 GO TO BEGIN.
023900 ADD 1 TO N.
023950 GO TO DATA-COLL.
024000 EQJ.
024050 PERFORM CLOSE-OUT-RT.
024100 MOVE SPACES TO PRINT-LINE.
024150 WRITE PRINT-LINE AFTER PAGE-TOP.
024200 CLOSE SAMPLE, RESULT.
024250 STOP RUN.
024300 BLANK-COL-STEP.
024350 ADD 1 TO N.
024400 IF N = 73 GO TO BEGIN.
024450 GO TO DATA-COLL.
024500 CARD-STEP-ONE.
024550 MOVE 2 TO N.
024600 MOVE CC (72) TO CC (2).
024650 READ SAMPLE AT END GO TO EQJ.
024700 PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.
024750 ADD 1 TO CARD-NR ON SIZE ERROR GO TO ALERT.
024800 COMPUTE M = N + 1.
024850 CARD-STEP-TWO.
024900 MOVE 1 TO N.
024950 MOVE CC (71) TO CC (1).
025000 MOVE CC (72) TO CC (2).
025050 READ SAMPLE AT END GO TO EQJ.
025100 PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.
025150 ADD 1 TO CARD-NR ON SIZE ERROR GO TO ALERT.
025200 COMPUTE L = N + 2.
025250 ADER.
025300 COMPUTE B = B + 33.
025350 ADD 1 TO FOL-FREQ (A, B) ON SIZE ERROR GO TO ALERT.
025400 COMPUTE B = B - 33.
025450 DAT-SHIFT.
025500 COMPUTE Y = X + 2.
025550 MOVE CB (X) TO CC (Y).
025600 ENG-FILL-1.
025650 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.
025700 ADD 1 TO L.
025750 IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-1.
025800 MOVE CBF-IN (L) TO ENG-ICRF (A).
025850 ADD 1 TO A.
025900 IF L = 11 MOVE ZERO TO L.
025950 ENG-CK-1.
026000 IF DC-ID NOT = 'A' OR DC-SEQ NCT = CARD-NR
026050 OR DC-LANG NOT = 'ENGLISH P' GO TO DEFAULT.
026100 ENG-FILL-2.
026150 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.
026200 IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-2.

```

```

0026250 ADD 1 TO L.
0026300 MOVE CRF-IN (L) TO ENG-DCRF (A, B).
0026350 IF B = 33, TO A.
0026400 ADD 1 TO A.
0026450 MOVE ZERO TO B.
0026500 ADD 1 TO B.
0026550 IF L = 11 MOVE ZERO TO L.
0026600 ENG-CK-2.
0026650 IF DC-ID NOT = 'B' OR DC-SEQ NOT = CARD-NR
0026700 OR DC-LANG NOT = 'ENGLISH P', GO TO DEFAULT.
0026750 ENG-FILL-3.
0026800 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.
0026850 IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-3.
0026900 ADD 1 TO L.
0026950 MOVE CRF-IN (L) TO ENG-DCRF (A, B).
0027000 IF B = 66
0027050 ADD 1 TO A.
0027100 MOVE 33 TO B.
0027150 ADD 1 TO B.
0027200 IF L = 11 MOVE ZERO TO L.
0027250 ENG-CK-3.
0027300 IF DC-ID NOT = 'C' OR DC-SEQ NOT = CARD-NR
0027350 OR DC-LANG NOT = 'ENGLISH P', GO TO DEFAULT.
0027400 SPAN-FILL-1.
0027450 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.
0027500 IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-1.
0027550 ADD 1 TO L.
0027600 MOVE CRF-IN (L) TO SPAN-ICRF (A).
0027650 ADD 1 TO A.
0027700 IF L = 11 MOVE ZERO TO L.
0027750 SPAN-CK-1.
0027800 IF DC-ID NOT = 'A' OR DC-SEQ NOT = CARD-NR
0027850 OR DC-LANG NOT = 'SPANISH P', GO TO DEFAULT.
0027900 SPAN-FILL-2.
0027950 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.
0028000 IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-2.
0028050 ADD 1 TO L.
0028100 MOVE CRF-IN (L) TO SPAN-DCRF (A, B).
0028150 IF B = 33, TO A.
0028200 ADD 1 TO A.
0028250 MOVE ZERO TO B.
0028300 ADD 1 TO B.
0028350 IF L = 11 MOVE ZERO TO L.
0028400 SPAN-CK-2.
0028450 IF DC-ID NOT = 'B' OR DC-SEQ NOT = CARD-NR
0028500 OR DC-LANG NOT = 'SPANISH P', GO TO DEFAULT.
0028550 SPAN-FILL-3.
0028600 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.

```

```

028650 IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-3.
028700 ADD 1 TO L.
028750 MOVE CRF-IN (L) TO SPAN-DCRF (A, B).
028800 IF B = 66, TO A.
028850 ADD 1 TO A.
028900 MOVE 33 TO B.
028950 ADD 1 TO B.
029000 IF L = 11 MOVE ZERO TO L.
029050 SPAN-CK-3.
029100 IF DC-ID NOT = 'C' OR DC-SEQ NOT = CARD-NR
029150 OR DC-LANG NOT = 'SPANISH P' GO TO DEFAULT.
029200 DEFAULT.
029250 MOVE ' RUN TERMINATED DUE TO INPUT SEQUENCE ERROR ' TO
029300 PRINT-LINE.
029350 WRITE PRINT-LINE AFTER PAGE-TOP.
029400 CLOSE SAMPLE, RESULT.
029450 STOP RUN.
029500 ALERT.
029550 WRITE PRINT-LINE FROM ERROR-LINE AFTER PAGE-TOP.
029600 CYCLE.
029650 READ SAMPLE AT END GO TO ECJ.
029700 IF NEXT-SAMPLE GO TO RESTART.
029750 GO TO CYCLE.
029800 RESTART.
029850 FORM INITIALIZE-RT.
029900 PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.
029950 PERFORM TO CARD-NR.
030000 ADD 1 TO N.
030050 GO TO DATA-COLL.
030100 CLOSE-OUT-RT.
030150 IF CARD-NR = 0 GO TO INITIALIZE-RT.
030200 PERFORM IND-TOTAL-CALC VARYING A FROM 1 BY 1 UNTIL A > 33.
030250 PERFORM IND-PROD-CALC VARYING A FROM 1 BY 1 UNTIL A > 33.
030300 PERFORM YULE-PROD VARYING A FROM 1 BY 1 UNTIL A > 33.
030350 COMPUTE K ROUNDED = (10000 * (S - 1LF-TOTAL)) /
030400 (1LF-TOTAL ** 2).
030450 COMPUTE Z-ENG ROUNDED = (K - 645.079) * (CARD-NR ** .5) /
030500 81.497.
030550 COMPUTE Z-SPAN ROUNDED = (K - 751.447) * (CARD-NR ** .5) /
030600 82.323.
030650 PERFORM LANG-CK-K-EXIT.
030700 WRITE PRINT-LINE FROM TITLE-AFTER PAGE-TOP.
030750 WRITE PRINT-LINE FROM TITLE-UL BEFORE 4.
030800 WRITE PRINT-LINE FROM COL-HEAD BEFORE 0.
030850 WRITE PRINT-LINE FROM COL-HEAD-UL BEFORE 2.
030900 MOVE BANK-FORM TO BANK-CNE.
030950 WRITE PRINT-LINE FROM C-LINE-ONE BEFORE 2.
031000 PERFORM CRILP-CALC VARYING A FROM 2 BY 1 UNTIL A > 33.

```



```

0035850 PERFORM ZEROING-TWO VARYING A FROM 1 BY 1 UNTIL A > 33
0035900 AFTER B FROM 1 BY 1 UNTIL B > 66.
0035950 READ SAMPLE AT END GO TO DEFAULT.
0036000 LANG-CK-K-
0036050 IF Z-ENG
0036100 MOVE 'NU DECSN' TO BANK-LANG,
0036150 MOVE Z-ENG TO Z.
0036200 IF Z-ENG
0036250 MOVE 'SPANISH' TO BANK-LANG,
0036300 MOVE Z-SPAN TO Z.
0036350 IF Z-ENG
0036400 MOVE 'ENGLISH' TO BANK-LANG,
0036450 MOVE Z-ENG TO Z.
0036500 IF Z > 4
0036550 0.000060 TO BANK-ALFA
0036600 MOVE 'K' TO BANK-EQUAL
0036650 GO LANG-CK-K-EXIT.
0036700 COMPUTE Z2 ROUNDED = (1 / 6) * Z2.
0036750 COMPUTE P = 1 - P ON SIZE ERROR GO TO ALERT.
0036800 COMPUTE P = 1 - P ON SIZE ERROR GO TO ALERT.
0036850 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0036900 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0036950 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037000 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037050 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037100 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037150 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037200 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037250 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037300 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037350 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037400 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037450 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037500 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037550 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037600 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037650 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037700 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037750 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037800 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037850 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0037900 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0037950 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0038000 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0038050 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0038100 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.
0038150 COMPUTE P = J - P ON SIZE ERROR GO TO ALERT.
0038200 COMPUTE P = J + P ON SIZE ERROR GO TO ALERT.

```

```

0382500 COMPUTE J = (P / 34.06451612903) * Z2.
0383000 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0383500 COMPUTE J - P ON (P / 36.060606061) * Z2.
0384000 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0384500 COMPUTE J + P ON (P / 38.05714285714) * Z2.
0385000 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0385500 COMPUTE J = P ON (P / 40.05405405405) * Z2.
0386000 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0386500 COMPUTE J - P ON (P / 42.05128205128) * Z2.
0387000 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0387500 COMPUTE J + P ON (P / 44.04878048780) * Z2.
0388000 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0388500 COMPUTE J = P ON (P / 46.04651162791) * Z2.
0389000 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0389500 COMPUTE J - P ON (P / 48.04444444444) * Z2.
0390000 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0390500 COMPUTE J + P ON (P / 50.042553191) * Z2.
0391000 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0391500 COMPUTE J = P ON (P / 52.040816326) * Z2.
0392000 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0392500 COMPUTE J - P ON (P / 54.039215686) * Z2.
0393000 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0393500 COMPUTE J + P ON (P / 56.037725849) * Z2.
0394000 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0394500 COMPUTE J = P ON (P / 58.03536) * Z2.
0395000 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0395500 COMPUTE J - P ON (P / 60.03509) * Z2.
0396000 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0396500 COMPUTE J + P ON (P / 62.03) * Z2.
0397000 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0397500 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0398000 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0398500 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0399000 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0399500 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0400000 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0400500 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0401000 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0401500 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0402000 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0402500 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0403000 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0403500 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0404000 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0404500 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.
0405000 COMPUTE J = P ON SIZE ERROR GO TO ALERT * Z2.
0405500 COMPUTE J + P ON SIZE ERROR GO TO ALERT * Z2.
0406000 COMPUTE J - P ON SIZE ERROR GO TO ALERT * Z2.

```

3989422804014 *

MOVE SPACES TO BANK-EQUAL.

LANG-CK-K-EXIT. EXIT.

CRILP-COMPUTE B = A - 1. ILP (5) + ILP (A).

IND-TOTAL-CALC. ILP (A) = ILP (5) + ILP (A).

IND-PRCB-CALC. ILF-TOTAL = ILF-TOTAL + ILF (A).

IF ILF (A) > 0 COMPUTE ILP (A) ROUNDED =

(ILF (A) / ILF-TOTAL), ELSE MOVE ZEROS TO ILP (A).

YULE-PRERE S = S + (ILF (A) ** 2).

DEPN-TWO-SUM FOL-FREQ-TOTAL = FOL-FREQ-TOTAL + FOL-FREQ (A, B).

DEPN-THREE-PROR IF FOL-FREQ (A, B) > 0 COMPUTE FOL-PRCB (A, B) ROUNDED =

IDEN-ONE
IDEN-ONE
IDEN-ONE

043050 MOVE ZEROS TO ILF (A), ILP (A).
043100 ZEROING-TWO.
043150 MOVE ZEROS TO FOL-FREQ (A, B), FOL-PROB (A, B).

APPENDIX J. LANGUAGE IDENTIFICATION PROGRAM (V-2)

```

000050 IDENTIFICATION DIVISION.
000100 PROGRAM-ID. LANGUAGE-IDENTIFICATION-TWO.
000150 AUTHOR. L. MORRISON D. RAU, USN.
000200 REMARKS. THIS PROGRAM IS PART OF MY MASTER OF SCIENCE IN SAMPLE
000250 MANAGEMENT THESIS AND IS DESIGNED TO IDENTIFY SHORT THE
000300 TEXT OF A LANGUAGE AS EITHER ENGLISH OR SPANISH.
000350 DECISION IS BASED ON STATISTICAL CHARACTERISTICS OF LETTERS,
000400 THE SPACE RATES AND THE INTERDEPENDENCY OF LETTERS IN
000450 ADJOINING WORDS.
000500 INPUT DATA MUST APPEAR IN THE FOLLOWING ORDER:
000550 4 ENGLISH D CARDS - ICRF TABLE VALUES
000600 106 ENGLISH F CARDS - ICRF TABLE VALUES X(2)/X(1)
000650 4 ENGLISH D CARDS - ICRF TABLE VALUES
000700 106 SPANISH D CARDS - ICRF TABLE VALUES
000750 106 SPANISH F CARDS - ICRF TABLE VALUES X(2)/X(1)
000800 106 SPANISH D CARDS - ICRF TABLE VALUES X(3)/X(1)
000850 106 SPANISH F CARDS - ICRF TABLE VALUES X(3)/X(1)
000900 1 HEAD CARD PRECEDING EACH GROUP OF SAMPLE CARDS.
000950 MAXIMUM OF SAMPLE CARDS PER TEST IS 999 OR 9,999
001000 INDIVIDUAL CHARACTER, INCLUDING SPACES.
001050 INDIVIDUAL DIVISION.
001100 ENVIRONMENT SECTION. IBM-360-67.
001150 CONFIGURATION SECTION. IBM-360-67.
001200 SPECIAL-NAMES.
001250 COI IS PAGE-TOP.
001300 INPUT-CONTROL SECTION.
001350 FILE-CONTROL. SAMPLE ASSIGN TO UR-S-INI.
001400 SELECT. RESULT ASSIGN TO UR-S-OUTI.
001450 DATA DIVISION.
001500 FILE SECTION.
001550 LABEL RECORDS ARE OMITTED
001600 BLOCK CONTAINS 5 RECORDS
001650 DATA RECORD IS SAMPLE-CARD, HEAD-CARD, DATA-CARD.
001700 SAMPLE FILLER.
001750 01 05 SAMPLE-FLD.
001800 05 05 FILLER.
001850 05 05 FILLER.
001900 05 05 FILLER.
001950 05 05 FILLER.
002000 05 05 FILLER.
002050 05 05 FILLER.
002100 05 05 FILLER.
002150 05 05 FILLER.
002200 05 05 FILLER.

```

002250	05	TEST-NAME	PIC X(20):	IDEN-TWO
002300	05	FILLER	PIC X(58):	IDEN-TWO
002350	05	DATA-CARD.		IDEN-TWO
002400	05	DC-ID	PIC A99.	IDEN-TWO
002450	05	DC-SEQ	PIC X(10):	IDEN-TWO
002500	05	DC-LANG	PIC 9V9(5), OCCURS 11 TIMES.	IDEN-TWO
002550	05	CRF-IN		IDEN-TWO
002600	FD	RE LABEL RECORDS ARE OMITTED		IDEN-TWO
002650		BLOCK CONTAINS 5 RECORDS		IDEN-TWO
002700		DATA RECORD IS PRINT-LINE		IDEN-TWO
002750		PRINT-LINE	PIC X(133).	IDEN-TWO
002800	01	WORKING-STORAGE SECTION.		IDEN-TWO
002850	77	A	PIC 99, COMP, SYNC.	IDEN-TWO
002900	77	B	PIC 99, COMP, SYNC.	IDEN-TWO
002950	77	N	PIC 99, COMP, SYNC.	IDEN-TWO
003000	77	S	PIC 9(12), COMP, SYNC.	IDEN-TWO
003050	77	J	PIC 99, COMP, SYNC.	IDEN-TWO
003100	77	M	PIC 99, COMP, SYNC.	IDEN-TWO
003150	77	X	PIC 99, COMP, SYNC.	IDEN-TWO
003200	77	Y	PIC 9(4), COMP, SYNC.	IDEN-TWO
003250	77	VC-TOTAL	PIC 9(4), COMP, SYNC.	IDEN-TWO
003300	77	F-CONS-CONS	PIC 9(4), COMP, SYNC.	IDEN-TWO
003350	77	F-VOW-CONS	PIC 9(4), COMP, SYNC.	IDEN-TWO
003400	77	F-CONS-VOW	PIC 9(4), COMP, SYNC.	IDEN-TWO
003450	77	F-VOW-VOW	PIC 9(4), COMP, SYNC.	IDEN-TWO
003500	77	F-FREQ-TOTAL	PIC 9(4), COMP, SYNC.	IDEN-TWO
003550	77	ILF-TOTAL	PIC 9(3), COMP, SYNC.	IDEN-TWO
003600	77	CARD-NR	PIC 9(13), V999.	IDEN-TWO
003650	77	K	PIC 9(15), V999.	IDEN-TWO
003700	77	Z-ENG	PIC 9(15), V999.	IDEN-TWO
003750	77	Z-SPAN	PIC 9(15), V999.	IDEN-TWO
003800	77	Z	PIC 9(15), V999.	IDEN-TWO
003850	77	Z	PIC 9(15), V999.	IDEN-TWO
003900	77	Z	PIC 9(15), V999.	IDEN-TWO
003950	77	Z	PIC 9(15), V999.	IDEN-TWO
004000	77	Z	PIC 9(15), V999.	IDEN-TWO
004050	77	Z	PIC 9(15), V999.	IDEN-TWO
004100	77	Z	PIC 9(15), V999.	IDEN-TWO
004150	77	Z	PIC 9(15), V999.	IDEN-TWO
004200	77	Z	PIC 9(15), V999.	IDEN-TWO
004250	77	Z	PIC 9(15), V999.	IDEN-TWO
004300	77	Z	PIC 9(15), V999.	IDEN-TWO
004350	77	Z	PIC 9(15), V999.	IDEN-TWO
004400	77	Z	PIC 9(15), V999.	IDEN-TWO
004450	77	Z	PIC 9(15), V999.	IDEN-TWO
004500	77	Z	PIC 9(15), V999.	IDEN-TWO
004550	77	Z	PIC 9(15), V999.	IDEN-TWO
004600	01	Z	PIC 9(15), V999.	IDEN-TWO

[illegible]

```

READ SAMPLE AT END GO TO DEFAULT.
ADD 1 TO CARD-NR.
PERFORM ENG-CK-3.
MOVE CRF-IN (1) TO ENG-DCRF (34, 68).
MOVE 1 TO A.
MOVE 0 TO L.
CARD-NR.
PERFORM SPAN-FILL-1 33 TIMES.
READ SAMPLE AT END GO TO DEFAULT.
ADD 1 TO CARD-NR.
PERFORM SPAN-CK-1.
MOVE CRF-IN (1) TO SPAN-ICRF (34).
MOVE 1 TO A.
MOVE 0 TO L.
CARD-NR.
PERFORM SPAN-FILL-2 155 TIMES.
READ SAMPLE AT END GO TO DEFAULT.
ADD 1 TO CARD-NR.
PERFORM SPAN-CK-2.
MOVE CRF-IN (1) TO SPAN-DCRF (34, 34).
MOVE 1 TO A.
MOVE 35 TO B.
CARD-NR.
MOVE 0 TO L.
CARD-NR.
PERFORM SPAN-FILL-3 155 TIMES.
READ SAMPLE AT END GO TO DEFAULT.
ADD 1 TO CARD-NR.
PERFORM SPAN-CK-3.
MOVE CRF-IN (1) TO SPAN-DCRF (34, 68).
MOVE ZEROS TO CARD-NR.
ALPHA. PERFORM READ-A-CARD.
MOVE 2 TO N.
BRAVO.
IF CL (N) = SPACE GO TO CK-ONE.
PERFORM FIRST-LETTER-CHECK THRU FIRST-
IF MATCH = ZERO GO TO CK-ONE.
ADD 1 TO ILF (A) ON SIZE ERROR GO TO A
COMPUTE M = N + 1.
IF M = 73 PERFORM SHIFT-TWO.
IF CL (M) = SPACE GO TO CK-TWO.
PERFORM SECOND-LETTER-CHECK THRU SECON
IF MATCH = ZERO GO TO CK-TWO.
CHARLIE.
ADD 1 TO FOL-FREQ (A, B) ON SIZE ERROR
PERFORM VC-TALLEY THRU VC-EXIT.
COMPUTE L = N + 2.
IF L = 73 AND CL (73) = ' ' PERFORM SH
IF L = 73 AND CL (73) NOT = ' ' PERFOR
IF CL (L) = SPACE GO TO CK-THREE.
PERFORM THIRD-LETTER-CHECK THRU THIRD-

```


026250	MOVE 1 TO N.	IDEN-TWO
026300	MOVE 3 TO L.	IDEN-TWO
026350	ALT-THREE.	IDEN-TWO
026400	MOVE CL (72) TO CL (1).	IDEN-TWO
026450	MOVE SPACES TO CL (2).	IDEN-TWO
026500	PERFORM READ-A-CARD.	IDEN-TWO
026550	MOVE 0 TO N.	IDEN-TWO
026600	MOVE 2 TO L.	IDEN-TWO
026650	ENG-FILL-1.	IDEN-TWO
026700	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	IDEN-TWO
026750	IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-1.	IDEN-TWO
026800	ADD 1 TO L.	IDEN-TWO
026850	MOVE CRF-IN (L) TO ENG-ICRF (A).	IDEN-TWO
026900	ADD 1 TO A.	IDEN-TWO
026950	IF L = 11 MOVE ZERO TO L.	IDEN-TWO
027000	ENG-CK-1.	IDEN-TWO
027050	IF DC-ID NOT = 'D' OR DC-SEQ NOT = CARD-NR	IDEN-TWO
027100	OR DC-LANG NOT = 'ENGLISH P' GO TO DEFAULT.	IDEN-TWO
027150	ENG-FILL-2.	IDEN-TWO
027200	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	IDEN-TWO
027250	IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-2.	IDEN-TWO
027300	ADD 1 TO L.	IDEN-TWO
027350	MOVE CRF-IN (L) TO ENG-DCRF (A, B).	IDEN-TWO
027400	IF B = 34, TO A.	IDEN-TWO
027450	ADD 1 TO A.	IDEN-TWO
027500	MOVE ZERO TO B.	IDEN-TWO
027550	ADD 1 TO B.	IDEN-TWO
027600	IF L = 11 MOVE ZERO TO L.	IDEN-TWO
027650	ENG-CK-2.	IDEN-TWO
027700	IF DC-ID NOT = 'E' OR DC-SEQ NOT = CARD-NR	IDEN-TWO
027750	OR DC-LANG NOT = 'ENGLISH P' GO TO DEFAULT.	IDEN-TWO
027800	ENG-FILL-3.	IDEN-TWO
027850	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	IDEN-TWO
027900	IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-3.	IDEN-TWO
027950	ADD 1 TO L.	IDEN-TWO
028000	MOVE CRF-IN (L) TO ENG-DCRF (A, B).	IDEN-TWO
028050	IF B = 68, TO A.	IDEN-TWO
028100	ADD 1 TO A.	IDEN-TWO
028150	MOVE 34 TO B.	IDEN-TWO
028200	ADD 1 TO B.	IDEN-TWO
028250	IF L = 11 MOVE ZERO TO L.	IDEN-TWO
028300	ENG-CK-3.	IDEN-TWO
028350	IF DC-ID NOT = 'F' OR DC-SEQ NOT = CARD-NR	IDEN-TWO
028400	OR DC-LANG NOT = 'ENGLISH P' GO TO DEFAULT.	IDEN-TWO
028450	SPAN-FILL-1.	IDEN-TWO
028500	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	IDEN-TWO
028550	IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-1.	IDEN-TWO
028600	ADD 1 TO L.	IDEN-TWO

028650	MOVE CRF-IN (L) TO SPAN-ICRF (A).	IDEN-TWO
028700	ADD 1 TO A.	IDEN-TWO
028750	IF L = 11 MOVE ZERO TO L.	IDEN-TWO
028800	SPAN-CK-1.	IDEN-TWO
028850	IF DC-ID NOT = 'D' OR DC-SEQ NOT = CARD-NR	IDEN-TWO
028900	OR DC-LANG NOT = 'SPANISH P', GO TO DEFAULT.	IDEN-TWO
028950	SPAN-FILL-2.	IDEN-TWO
029000	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	IDEN-TWO
029050	IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-2.	IDEN-TWO
029100	ADD 1 TO L.	IDEN-TWO
029150	MOVE CRF-IN (L) TO SPAN-DCRF (A, B).	IDEN-TWO
029200	IF B = 34,	IDEN-TWO
029250	ADD 1 TO A.	IDEN-TWO
029300	MOVE ZERO TO B.	IDEN-TWO
029350	ADD 1 TO B.	IDEN-TWO
029400	IF L = 11 MOVE ZERO TO L.	IDEN-TWO
029450	SPAN-CK-2.	IDEN-TWO
029500	IF DC-ID NOT = 'E' OR DC-SEQ NOT = CARD-NR	IDEN-TWO
029550	OR DC-LANG NOT = 'SPANISH P', GO TO DEFAULT.	IDEN-TWO
029600	SPAN-FILL-3.	IDEN-TWO
029650	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	IDEN-TWO
029700	IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-3.	IDEN-TWO
029750	ADD 1 TO L.	IDEN-TWO
029800	MOVE CRF-IN (L) TO SPAN-DCRF (A, B).	IDEN-TWO
029850	IF B = 68,	IDEN-TWO
029900	ADD 1 TO A.	IDEN-TWO
029950	MOVE 34 TO B.	IDEN-TWO
030000	ADD 1 TO B.	IDEN-TWO
030050	IF L = 11 MOVE ZERO TO L.	IDEN-TWO
030100	SPAN-CK-3.	IDEN-TWO
030150	IF DC-ID NOT = 'F' OR DC-SEQ NOT = CARD-NR	IDEN-TWO
030200	OR DC-LANG NOT = 'SPANISH P', GO TO DEFAULT.	IDEN-TWO
030250	DEFAULT.	IDEN-TWO
030300	MOVE . RUN TERMINATED DUE TO INPUT SEQUENCE ERROR . TO	IDEN-TWO
030350	PRINT-LINE.	IDEN-TWO
030400	WRITE PRINT-LINE AFTER PAGE-TOP.	IDEN-TWO
030450	CLOSE SAMPLE, RESULT.	IDEN-TWO
030500	STOP RUN.	IDEN-TWO
030550	ALERT.	IDEN-TWO
030600	WRITE PRINT-LINE FROM ERROR-LINE AFTER PAGE-TOP.	IDEN-TWO
030650	CYCLE.	IDEN-TWO
030700	READ SAMPLE AT END GO TO ECJ.	IDEN-TWO
030750	IF NEXT-SAMPLE GO TO RESTART.	IDEN-TWO
030800	GO TO CYCLE.	IDEN-TWO
030850	RESTART.	IDEN-TWO
030900	PERFORM INITIALIZE-RT.	IDEN-TWO
030950	PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.	IDEN-TWO
031000	ADD 1 TO CARD-NR.	IDEN-TWO

```

031050 MOVE 2 TO N.
031100 GO TO BRAVO.
031150 CLOSE-CUT-RT.
031200 SUBTRACT 1 FROM ILF (34).
031250 PERFORM IND-TOTAL-CALC VARYING A FROM 1 BY 1 UNTIL A > 34.
031300 PERFORM IND-PRGB-CALC VARYING A FROM 1 BY 1 UNTIL A > 34.
031350 PERFORM YULE-PRGB VARYING A FROM 1 BY 1 UNTIL A > 34.
031400 COMPUTE K ROUNDED = (10000 * (S - ILF-TOTAL)) /
031450 (ILF-TOTAL ** 2).
031500 COMPUTE Z-ENG ROUNDED = (K - 747.77) * (CARD-NR ** .5) /
031550 95.973.
031600 COMPUTE Z-SPAN ROUNDED = (K - 806.481) * (CARD-NR ** .5) /
031650 77.81.
031700 PERFORM LANG-CK-K THRU LANG-CK-K-EXIT.
031750 WRITE PRINT-LINE FROM TITLE-AFTER PAGE-TCP.
031800 WRITE PRINT-LINE FROM TITLE-UL BEFORE 4.
031850 WRITE PRINT-LINE FROM COL-HEAD BEFORE 0.
031900 WRITE PRINT-LINE FROM COL-HEAD-UL BEFORE 2.
031950 MOVE BANK-FORM TO BANK-ONE.
032000 WRITE PRINT-LINE FROM G-LINE-CNE BEFORE 2.
032050 PERFORM CRILP-CALC VARYING A FROM 2 BY 1 UNTIL A > 34.
032100 PERFORM DEPN-TWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 34.
032150 AFTER DEPN-TWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 34.
032200 PERFORM DEPN-TWO-PRGB VARYING A FROM 1 BY 1 UNTIL A > 34.
032250 AFTER DEPN-TWO-PRGB VARYING A FROM 1 BY 1 UNTIL A > 34.
032300 MOVE ZEROS TO S.
032350 PERFORM YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 34.
032400 AFTER YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 34.
032450 COMPUTE K ROUNDED = (10000 * (S - FOL-FREQ-TOTAL)) /
032500 (FOL-FREQ-TOTAL ** 2).
032550 COMPUTE Z-FREQ ROUNDED = (K - 87.93) * (CARD-NR ** .5) /
032600 31.618.
032650 COMPUTE Z-SPAN ROUNDED = (K - 108.784) * (CARD-NR ** .5) /
032700 30.665.
032750 PERFORM LANG-CK-K THRU LANG-CK-K-EXIT.
032800 MOVE BANK-FORM TO BANK-TWO.
032850 WRITE PRINT-LINE FROM Q-LINE-TWO BEFORE 2.
032900 PERFORM CR-CK-FOL-PRGB-CALC VARYING A FROM 1 BY 1 UNTIL A > 34.
032950 AFTER CR-CK-FOL-PRGB-CALC VARYING A FROM 1 BY 1 UNTIL A > 34.
033000 MOVE FOL-FREQ-TOTAL TO FET2.
033050 MOVE ZEROS TO FOL-FREQ-TOTAL.
033100 PERFORM DEPN-TWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 34.
033150 AFTER DEPN-TWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 34.
033200 PERFORM DEPN-TWO-PRGB VARYING A FROM 1 BY 1 UNTIL A > 34.
033250 AFTER DEPN-TWO-PRGB VARYING A FROM 1 BY 1 UNTIL A > 34.
033300 MOVE ZEROS TO S.
033350 PERFORM YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 34.
033400 AFTER YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 34.

```


122

[illegible]

LIST OF REFERENCES

1. Nakamura, Y., "Identification of Languages with Short Sample Texts: A Linguometric Study," Library and Information Science, No. 9, pp 459-81, 1971.
2. Yule, G. U., The Statistical Study of Literary Vocabulary, Cambridge, 1944.
3. Massey, F. J., jr., "A Note on the Estimation of a Distribution Function by Confidence Limits," Annals of Mathematical Statistics, v. 21, pp 116-19, 1950.
4. Birnbaum, Z. W., "Numerical Tabulation of the Distribution of Kolmogorov's Statistic for Finite Sample Size," Journal of the American Statistical Association, v. 47, pp 425-41, September 1952.
5. Birnbaum, Z. W. and Tingey, F. H., "One-Sided Confidence Contours for Probability Distribution Functions," Annals of Mathematical Statistics, v. 22, pp 592-96, 1951.
6. Miller, L. H., "Table of Percentage Points of Kolmogorov Statistic," Journal of the American Statistical Association, v. 51, pp 111-21, March 1956.
7. Smirnov, N., "Tables for Estimating the Goodness of Fit of Empirical Distributions," Annal of Mathematical Statistics, v. 19, pp 279-81, 1948.
8. Lapin, L. L., Statistics for Modern Business Decisions, Jovanovich, 1973.
9. Kelley, T. L., The Kelley Statistical Tables, p. 3, Macmillan, 1938.
10. Altenbernd, L. and Lewis, L. L., Introduction to Literature: Stories, 2nd ed., Macmillan, 1970.
11. Barrett, L. L., Five Centuries of Spanish Literature, Dodd, Mead, & Co., 1963.
12. Medina, Jose Ramon, Antologia Venezolana, Editorial Gredos, 1962.
13. Instituto Internacional de Literatura Iberoamericana, An Anthology of Spanish American Literature, 2nd ed., Appleton-Century-Crofts, 1946.
14. Stallman, R. W. and Watter, R. E., The Creative Reader, 2nd ed., Ronald Press, 1962.

15. Saturday Evening Post, the, Best Modern Short Stories, Curtis
Books, 1965.