PHONOSTATISTIC METHODS

Gary Simons
Summer Institute of Linguistics
and Cornell University

0. INTRODUCTION

Phonostatistics means any analytical technique which seeks to quantify the phonological differences between speech groups. Phonostatistic methods have some advantages over the more common lexicostatistic technique. In addition, phonostatistic results serve to complement the results of other language survey techniques in the analysis of language variation. It is for these reasons that I here discuss and evaluate twelve different phonostatistic methods that various investigators have proposed.

Phonostatistic methods may be applied for either diachronic or synchronic purposes (Simons 1977). In the case of a diachronic study, phonostatistic measures can be used as a measure of the phonological divergence between speech groups. Some methods may even indicate a relative time depth of divergence. In the case of a synchronic study, phonostatistic measures are used primarily as an indicator of the phonetic similarity of cognate words. Such a measure is used in helping to estimate the amount of intelligibility possible between two speech groups.

In this paper I first discuss some of the advantages of phonostatistics over lexicostatistics. In the second section, a full explanation of the twelve different phonostatistic methods is given. In section three, I evaluate the different methods. Finally, in the conclusion I suggest which methods are best for a diachronic study and which are best for a synchronic study, ending up with suggestions on how to relate phonostatistic measures and lexicostatistic measures to intelligibility and how to compute the phonostatistic measures.

1. ADVANTAGES OF PHONOSTATISTIC METHODS OVER LEXICOSTATISTIC METHODS

1.1 Sampling problem not critical

In phonostatistics, sampling methods are not as critical as in lexicostatistics. In lexicostatistics there is a sampling problem involved with eliciting "basic vocabulary". With the assumption that there is a special subset of the vocabulary that is more diagnostic of language relationships than any of the rest of the vocabulary, sampling becomes critical. Not just any list of words can be used for lexicostatistics.

In phonostatistics this is not so true. Almost all of the comparable phonological elements between two or more speech groups can be found in a limited corpus (McKaughan 1964:118). Whether the corpus is basic vocabulary or not should make no difference. Once a proper corpus is found, adding more items to it should not change the results of the comparisons significantly.

McKaughan's (1964) study of divergence in four Eastern Highlands languages of Papua New Guinea clearly shows the relative effect of sampling on lexicostatistic and phonostatistic relations. He first computed lexicostatistic relations on the basis of a 100-word list, a 200-word list and a 345-word list. For the 100 word comparison, the relation between the four languages ranged from 58% to 75%; for the 200-word comparison, they ranged from 36% to 52%; and for the 345-word comparison, they ranged from 25% to 34%. In each case there is a very dramatic drop in the cognate percentages as a larger sample of words is compared. This indicates certain vocabulary items do have a higher retention rate than others, in other words, there is a basic vocabulary (1964:100). When comparing the very same lists for phonostatistic relations, however, McKaughan found he got essentially the same results on each list. This indicates almost all of the comparable phonological elements can be found in a limited corpus (1964:118), and thus, sampling is not as critical in phonostatistics.

1.2 Generative elicitation

In lexicostatistics the elicitation method is critical; in phonostatistics there is much more freedom. In lexicostatistics one is not really measuring how many lexical forms are the same between two speech groups, rather one is measuring how many meanings have the same lexical form in the two speech groups. If the word for 'blood' in language A is historically cognate with the word for 'red' in language B, but the word for 'blood' in B is non-cognate, then in lexicostatistics A and B score no relationship for the item 'blood'. In lexicostatistics, semantic shift and the total loss of a form are treated equally. In phonostatistics this is not the case; we want to measure phonological shifts. When studying the phonological relation between A and B, we want to know that the word for 'blood' in A and the word for 'red' in B are the same root, for this root will certainly tell us something of the phonological differences between A and B.

This suggests that lexicostatistics and phonostatistics should have different strategies for eliciting data. In lexicostatistics we are comparing meanings so it is necessary to elicit roots by meaning. In phonostatistics we are comparing sounds so ideally we should elicit roots by sound. That is, instead of asking, "What is your word which means 'house'?", we can ask "Do you have a word that sounds like _____ and means something like 'house'?" or "In village A they say _____ for 'house'. What do you say?" Such elicitation techniques are forbidden in lexicostatistics; they are ideal for phonostatistics.

This method of elicitation could be called "generative elicitation." Vern Carrol (1966) was the first to speak of generative elicitation techniques in lexicography for eliciting new dictionary entries. The method suggested here for field elicitation of phonological data is in line with a technique suggested by Carrol of using dictionaries of other languages to suggest possible words (1966:67).

1.3 A more refined measure of historical relationship

Grimes (1964:49) suggests that the phonostatistic method he developed in Grimes and Agard (1959) and in Grimes (1964) is a more refined index of historical relationships than is lexicostatistics. This is because his method is based on comparative phonology and compares only items that are proven to be historically related. In lexicostatistics, however, even when comparative statements are available, the problem of determining cognates still remains. When all the segments do not correspond perfectly there is always the danger of rejecting a cognate that should have been accepted, or accepting as cognate what is not cognate at all.

McKaughan (1964:118) agreed with Grimes in concluding that phonostatistics is a more refined measure of linguistic relationship than lexicostatistics. He qualified this, however, by stating that the phonostatistical measure has its greatest value when the separation between languages is not very great. When the separation between languages increases, the value of phonostatistics decreases while lexicostatistics retains its value. Then as distance increases lexicostatistics loses its value and structural studies become the most fruitful.

2. SOME PHONOSTATISTIC METHODS

In this section twelve different phonostatistic methods are reviewed and explained. The reader should be able to get an idea of the many possibilities for phonostatistic measures. In the next section of the paper, the methods will be evaluated and suggestions made about the methods most practical for analyzing survey data.

2.1 Degrees of difference approaches

In the degrees of difference approaches, corresponding phonological items are compared and the differences between them are quantified as to the degree of difference. That is, items which hardly differ show a stronger relationship than items which are quite different. For diachronic studies, these methods are based on the observation by Austin (1957:544) that phonological change usually proceeds by minimal steps along one phonetic dimension at a time. Thus we assume that the greater the phonological differences between corresponding phonemes, the greater the time depth of divergence. Thus degrees of difference give a relative measure of time depth. For synchronic studies, these methods are based on the assumption that the more phonetically similar corresponding words are, the more likely it is that the native speaker from one group would understand the word if he heard it spoken in the speech of another group.

2.1.1 Grimes and Agard 1959

In their paper "Linguistic Divergence in Romance", Grimes and Agard published the first description of a phonostatistic method. The data for their study was 169 sets of phonological correspondences between seven Romance languages. The correspondence sets, which included almost all the important sound correspondences in Romance, were determined by the comparative method.[1]

The next step was to compare the corresponding sounds between each pair of languages. With seven languages, this involved 21 sets of comparisons for each correspondence set. If the corresponding sounds were phonetically the same, then that fact was recorded. If they were different, the next step was to determine how different.

The difference between sounds was quantified in accordance to a scale based on Pike's concept of rank of stricture (K.Pike 1943:129ff, 1967:329ff; E. Pike 1954:25-41). They saw that all the phonetic distinctions necessary to deal with the Romance data could be organized into six independent articulatory dimensions. Each of these dimensions, along with the scale used for ranking each sound according to each of the dimensions, are given below (1959:602).

(1) Point of articulation:

1 bilabial, 2 labiodental, 3 interdental, 4 apical or front vowel, 5 laminal or central vowel, 6 dorsal or back vowel, 7 glottal.

(2) Constriction of the airstream in the median line of the mouth:

1 closure, 2 local friction or loss of a vowel, 3 semivowels, 4 high vowel or loss of a consonant, 5 mid close vowel, 6 mid open vowel, 7 low vowel.

(3) Effective timing of the central constriction:

1 effective lack of constriction due to lateral opening to bypass the median constriction, 2 momentary constriction or flap, 3 "normal" constriction, 4 lenghtened constriction, including trill.

(4) Secondary shaping of the air stream in the mouth:

0 no secondary shaping, 1 vocoidal shaping, such as palatal or labial offglide or lip rounding, 2 contoidal shaping, such as affrication.

(5) Velic action:

0 velic closed, 1 velic open.

(6) Laryngeal action:

1 vocal cords vibrating, 2 vocal cords open.

Each sound was then converted to a vector of six numbers corresponding to each of the dimensions. For instance, the phone [p] is coded as follows: (1) point of articulation is 1, bilabial; (2) degree of constriction is 1, closure; (3) effective timing is 3, "normal" constriction; (4) secondary shaping is 0, no secondary shaping; (5) velic action is 0, velic closed; and (6) laryngeal action is 2, vocal cords open. Thus [p] is represented as the vector 113002. As another example, the nasalized vocoid [ũ] is coded as: 6

back vowel, 4 high vowel, 3 "normal" constriction, 1 lip rounding, 1 velic open, 1 vocal cords vibration, or 643111. The vector equivalents of all the phones in the Romance data are given in Table 1. A 9 indicates that the sound feature is irrelevant to the computation. For instance, in the glottal stop 713999, any supralaryngeal action is irrelevant. In loss of consonants 949999 and loss of vowels 929999, only degree of opening and closure counts.

a	473001	õ	653111
ã	473011	ɔ	663101
ɑ	573001	ɔ̃	663111
ɑ̃	573011	ø	453101
b	113001	ø̃	453111
b̄	123001	œ	463101
c	413202	œ̃	463111
č	513202	p	113002
d	413001	r	412001
d'	413101	r̃	414001
đ	323001	ṛ	612001
e	453001	ṛ̃	614001
ẽ	453011	s	423002
ɛ	463001	ś	423102
ɛ̃	463011	š	523002
f	223002	t	413002
g	613001	t'	413102
ǧ	513201	θ	323002
ɡ	623001	u	643101
i	443001	ũ	643111
ĩ	443011	v	223001
j,i̭	533001	w,ụ	133001
k	613002	ẉ	633101
l	411001	x	623002
ɫ	411101	y	443101
ɪ	511001	ỹ	443111
m	113011	ɥ	433101
n	413011	z	423001
ñ	513011	ź	423101
ŋ	613011	ž	523001
~	949919	ʒ	413201
o	653101	ʔ	713999

loss of consonant	949999
loss of vowel	929999
loss of nasal consonant with retention of nasalization	949919
assimilable nasal N	913011

Table I. Vector Equivalents

The degree of difference between two sounds is calculated by summing the numerical difference between each component in their vectors. Thus the degree of difference between [ƀ] and [p] is 2--fricative versus stop and voiced versus voiceless. That is,

[ƀ]	1 2 3 0 0 1
[p]	1 1 3 0 0 2
Numerical differences	0+1+0+0+0+1 = 2

The degree of difference is calculated for each pair of languages for all correspondence sets. This results in a tabulation of how many sets show d degrees of difference, where d represents 0 to the highest degrees of difference found in the data. For instance, the comparisons of Portuguese with Spanish and Portuguese with Rumanian result in the following:

Number of sets which show d degrees of difference

d	0	1	2	3	4	5	6	7
Po - Sp	84	50	15	12	2	5		
Po - Ru	63	28	16	31	19	8	0	2

Finally, the mean degrees of difference for all pairs of languages is computed as the measure of phonological difference. The mean degrees of difference is equal to the sum of degrees of difference for all correspondence sets, divided by the number of correspondence sets. Thus, the mean degrees of difference for Po - Sp and Po - Ru are:

$$\text{Po - Sp:}\quad \frac{84\text{x}0 + 50\text{x}1 + 15\text{x}2 + 12\text{x}3 + 2\text{x}4 + 5\text{x}5}{168} = 0.77$$

$$\text{Po - Ru:}\quad \frac{63\text{x}0 + 28\text{x}1 + 16\text{x}2 + 31\text{x}3 + 19\text{x}4 + 8\text{x}5 + 0\text{x}6 + 2\text{x}7}{167} = 1.70$$

Thus, on average, sound correspondences between Portuguese and Spanish differ by 0.77 degrees of differences, whereas those between Portuguese and Rumanian differ by 1.70 degrees of difference. We conclude that Rumanian has diverged further from Portuguese than has Spanish.

2.1.2 Grimes 1964

This paper by Grimes, "Measures of liguistic divergence", reports on some refinements of the methods used in Grimes and Agard 1964. The same Romance data are used and the phonological differences are quantified in the same way. The difference is in how measures of divergence are computed.

Grimes proposes that the number of correspondence sets that differ by 0,1,2,...,n degrees of differences be plotted on a graph. The degrees of difference would be numbered from left to right, the number of sets from bottom to top. If two groups have only recently diverged, the great majority of corresponding sounds will be identical. A few will have one degree of difference, maybe two. This distribution would appear as a very steep line on the graph as described by Grimes. If two groups have diverged for a long time, fewer correspondences will be identical. Many more will involve one or two degrees of differences. Some will involve 3, 4, or more degrees of differences. On the graph this would appear as a line with a less steep slope. These two possibilities are illustrated in Figure 1.

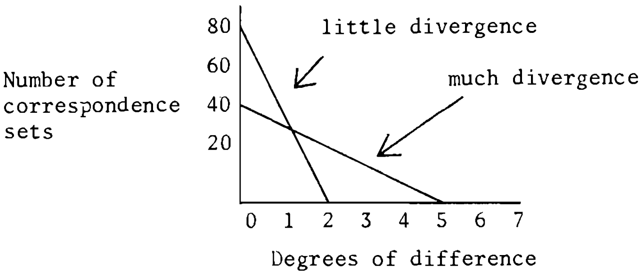


Figure 1

The slope of the line is thus an index of linguistic distance. The mathematical computation of the slope is a problem in regression analysis.[2] Grimes uses a linear regression, though suggests that an expoential function gives a better fit to the data.

When the data points are plotted as described above, not all the points lie exactly on a straight line; they will be scattered around the line. All the points may be very close to the line, or some may be at a considerable distance away. A correlation analysis[3] is the statistical measurement of the scatter around the regression line. Grimes (1964:46) suggests that "the scatter measurement reflects a disturbance of the ideal picture of phonological divergence brought about by random innovations". That is, the greater the scatter, the more the process of random innovation has been disturbed.

2.1.3 McKaughan 1964

In his paper entitled "A study of divergence in four New Guinea languages", McKaughan used an adaptation of the method suggested in Grimes and Agard 1959. He, too, sets up a scale based on rank of stricture. He found that a simplified scheme could be used to uniquely represent all the distinctions of the phonologies of the languages being studied. Only four dimensions were used with the following values:

(1) Point of articulation:

1 bilabial, 2 alveolar, 3 velar, 4 glottal, 5 front vowel, 6 central vowel, 7 back vowel.

(2) Degree of constriction of the air stream:

1 closure, 2 local friction or loss of a vowel, 3 semivowel, 4 high vowel or loss of a consonant, 5 mid vowel, 6 low vowel.

(3) Velic action:

0 velic closed, 1 velic open.

(4) Laryngeal action:

1 vocal cords vibrating, 2 vocal cords open.

The vector equivalents of all the sounds compared in McKaughan's study are given in Table 2.

p	1102	t	2102	k	3102	q	4199
ꝑ	1202	s	2202	k^w	3202	h	4299
b	1101	d	2101	g	3101		
ƀ	1201	r	2201	g	3101		
w	1301	y	2301				
m	1111	n	2111				
i	5401	ʌ	6501	u	7401		
e	5501	a	6601	o	7501		
ë	5601			ä	7601		
Loss of consonant	9499			Loss of vowel	9299		

Codification

Table 2

The method of McKaughan differs from that of Grimes and Agard in that the phonostatistic approach is usable at a stage in the investigation far ahead of the time when the results of the comparative method are available (1964:103). McKaughan suggests that the phonostatistic method be applied to languages which have already been shown to be related by a lexicostatistic study. The words which are considered as probable cognates (by the inspection method rather than the comparative method) in such a study are compared by the phonostatistic approach in order to determine the phonological divergence of cognate words. Thus, instead of comparing known phoneme correspondence sets to compute the mean degrees of difference for correspondence sets, McKaughan compares probable cognates phone by phone to compute the mean degrees of difference for cognate words.

His method is illustrated by the following example (1964:104).

The Tairora word for 'ear' is *ato*. The Asempa word is *aqʌ*. The degrees of difference for these two cognate words is computed as follows:

(*aqʌ*)	6601	4199	6501
(*ato*)	6601	2102	7501
	0000	2000	1000

degrees of difference = 3

The degrees of difference for each pair of cognate words is summed and divided by the number of cognates compared to obtain the mean degrees of difference for cognate words.

2.1.4 Ezard

In trying to relate a phonostatistic measure of degrees of difference to lexicostatistic or intelligibility relations, it would be helpful to have a phonostatistic method which yielded results in a comparable scale, that is, a percentage scale. For instance, the study by McKaughan yields mean degrees of difference measures in the range of .80 to 3.76, where the lowest numbers represent the closest relationships. Such results are not easily compared with lexicostatistic percentages which range from 0 to 100, where the highest numbers represent the closest relationships.

Bryan Ezard (personal communication) has suggested a means of computing phonostatistic relations as percentages.[4] He follows the method of McKaughan in quantifying the degrees of difference between cognate words. He then takes the computation a step further. Before computing the degrees of difference, all of the vector equivalents of the vowels are compared to see what the greatest possible number of degrees of difference between vowels is. Likewise, the consonants are compared to see what the greatest possible number of degrees of difference between consonants is. For instance, from the table of vector equivalents used by McKaughan (Table 2) we see that the maximum difference between vowels is 4 (eg. i 5401 and ä 7601, or ë 5601 and u 7401). For consonants, the maximum difference is 5 (eg. m 1111 and k^w 3202, or w 1301 and k 3102).

For each pair of segments, the number of degrees of difference is compared to the maximum number of degrees of difference for that segment type, that is, consonant or vowel. If the number of degrees of difference equals the maximum, then the two segments are maximally different, that is, 100% different. If there are zero degrees of difference, then the segments are 0% different; that is, they are maximally, or 100%, the same. The percentage of difference for any two segments is obtained by dividing the degrees of difference (d) for the segments by the maximum degrees of difference (m) for segments of that type and multiplying the result by 100. The percentage of sameness (P) is 100 minus the percentage of difference. That is,

$$P = 100 - \frac{100d}{m} \qquad \text{or} \qquad 100\ (1 - \frac{d}{m})$$

The percentage of difference for a whole word or word list is 100 times the sum of the degrees of difference for each segment compared divided by the sum of the maximum degrees of difference for all the segments compared. Thus, the percentage of sameness is

$$P = 100 \left(1 - \frac{\sum_{i=1}^{n} d_i}{\sum_{i=1}^{n} m_i} \right)$$

where, n is the total number of segments compared, d_i is the degrees of difference for the ith pair of segments compared, and m_i is the maximum degrees of difference for the segment type of ith pair compared. The sum of the maximum degrees of difference can be computed by adding the product of the number of consonants (n_c) and the maximum degrees of difference for consonants (m_c) to the product of the number of vowels (n_v) and the maximum degrees of difference for vowels (m_v). That is,

$$P = 100 \left(1 - \frac{\Sigma d}{n_c m_c + n_v m_v} \right)$$

2.1.5 Ladefoged 1968

Peter Ladefoged, in his study of cross-language communication in Uganda, used a phonostatistic measure based on binary distinctive features (see also Casad 1974:120-121). The rank of stricture method used by Grimes and Agard (1959), Grimes (1964), and McKaughan (1964) can be thought of as a method based on "n-ary" features. For instance, McKaughan uses the feature "point of articulation", which has seven distinctive values. In Ladefoged's method only binary distinctive features, as commonly used by the generative school of phonology, are employed. The codification of the segments (analogous to the table of vector equivalents in Grimes and Agard) is simply the binary feature matrix produced by the phonological analysis. To quantify the difference between two segments, one simply compares them feature by feature and adds up the number of features that are different. As an example, the feature specifications for four English consonants are given below (Chomsky and Halle 1968:177) along with a small table containing the number of features that are different between them.

	p	v	m	t
vocalic	-	-	-	-
consonantal	+	+	+	+
anterior	+	+	+	+
coronal	-	-	-	+
voice	-	+	+	-
continuant	-	+	-	-
nasal	-	-	+	-
strident	-	+	-	-

Number of features Different

	p	v	m
v	3		
m	2	3	
t	1	4	3

Ladefoged, like McKaughan, computed his phonostatistic measures on the basis of wordlists rather than established correspondence sets. By summing the number of features different between comparable segments in a pair of cognate words, one computes the number of feature differences for that pair. By summing the number of feature differences for all cognate pairs and dividing by the number of cognate words, one computes the average number of features different per cognate word.

2.1.6 Larsen 1975

Robert Larsen, in "Difference coefficients for determining dialect relatedness", uses a modified lexicostatistic approach which scales cognates on the basis of phonetic similarity to study the relationships between the dialects of Orokaiva in Papua New Guinea. It differs from Allen and Hurd 1963 in that the scale is based on the number of features (in the generative phonology sense) different rather than the number of phonemes different. He uses the values from 1 to 5 to represent the following degrees of difference (1975:12):

1, words with no difference;
2, words with just one feature different;
3, words with two feature changes or loss or addition of a segment;
4, words that are still recognizably cognate but have at least a combination of three feature differences or one or two each on two or more phones;
5, words with no recognizable common heritage except meaning.

When comparing two dialects, the sum of all values assigned to

The phonological relations are now quantified by counting the number of isoglosses which separate the dialects. We see that A and B are separated by one isogloss, A and C are separated by two, B and D are separated by three, and so on. The complete matrix of relations is:

A				
1	B			
2	1	C		
4	3	2	D	
4	3	2	0	E

2.2.2 Healey and Healey 1961

In their paper, "Dusun dialect comparison", Alan and Phyllis Healey describe a phonostatistic method which is also based on comparing correspondence sets found through the comparative method. All the sound correspondences between two dialects were compared to determine how many of the correspondences had identical sounds (that is, represented no change). In the case where one sound corresponded to two others through split, merger, or conditioning, half credit was given if one of those two sounds was the same. The number of identical correspondences was divided by the total number of correspondences compared to reach a percentage of identical sound correspondences.

This measure of phonological similarity was used in conjunction with the cognate percentage for a pair of dialects to compute an "index of expected mutual intelligibility". This was obtained by multiplying the percentage of cognates by the percentage of identical sound correspondences. This is based on the assumption that the number of cognate words a speaker from another dialect will understand is in proportion to the number of identical sound correspondences between the two dialects.

2.3 Typological approaches

The degrees of difference and isogloss approaches examine the phonological differences in their context, that is, corresponding segments in cognate words are compared. The typological approaches look only at the phonological systems and compare them. In a typological approach the fact that two languages have a /p/ phoneme counts as a sameness, even though the /p/ in one language may always correspond to a /b/ in the other language. In this situation, putting the /p/'s in the context of discourse, one would see that there is no sameness between the /p/'s.

2.3.1 Fodor 1965

The phonostatistic measures used by Fodor, in his monograph, "The rate of linguistic change", are measures which characterize a single phonological system, rather than give the relationship between two systems. By comparing the characterization of one system with the characterization of another system, one can get an idea of the relation between the two systems.

Fodor suggests the following as possible quantitative characterizations of a phonemic system (1965:45): absolute number of the phonemes, absolute number of vowels and consonants, the ratio of vowels to consonants, absolute numbers and ratios of the detailed genetic and/or acoustic types of phonemes (e.g. stops vs. fricatives, labials vs. non -labials), all these data detailed according to code and message, the load and the arrangement of the phonemes. For instance, the ratio of vowels to consonants would be obtained by dividing the number of vowel phonemes in the system by the number of consonant phonemes. If this ratio were computed for a number of languages, then they could be compared as to the ratio of vowels to consonants in their phonemic systems. Fodor suggests a method for doing this. He divides the ratio in one language by the ratio in another language. This ratio of ratios is a quantification of the relation between languages.

Fodor's distinction between measures in relation to code and those in relation to message is an important one (1965:44). Code refers to the linguistic system itself; message refers to its use in discourse. The ratio of vowels to consonants in the code would be computed as discussed in the previous paragraph. The ratio of vowels to consonants in the message would be computed by dividing the number of vowel occurrences in a representative sampling of discourse by the number of consonant occurrences. For instance, Fodor gives the ratio of vowels to consonants in the code and in the message for Italian, English, and Czech as (1965:44):

	Italian	English	Czech
	Vowel/Consonant	Vowel/Consonant	Vowel/Consonant
Code	7:26	18:26	12:24
Message	85:100	59:100	69:100

Inspection of these data shows how significant the distinction between measures in relation to code and to message is. Italian has, by far, the least number of vowels in the code, but the highest occurrence of them in the message. Conversely, English has the highest number of vowels in the code, but they occur the least frequently in the message.

2.3.2 Voegelin et al. 1963

This method compares the phonemic systems of all the languages of a family, a sub-family, or any other group of related languages, in order to derive an index of synchronic phonological divergence for that group of languages. Comparing the indexes of phonological divergence gives an idea of the relative diversity or homogeneity of the phonemic systems within a group of languages. The method works (1963:5)

> "...by constructing, for the sample available, a maximum non-existent phonemic system which shows every linear distinction made by any language in the sample...and also by constructing from the same sample a minimum non-existent phonemic system which shows no phonemic distinction lacking in any language in the sample.... Having constructed a minimax pair of non-existing phonemic systems for a given sample of languages, we compute and list the particular intervals (between the fewest and most stops, fricatives, nasals, and so on); finally, we add up the particular intervals and call the sum the general index. This serves as our index for synchronic diversity for a given sample."

Any differences due only to series generating components like voicing, prenasalization and nasalization (consider the three series p t k, b d g, mb nd ŋg) are ignored in computing the index of diversity (Wurm 1964b:3). The following example of a computation for two languages is from Wurm (1964b:6-7).

Gahuka Sub-family

Gahuka

p	t	k	?
b		g	
m	n		
	s		h
	z		
	r		
w			
i		u	
e		o	
	a		

Benabena

p	t	k	?
b		g	
m	n		
f	s		h
	r		
	y		
i		u	
e		o	
	a		

Non-existent minimax systems

Maximum				Minimum			
p	t	k	?	p	t	k	?
b		g		b		g	
m	n			m	n		
f	s		h		s		h
	z						
	r				r		
w	y						
i		u		i		u	
e		o		e		o	
	a				a		

Stops:	4 max. to 4 min. = 0 interval
Nasals:	2 max. to 2 min. = 0 interval
Fricatives:	3 max. to 2 min. = 1 interval
Liquids:	1 max. to 1 min. = 0 interval
Semi-vowels:	2 max. to 0 min. = 2 intervals
Vowels:	5 max. to 5 min. = 0 interval

Index of phonological diversity = 3

2.3.3 Typological features

The comparison of typological features is a common approach to language comparison. A paper by S.A. Wurm, "Australian New Guinea Highlands languages and the distribution of their typological features" (1964a), is used to demonstrate this approach.

Wurm lists 21 features of the phonology and grammar of New Guinea Highlands languages which he feels are important in classifying them. Five of these deal with phonology. They are (1964:82): the presence of the glottal stop phoneme (gS), laterally released stop phonemes (lS), somewhat complex syllable structures (cS), initial consonant clusters (iC), and nasal vowels (nV).

Each language is examined for each feature and is scored as having or not having a given feature. The results are displayed in a table with the features labelling the columns and the languages labelling the columns and the languages labelling the rows. At the intersection of each language with each feature is a plus (+) or a minus (-) indicating the presence or absence of the feature. A portion of Wurm's table (1964a:84-85) is reproduced in Table 3.

Language	gS	1S	cS	iC	nV
Gadsup	+	-	-	-	-
Kanite	+	+	-	-	-
Gahuku	+	-	-	-	-
Siane	-	-	-	-	-
Chimbu	-	+	+	+	-
Duna	-	-	-	-	+

Table 3

This is as far as Wurm takes the quantitative analysis. However, the phonological relatedness can be quantified from the data as it is given in Table 3. To compare two languages, the features are compared one by one and the number of identical features is added up. This number, the number of typological features shared by the two languages, is then a measure of phonological relatedness. The number of shared features could be divided by the total number of features to adjust all figures to a percentage scale. Number of shared features and percentage of shared features is given below for the data of Table 3.

Number of shared features

Gadsup					
4	Kanite				
5	4	Gahuku			
4	3	4	Siane		
1	2	1	2	Chimbu	
3	2	3	4	1	Duna

Percentage of shared features

Gadsup					
80	Kanite				
100	80	Gahuku			
80	60	80	Siane		
20	40	20	40	Chimbu	
60	40	60	80	20	Duna

The above method of computing relationships is equivalent to an isogloss approach. The distribution of typological features could just as well have been mapped by enclosing all languages which have a feature within an isogloss on a map. The number of lines separating the languages is then the number of features not shared.

3. EVALUATION OF METHODS

In the previous section, twelve different approaches to quantifying the phonological differences between speech groups were presented in order to give an idea of the wide range of possibilities. In this section the methods will be evaluated, particularly in view of their applicability for diachronic studies as measures of genetic relationship or for synchronic studies as predictors of intelligibility. All the phonostatistic methods reduce the phonological relations between speech groups to a number. The first two sections of the evaluation concern the numbers themselves and what they mean. The remaining three sections deal with the data on which the computations are based, and different ways of quantifying the differences.

3.1 Do the results have relative meaning?

The first question we ask is "Do the results have relative meaning?" That is, does the relative difference between the values correspond to relative relatedness of speech groups? For example, given any phonostatistic method, there is a possible computed value, call it z which represents total sameness. Then, given that the relationship computed between languages a and b is x and between languages a and c is y, and that the value x is nearer to z than y, can we posit that a and b are more closely related than a and c, either diachronically or synchronically?[5] In the case of the degrees of difference and isogloss approaches we can; for the typological approaches we cannot.

Two totally unrelated languages which have never had contact could have the same ratio of vowels and consonants in their phonemic inventories. Thus they would show perfect sameness on one of Fodor's typological measures. A closely related language which had undergone some change would show less relationship. Furthermore, two totally unrelated languages, which have never had contact, could have identical inventories of phonemes. Thus they would show total relatedness on most of Fodor's measures and on the two other typological approaches. If our purpose in measuring phonological relations is to predict or explain communication (intelligibility) between speech groups, or to posit historical relationships, then any method which scores perfect relationship between groups which have no genetic relationship or contact would be of little use. Even their use among speech groups which are strictly related would be questionable, since the control case of no relationship yields inconsistent and meaningless results.

The typological measures would have value, however, in a model which would attempt to predict how well members of one speech community would be able to learn the speech of another community. For instance, the native speaker of English would probably have an easier time mastering the phonemes and sound system of Hawaiian than he would for the closely related language German with its uvular /r/ and velar fricative.

3.2 Do the results have absolute meaning?

An important criterion for evaluating phonostatistic measures is whather the computed results have an absolute meaning or only a relative meaning. That is, do the numbers correlate directly to something in the real world which the linguist understands and is comfortable talking about, or do the numbers have meaning only in relation to one another?

The method of Grimes (1964) is an example of a measure which has relative meaning only. What does it mean to say that the measure of phonological divergence--the absolute value of the slope of the regression line is 10? It has relative meaning only in that 10 represents more divergence than 7 and less than 13. The absolute meaning of 10 is hidden in the method of regression analysis. The linguist has no intuitive feel for what it means to have a divergence measure of 10, or a measure of 7. In contrast to this, the method of Grimes and Agard (1959), is an example of a measure which has absolute meaning. The linguist understands what a degree of difference is and what a correspondence is, so he can easily understand the absolute meaning of the mean degrees of difference--the average number of degrees of difference between the correspondences of two speech groups. The linguist has an intuitive feel for what it means to have a mean degrees of difference of 2, or of 4.

The other measures which lack absolute meaning are the degrees of difference methods of Allen and Hurd (1963) and Larsen (1975). Both of these methods are hybrid lexicostatistic-phonostatistic methods in which cognates and non-cognates are ranked in a single scale for degrees of difference. The result is a measure which has no absolute meaning in either lexicostatistic or phonostatistic terms. For instance, a relation of 75% by the Allen and Hurd method does not mean that the two groups have 75% cognates, nor does it mean they are 75% phonologically similiar. It is impossible to say what 75% does mean. The best we can do is give it relative meaning and say that it represents a closer relation than 70% and a further relation than 80%. The method of Larsen (1975) has the same fault. A coefficient of differences of three does not necessarily mean that the words between the two speech groups have a mean difference of two features. It could mean that half of the words are non-cognate and the other half are identical.

We suggest that if an investigator wishes to combine lexical and phonological relationship, he should compute a lexicostatistic relation with absolute meaning (e.g. percentage of cognates) and a phonostatistic relation with absolute meaning. Then if a single number is required, divise a means of combining the two separate measures. Only in this way can be investigator, and those who later read his materials, assign meaning to the numbers and truly understand them.

The methods which have absolute meaning and the meanings they have are: Grimes and Agard (1959)--mean degrees of difference per correspondence set; McKaughan (1964)--mean degrees of difference between cognate words: Ezard--percentage of similar features between cognate words; Ladefoged (1968)--number of distinctive features different between cognate words; Larsen (1975)--number of phonological isoglosses separating two dialects; and Healey and Healey (1961)--percentage of corresponding phonemes which are the same.

3.3 Do the results measure relations in code or message--correspondence sets or word lists?

If the purpose of one's study is to use phonostatistic measures as predicters of ease of communication between groups, then a measure based on relations in the message would be best. It is not the fact of the correspondences which will affect intelligibility, but the frequency of them. A correspondence which shows up in 10% of the vocabulary will have much greater effect on intelligibility than one which has only two attestations. On the other hand, if the purpose is diachronic, to determine genetic divergence, then the simple fact of the correspondences is probably more important, and a measure based on relations in the code would be best.

The phonostatistic methods which measure relations in the code are the following. The methods of Grimes and Agard (1959), Grimes (1964), Larsen (1975, isogloss approach), and Healey and Healey (1961) are all based on correspondence sets. They measure code relations in that the result in no way takes account of the frequency of correspondences. All correspondences have equal weight in the computation. All the typological approaches (except for Fodor's measures in relation to message) measure relationship between codes.

The remaining methods--McKaughan (1964), Ezard, Ladefoged (1968), Larsen (1975, degrees of difference), and Allen and Hurd (1965)--all measure message relations in that the phonostatistic relations are computed on the basis of word lists. Though word lists do not constitute messages in the same way that the discourses used by Fodor do, they have a considerable advantage over the correspondence set approaches in that the correspondences occur in context. The amount of weight which a given correspondence has in the final measure of phonostatistic relationship will be in proportion to the number of times it occurs in the word list. Most of the rare correspondences will not even figure into the computation since they wouldn't be likely to appear in a word list. The word list frequencies would at least approximate the message frequencies.

We have seen the correlation of diachronic purpose with code measures and correspondence set methods and of synchronic purpose with message measures and word list methods. There is no reason to insist that a diachronic study be limited to a code measure. A diachronic measure which weighted the different correspondences according to their frequency of occurrence could well be of value. We must insist, however, that a diachronic study use a method based on correspondence sets. A phonostatistic measure can reflect genetic, historical relationship only if it is based on items proven to be historically related. Nevertheless, McKaughan suggests that the word list method he uses (1964:103) can profitably be applied before the results of the comparative method are available. However, such an application must be considered preliminary.

For a synchronic study of communication we do insist on a message measure, and a word list method. The wordlist approach for a synchronic study is not a shortcut method, as McKaughan suggests it is for a diachronic study. Rather, it is the best alternative available at present. A study of communication requires a measure of message relations, and a comparison of word lists is the best method we have at present to do this.

3.4 *Degrees of difference or isogloss?*

This question has already been answered for synchronic studies. Since a message measure is required, the isogloss methods are eliminated. The question still remains for diachronic studies. As stated in section 2.1, degrees of difference methods are based on the observation by Austin (1957:544) that phonological change usually proceeds by minimal steps along one phonetic dimension at a time. Thus we assume that the greater the phonological difference between corresponding phonemes, the greater the time depth of divergence. Thus degrees of difference give a relative measure of time depth whereas the yes-no answers of isogloss approaches do not. Thus degree of difference methods are to be preferred. Measures based on isogloss approaches, however, still have the advantage of being much easier to compute.

3.5 *Rank of stricture or distinctive features?*

The question remains whether quantification based on rank of stricture or distinctive features is the best for the degrees of difference approach. Actually neither is better. The decision largely depends on which the investigator feels most comfortable with.

We can, however, offer an evaluation metric for possible degrees of difference scales. Degrees of difference approaches are based on the fact that phonological change usually proceeds by minimal steps. Let us define a minimal step as a change for which we can hypothesize no intermediate stage. We then require that any minimal step in the data is one degree of difference. For instance, for a change from [p] to [w] we would posit two intermediate stages, [ᵽ] then [ƀ]. Between [p] and [ᵽ] we cannot posit an intermediate stage so [p] to [ᵽ] is a minimal step and is one degree of difference. Likewise, from [ᵽ] to [ƀ], and [ƀ] to [w], are each one degree of difference. From

[p] to [w], is then three minimal steps or three degrees of difference.

The evaluation procedure is now illustrated with a few examples. If [ɖ] to [y] is a common correspondence in the data, then is the rank of stricture scale of Grimes and Agard (1959) or McKaughan (1964) better? [ɖ] to [y] is a minimal step so we require that they be separated by one degree of difference. By the Grimes and Agard scale [ɖ] and [y] differ in degree of constriction (1 degree) and all the other dimensions agree. By the McKaughan scale, [ɖ] and [y] differ in point of articulation (2 degrees) as well as in degree of constriction (1 degree). Thus the Grimes and Agard scale, where points of articulation for consonants are made to correspond to the points of articulation for vowels, is better in this case.

A distinctive feature approach is now compared for the same example, [ɖ] to [y]. The features used here are from Schane (1973:26-31). The change from [ɖ] to [y] involves a change from the class of oral cavity obstruents to the semivowels. This already involves a change in two features, sonorant and consonantal, and thus is a difference of two. There are further complications in that [y] would be specified by the vowel features +high, -back, and -round, all of which are irrelevant to the consonant [ɖ]. The consonant features +anterior, and -coronal of [ɖ] are irrelevant to [y]. In the case of [ɖ] to [y], the Grimes and Agard scale is the only one that satisfies the evaluation criterion.

If a change from [k] to [p] was common in the data, and there were no labio-velar stops, an intermediate stage would be improbable. Thus [k] to [p] would be a minimal step. By both of the rank of stricture scales, velar articulation to labial is three degrees of difference, which fails to satisfy the evaluation criterion. A distinctive feature approach, however, meets the criterion. Velar articulation is -anterior and -coronal, while labial articulation is +anterior and -coronal, a difference of one feature.

We have seen that in some cases the rank of stricture scale meets the evaluation criterion and in other cases the distinctive features do. This suggests a third means of devising a degree of difference scale--one constructed by applying the evaluation criterion to the data so that it is always met. The method would be this. First, write down all the correspondences in the data. Then, for each pair of corresponding phones (even if the data are phonemicized, the phonetic realizations must be compared) determine if they represent a minimal step. If not, then posit the necessary intermediate stages, each of which must be a minimal step from the preceeding stage. The degrees of difference is then the number of minimal steps required to get from the one sound to the other. Finally, construct a two-dimensional table with as many rows and columns as there are phones which occur. Label the rows down the left side each with a different phone. Do the same along the top for the columns. Then for all the corresponding pairs, write the degrees of difference in the square where the two sounds intersect. This table then serves as a reference for looking up degrees of difference when later doing the phonostatistic computations.

The evaluation criterion of minimal steps is certainly correct for a diachronic study. For a synchronic study, one could object. Since the degrees of difference is meant to indicate relative ease of understanding, one could argue that a correspondence of [k] to [p] would be harder to understand than [k] to [g], for instance, even though both are a minimal step historically. If, however, [k] to [p] is a regular correspondence that people are aware of, this would not be so. The investigator must judge what is best in his own case.

4. CONCLUSION

4.1 Diachronic studies

In the preceeding evaluation, we concluded that measures with absolue meaning are to preferred. It was determined that for diachronic studies the phonostatistic measure should be based on items proven to be historically related, that is, correspondence sets. It was also suggested that the degrees of difference approaches give a more refined measure of phonological divergence.

The only method discussed in section two which meets all these criteria is the degrees of difference method of Grimes and Agard (1959). Such a method, based on distinctive features rather than rank of stricture, would be just as satisfactory.

4.2 Synchronic studies

In the preceeding evaluation we concluded that measures with absolute meaning are to be preferred. In addition, it was determined that for synchronic studies a measure of relations in the message was required. The word list approaches are the best methods for this currently available.

Of the methods discussed in section two, three meet these criteria: McKaughan (1964), mean degrees of difference for cognate words; Ladefoged (1968), average number of distinctive features different between cognate words; and Ezard, percentage of phonetic similarity between cognate words.

4.2.1 Relating lexicostatistic and phonostatistic measures in predicting intelligibility

We have already discussed how Healey and Healey (see section 2.2.2) computed an "index of expected mutual intelligibility" by multiplying the percentage of cognates by the percentage of identical sound correspondences. Because both the lexicostatistic measure and the phonostatistic measure are percentages, the index is very easily computed. Of the phonostatistic methods recommended above for synchronic studies, only Ezard's method yields a percentage of phonetic similarity. Thus his measure can be directly related to percentage of cognates to predict intelligibility. It would still be possible to use the methods of McKaughan (1964) and Ladefoged (1968); however, the formula for relating the lexicostatistic measure to the phonostatistic one would be more complex. In those methods, a low phonostatistic measure means a high degree of relationship, whereas for the percentage of cognates the reverse is true. Thus, the phonostatistic measure would have to be inverted by subtracting it from some consonant, or dividing it into a constant.

In actual fact, the method of simply multiplying the cognate percentage by the percentage of phonetic similarity is probably too simplistic (disregarding all social factors, of course). This is because the percentage of cognate words that are understood most likely would not equal the percentage of phonetic similarity of cognate words. If the relation between these two percentages is a direct proportion, then the phonostatistic measure will first have to be adjusted by multiplying by some constant. However, it probably is not that simple; the degrees of difference for each word and the frequency of correspondences may have to be considered. The relation between the degrees of difference for cognate words and the likelihood of understanding may be exponential instead of multiplicative. That is, given a pair of words that differs by one degree of difference and another that differs by two degrees of difference, the likelihood that the first pair will be understood may be four times as great as the likelihood for the second pair, rather than twice as great. No doubt, the frequency of occurrence of correspondences is a factor also. That is, if the difference between a pair of cognate words is a common correspondence, it is more likely that it will be understood than a pair which differs by an uncommon correspondence. However, as soon as an individual has heard enough of the speech of another group to be aware of the frequency of certain correspondences, he has had enough contact that we could attribute the understanding to the social factor of contact rather than to the linguistic factor of frequency of the correspondence. There is room for much investigation to be done as to the relation of lexicon, phonology, grammar, and social factors to intelligibility.

4.2.2 Computing the phonostatistic measures from word list data

Here we outline a method by which the computation of phonostatistic measures for a synchronic study can be performed mechanically. If a computer is available, then a program can be written to process the data. If not, the method can still be followed mechanically by any individual who does not mind clerical work, regardless of his linguistic expertise or familiarity with the data. Such a mechanical computation of the phonostatistic measures requires that the linguist first make all the decisions. He must provide decisions on three things: the cognate set assignments, the formatting of the data such that all phones to be compared occur in the same position in the word, and the matrix of degrees of difference for all the possible correspondences.

All of the three phonostatistic methods recommended for a synchronic study are based on a preliminary lexicostatistic analysis of the data. The phonostatistic measures are then computed for pairs of cognate words. Thus the linguist must first make all of the cognate decisions and assign the individual forms to different cognate sets (Sanders 1977).

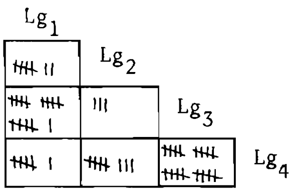
After the cognate sets have been assigned, the linguist must decide which phones in the cognate words are to be compared. He records his decisions by formatting the words in such a way that the corresponding phones in the words occur in the same character positions in the words. This is done by inserting spaces where segments have been lost. If two words have unequal numbers of phones and the extra phones in the longer word are additions, rather than the missing phones in the shorter word being lost, the linguist has a number of options. He may count them as losses anyway, he may remove the extra phones from the longer word, or he may insert a symbol such as "-" in the shorter word to indicate that this position in the word is not to be compared to any other word. Otherwise, when a space is compared to a character it represents the loss of a phone.

This process is illustrated with the word for 'hair' from five Austronesian languages of Papua New Guinea (Hooley 1971:109,112): Wagau ƀəlus, Mapos ƀis, Manga ƀarus, Zenag ƀəluhu, and Towangara niƀeru. By inspecting the data we determine that the ui- on the Towangara form is something that has been added to the root, so it will not be recorded in the formatted data. We assume that all remaining cases of an unequal number of phones between words are due to phonological loss. The vowel of the Mapos form appears to correspond to the second vowel in the other forms. All the decisions are now made; the formatting is as follows.

Wagau:	ƀəlus
Mapos:	ƀ is
Manga:	ƀarus
Zenag:	ƀəluhu
Towangara:	ƀeru

Finally, the linguist must provide the matrix of degrees of difference between corresponding phones. One must remember to include the space as one of the "phones" in the matrix; this is to represent loss of a phone. Whether the rank of stricture, distinctive feature, or minimal step method of computing degrees of difference is used, the matrix should be set up as described in section 3.5.

Once these three things are provided, the phonostatistic measures can be computed mechanically. The "computer", whether it be machine or human, goes down the whole word list and does the following for each item. The word in one list must be compared to the word for every other list. This can be assured by comparing the words in a regular order. The word in the first list should be compared to the word in the second list, then the third list, and so on until it is compared to the last list. Then the word in the second list is compared to the word in the third list, then the fourth list, and so on. By the time the word from the second to last list is compared to the word in the last list, every possible pairing of words has been compared. To make a comparison, first check the cognate set assignments for the two words. If they are not equal then the words are not comparable; take the next word pair. If they are cognate, then they must be compared phone by phone. The first phone of the first word is found on the left hand side of the degrees of difference matrix. The first phone of the second word is found on the top of the degrees of difference matrix. At the intersection of the row and column is found the degrees of difference for the two phones. The degrees of difference for the two phones is added to a running total of the degrees of difference between cognate words in the two languages. For computation by hand, these running sums for all the sets of combinations of languages are best kept by marking tallies in a matrix with a large box for each of the language pairs. That is,



For Ezard's method, a second matrix must be maintained with the running sum of the maximum degrees of difference for each of the phones compared. For the McKaughan (1964) method or the Ladefoged (1968) method, a second matrix must be maintained in which the running total of the number of cognate words is kept.

After all the corresponding phones in all the cognate words have been compared and the degrees of difference tallied, the phonostatistic measures are finally computed by performing a division. In Ezard's method, each value in the degrees of difference matrix is divided by the corresponding value in the maximum degrees of difference matrix. This value is then multiplied by 100 and subtracted from 100 to obtain the percentage of phonetic similarity for cognate words. For the other two methods, each value in the degrees of difference matrix is divided by the corresponding value in the number of cognates matrix to obtain the average degrees of difference for cognate words.

NOTES

[1] For descriptions of the comparative method, see Bloomfield 1933, Hockett 1958, Hoenigswald 1960, Longacre 1967, and Pike 1950.

[2] See a standard statistics text for an explanation of how to perform regression analysis (e.g. Blalock 1960:279-285).

[3] See a standard statistics text for an explanation of how to perform correlation analysis (e.g. Blalock 1960:285-289).

[4] Ezard uses a simplified version of the method he suggests in another paper in this volume, "Tubetube's place among the Milne Bay Province languages: a synchronic study."

[5] By synchronic relatedness I mean the combination of sameness due to genetic relationship and sameness due to contact and borrowing, factors which both contribute to intelligibility.

BIBLIOGRAPHY

Allen, Jerry and Conrad Hurd, 1963. *Languages of the Cape Hoskins patrol post division of the Talasea Sub-district, New Britain.* Pt. Moresby: Department of Information and Extension Services.

——— and ———, 1965. *Languages of the Bougainville District.* Pt. Moresby: Department of Information and Extension Services.

Austin, William, 1957. Criteria for phonetic similarity. *Language* 33:539-44.

Blalock, H., 1960. *Social statistics.* New York: McGraw-Hill.

Bloomfield, Leonard, 1933. *Language.* London: George Allen and Unwin.

Carrol, Vern, 1966. Generative elicitation techniques in Polynesian lexicography. *Oceanic linguistics* 5:59-70.

Casad, Eugene H., 1974. *Dialect intelligibility testing. Summer Institute of Linguistics publications in linguistics and related fields,* No.38.

Chomsky, Noam and Morris Halle, 1968. *The sound pattern of English.* New York: Harper and Row.

Deibler, Ellis and David Trefry, 1963. *Languages of the Chimbu Sub-district.* Pt. Moresby: Department of Information and Extension Services.

Fodor, Istvan, 1965. *The rate of linguistic change.* The Hague: Mouton.

Grimes, Joseph E., 1964. Measures of linguistic divergence. In *Proceedings of the ninth international congress of linguists,* edited by H.G. Lunt. The Hague: Mouton, pp.44-50.

—— and Frederick Agard, 1959. Linguistic divergence in Romance. *Language* 35:598-604.

Healey, Alan and Phyllis Healey, *Dusun dialect comparison.* Ms.

Hockett, Charles, 1958. *A course in modern linguistics.* New York: Macmillian.

Hoenigswald, H., 1960. *Language change and linguistic reconstruction.* Chicago: University of Chicago Press.

Hooley, Bruce A., 1971. Austronesian languages of the Morobe District, Papua New Guinea. *Oceanic linguistics* 10:79-151.

Ladefoged, Peter, 1968. The measurement of cross-language communication. Uganda language survey working paper. *Educational resources information center document ED* 024 920 (*AL* 001 284). Washington: United States Office of Health, Education and Welfare.

Larsen, Robert E., 1975. *Difference coefficients for determining dialect relatedness.* MA thesis, University of North Dakota.

Longacre, Robert, 1967. Systematic comparison and reconstruction. In *Handbook of Middle American Indians,* Vol.5, edited by N. McQuown, Austin: University of Texas Press.

McKaughan, Howard, 1964. A study of divergence in four New Guinea languages. *American anthropologist* 66:98-120.

Pike, Eunice, 1954. Phonetic rank and subordination in consonant patterning and historical change. *Miscellanea phonetica* 2:25-41.

Pike, Kenneth L., 1943. *Phonetics: a critical analysis of phonetic theory and a technic for the practical description of sounds.* Ann Arbor: University of Michigan Press.

———, 1950. *Axioms and procedures for reconstructions in comparative linguistics: an experimental syllabus.* Glendale, California: Summer Institute of Linguistics.

———, 1967. *Language in relation to a unified theory of the structure of human behavior.* The Hague: Mouton.