TABLE OF CONTENTS

A TYPOLOGICAL INVESTIGATION OF THE STRUCTURE OF CONSONANT INVENTORIES

WILLIAM D. RAYMOND

1. INTRODUCTION. The phonetic inventories of the world's languages vary dramatically in the number of segments they contain. The smallest inventories comprise around a dozen segments, while the largest have upwards of 140. Despite this spread, the number of segments used by the majority of languages falls within a much narrower range.[1] Though variation across a factor of ten in inventory size seems considerable, the total number of articulatorily possible segments is somewhere in the neighborhood of 1000—another order of magnitude greater than the largest systems.[2] No language even approaches this number of distinctions in its segment inventory.

Phonetic inventories also vary greatly in composition. The vast majority of articulatorily possible segments are relatively to exceedingly rare—indeed many may not occur at all. However, the members of a small subset of the total possible segment set occur widely. As a consequence, all languages share some common segments: no two languages have disjoint segment inventories.

With observations such as the preceding ones we begin to explore the makeup of phonetic inventories. The focus of this paper is to explain and to begin to quantify inventory structure—specifically that of consonant inventories. Within a framework provided by the concept of phonetic space, organizing principles motivated by cross-linguistic phonetic observations will be presented that seem to apply universally to phonetic inventory structure. Using the framework of phonetic space, a variety of models will be suggested which will allow the principles to be investigated. The models are spatial representations of segment space, some of which have metrics, or measures, defined over them allowing the distance between segments to be determined. The investigations will seek both evidence for the principles asserted and directions in which statement of the principles can be made more precise. Through this process principles will be embodied within the models, giving them predictive power for the composition of individual inventories and perhaps for some phonological processes.

The approach I take in investigating models and principles is a typological one, in which the composition of inventories from a large set of genetically and areally diverse languages will be used to study inventory structure. The language data employed were taken from Maddieson 1984, in which the author details his digital compilation and subsequent typological investigations of segment inventory data from 317 languages. The data base of segments—known as UPSID—is compiled from languages which 'have been chosen to approximate a properly constructed quota sample on a genetic basis of the world's extant languages,' with one language chosen from each small family grouping (Maddieson 1984:5). Maddieson's sample is thus in accord with the recommendations of Bell 1978 for construction of a quota sample based on genetic groupings, given the accuracy of current genetic classifications.

I have transferred a portion of the data from Maddieson on each language (less !Xũ in order to avoid its complex system of clicks) to computers of the Statistical Computing Center at the University of Colorado. SAS (a statistics and graphics package) was used to perform statistical analyses and produce the output contained in the tables of the appendix. The data I have tabulated for each UPSID language (by counting segments) consist of the number of consonants, vowels, stops, stop series, stop places, affricates, fricatives, labials, palatals, and ejective stops.

2. FRAMEWORK FOR THE MODELS: PHONETIC SPACE. Fundamental to the models and the paper is the concept of phonetic space. Such a construct is implicit in any phonological feature system. We can assume that the segments of a language's phonetic inventory lie in an abstract space that will accommodate all segments that might appear in natural language. The dimensions of phonetic space can be variously defined using articulatory, acoustic, auditory, or even perceptual parameters, singly or in combination. Sets of values for these parameters specify the location of segments within the space.

As an example, the articulatory dimensions of PLACE and MANNER can be used as the ensions specifying the coordinates of consonant segments in the phonological space. MANNER is over term for many articulatory parameters—such as ASPIRATED, STOP, AFFRICATE, or TAP—ch taken together with PLACE uniquely identify a given consonant segment. Both MANNER and CE are thus multivalued parameters. Vowels are readily specified by using articulatory (PLACE = H FRONT, e.g.) or acoustic (formant frequencies) parameters. As a third example, the Chomsky-le feature system (Chomsky & Halle 1968) is a combination of broad, binary-valued articulatory . CORONAL) and acoustically based (e.g. STRIDENT) parameters. The use to which a framework honetic space is to be put dictates appropriate dimensions for it.

Within the framework of phonetic space we can begin to suggest models and identify principles ch give segment inventories structure. It is the whole space we will be exploring, while the npositions of the individual inventories are the samples by which we probe its organization.

3. ORGANIZING PRINCIPLES: METRICS IN PHONETIC SPACE. To begin to investigate the cture of phonetic space, we need general principles, motivated by the observations others have le and perusal of available data, which can be used as starting points. In this section, I will gest three.

Maddieson discusses the concept of 'phonetic distance' as a possible factor that might help lain the structure of segment inventories. He notes that certain segments, which vary minimally ng one (in his examples acoustic) parameter and hence are perceived as being very similar, do not occur in any languages in his sample (or only do so rarely). For example, no language (in SID) contains both a voiceless lateral fricative and a voiceless lateral approximant (Maddieson 4:14). We can generalize from such observations to argue that there is a minimum 'separation' is necessarily maintained between any two segments in an inventory, an observation which, ugh not novel, precise, or universally accepted, has an intuitive appeal. The concept of separation plies it is possible to measure how 'close' two segments are to each other. In a mathematical del, the means by which a distance measure is defined is called a metric, and a space in which asures can be made is termed a metric space.

A principle of minimum separation could help explain why attested inventories are so much ller than the theoretical maximum. If an inventory contains a very large number of segments, the ustic differences between some pairs of members will become too small to be clearly produced or dily perceived, since at some size it would necessarily have to include pairs of segments which are er observed to contrast. Of course, other, cognitive limits are most likely also involved and may vide the better explanation of size limits globally.

Given the preceding observations, I will accept minimum separation as the principle:

(P1) Any pair of segments from a language's inventory must be separated by a minimum phonetic distance.

The problem with quantifying or testing this principle is to define how phonetic distance is to neasured given a pair of segments. We are making the fundamental assumption that phonetic ce is organized so that it is POSSIBLE to measure the distance from one point in it to any other nt, i.e. that it is a metric space. For the time being, I will not attempt to define the dimensions r which a metric could be imposed, but will return to this issue later, at least for some spaces of phonetic space.

If segments should not be too similar, we might hypothesize that indeed they should be as erent as possible, that is, that the acoustic and/or articulatory differences among the members of ment inventories would be maximized in any language. As has been frequently noted (by, ong others, Ohala 1980, Maddieson 1984, Lindblom 1987), this is not the case. If it were true, might expect to find languages whose consonant inventory consisted, for example, of segments h of which had a manner of production and a place of articulation both differing from those of other segments. Examination of language data turn up no instances of this type of inventory ribution.

Might segment inventories, then, be random subsets of the phonetic space—so long as they do not violate P1? If this were the case we would expect to find pairs of inventories having no segments in common, since the probability of two random samplings of segments chosen from the pool of theoretically possible segments overlapping is not close to certainty.[3]

What we do find is that some segment types occur with great frequency in languages, and other less common ones are employed less often. In structuralist terminology this is a frequency measure of markedness. A strict hierarchy cannot be constructed, though; variation within some parameters is possible. For example, /b/ occurs in 62.8% of UPSID languages, while /p/ is much more common, occurring in 82.6%; however, we cannot conclude with certainty that if a language contains a /b/ it will contain a /p/.[4] Maddieson notes other similar hierarchical tendencies (Maddieson 1984:13-15) for individual segments, but they provide only local, sporadic constraints, or limited hierarchies (e.g. /p/ < /k/ < /t/ in almost all cases, though cf. Bell 1984 for a more thorough discussion of these and similar implicational hierarchies).

Perhaps we should focus on segment classes. Maddieson presents what he calls the 'modal' inventory of 21 segments— 21 is the modal inventory size in the UPSID sample—containing the 20 most common segments and one other chosen from among a small set of less common but similarly frequent segments, such as /z/, /ts/, /x/, /v/, or /ʤ/. One modal inventory is the one displayed in Fig. 3 (in §5 below). Maddieson's modal inventory is very similar to the 'basic' set proposed by Lindblad 1980 (discussed in Lindblom 1987) from a sample of 189 languages. Though Maddieson provides no comprehensive list of segments ordered by frequency, a survey of his data indicates that the modal segments are very much more common than almost all other segments, each being found in from over 100 to almost 300 of the UPSID languages. The non-modal segments are all found in fewer than 50 of the UPSID languages each, and in most cases in fewer than 10.

Lindblad subdivides non-basic segments into 'elaborated' and 'complex' sets, where elaborated segments use a non-basic PLACE (e.g. UVULAR or PALATAL) or MANNER (e.g. a non-pulmonic-egressive phonation type or a secondary articulation), and complex segments involve two or more elaborations. He thus defines a markedness hierarchy based on structural criteria of number of articulatory gestures and place of articulation. His sample shows that elaborated segments only begin to be used by languages when their consonant inventory size reaches about 10, and complex segments only in inventories of at least 25-30. Thus, the presence of complex stops implies the presence of elaborated stops, and elaborated stops imply basic stops. Significantly, this suggests that segment class hierarchies can be investigated as a function of inventory size.

We thus have an intimation of a hierarchy of nested classes. It is not clear at this point whether we can subdivide the classes finely enough to include only single segments, or whether indeed a hierarchy can be made precise. However, I will use the above observations as motivation for the principle:

(P2) There is a tendency for segments to be arranged in a hierarchy of nested segment classes, with the most common segments in the lowest class and less common ones in higher classes. If languages L_1 and L_2 have inventories of size I_1 and I_2, respectively, with $I_2 > I_1$, L_1 will tend not to contain segments from any segment class higher than the highest segment class comprising L_2.

P2 says that language inventories will, as a tendency, include segments from classes higher on a putative segment class hierarchy only as inventory size increases. In fact, given the boundedness of these classes, inventories MUST occupy the outer segment classes as their size increases or they would eventually violate P1.

Keeping in mind the notion of a markedness hierarchy of segment classes, let's revisit the notion of maximization of phonetic distance. We have seen that maximization clearly does not occur over the entire phonetic space, so if the notion is valid at all it must be constrained in some way. If we accept P2, one constraint on maximization would be that as inventory size increases,

https://scholar.colorado.edu/cril/vol12/iss1/1
DOI: https://doi.org/10.25810/8d78-9s92

inventories will occupy an ever larger subspace of phonetic space as dictated by a hierarchy of nested segment classes. In other words, inventories grow outward from a fixed, central 'core.' But within the bounded subspace occupied by a given language, the distribution of segments is clearly still not random. For example, one type of elaborated segment is an aspirated stop. When aspiration is used as a distinctive feature, it is generally distinctive at every place of articulation at which non-aspirated stops occur.[5] Similar observations can be drawn about other consonant types.

We can explore the distribution of segments within the nested classes by looking for slices of phonetic space orthogonal to the nested classes identified so far. Such classes will be present regardless of inventory size, and the number of members from such a class will be a constant proportion of the total number of segments. Lindblom 1987 identifies obstruents as such a class. He shows that as the total number of consonants increases, the ratio of obstruents to non-obstruents (sonorants) remains fairly constant, at 70% obstruents to 30% sonorants. This is only a gross measure of distribution, but it suggests that segments continue to be 'spread out' over the subspace they occupy as the size of the space grows; they do not 'clump up,' becoming predominantly obstruents, for example.

We can thus speculate that the following principle holds:

> (P3) Inventories tend to maximize separation of segments within that part of the phonetic space bounded by the highest segment class H_i from the hierarchy of P2 from which they contain members, and within each class contained by H_i. Maximization is achieved by maintaining constant proportions of segments in every nested class from each of a set of subspaces which partition the nested classes.

Lindblom 1987 gives what can be considered a functional explanation of this principle. He asserts that segments can be said to inhabit an abstract space, which he calls an 'acoustic-perceptual space,' and claims that the organization of inventories is motivated by the principle that segments included will have 'sufficient perceptual contrast at acceptable articulatory cost' (Lindblom 1987:8).

We now have two principles, P2 and P3, which constrain inventory growth by asserting an even distribution of segments among partitions created by two sets of orthogonal subspaces. P1, however, refers directly to phonetic distance. We must thus return to the problem of measuring separation. While we are not sure how to measure distance between ANY two segments, there are some subspaces which lend themselves more readily to a definition of a measure. Let's turn our attention to one such subspace.

4. A BACKGROUND MODEL: PHONETIC DISTANCE IN VOWEL SUBSYSTEMS. Liljencrants & Lindblom 1972 discuss a model used to predict the makeup of vowel systems of a given size. Vowel systems are a good starting point for exploring the concept of phonetic distance because the subspace in which simple oral vowels are produced is conceptually much more continuous and homogenous than the consonant subspace.

As a starting point, consider the chart traditionally used by linguists to represent vowel systems (Fig. 1). This is essentially a phonetic space with articulatory dimensions. A metric could be defined on this space in terms of physical distance from the cardinal vowels at the vertices of the space.

What makes vowels distinct in language, however, is not their articulatory separation per se, but more importantly their acoustic separation. Acoustically, vowels can be represented in terms of formant frequencies. If we construct the vowel space using the first three formants as dimensions, we have the space shown in Fig. 2. The most obvious metric for this space is a three-dimensional Euclidean distance measure within the formant dimensions.

A further refinement of the model comes from the observation that vowel quality (and hence salience) is not just a function of acoustic quality, but is influenced by perceptual factors as well. This is taken into account by Liljencrants & Lindblom by mapping the formant frequencies into the perceptually more realistic, quasilogarithmic mel scale. The vowel space with a Euclidean distance metric defined in terms of mel values is essentially the space the authors use for their model.

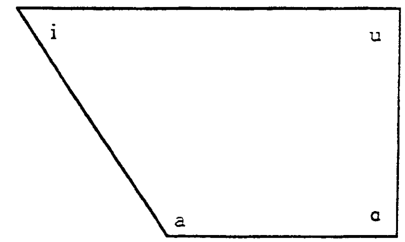


FIGURE 1. The articulatory vowel space.

The hypothesis Liljencrants & Lindblom wish to test is that vowels tend to maximize their distance from one another, i.e. that phonetic distance, as measured in terms of formant frequencies, is maximized in the vowel space, regardless of the number of vowels in a language's vowel inventory. In analogy to electrical charges, they assume the vowels repel each other with a force inversely proportional to the square of their distance. When their model is applied to inventories of a given size, it predicts configurations similar to natural inventories with the same number of segments. For example, a natural observable tendency the model duplicates is that in attempting to maximize their distance from one another, vowels align themselves along the periphery in small inventories; interior vowels are only introduced when the inventory reaches a critical size. (Taking even more perceptual factors into consideration, Lindblom 1986 refines the model further.)

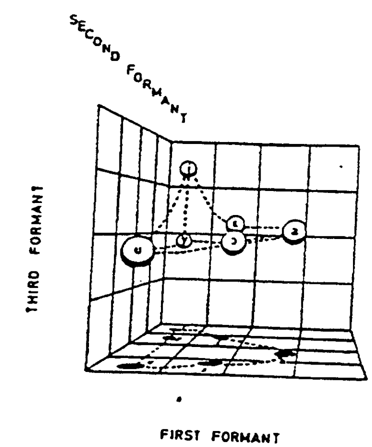


FIGURE 2. The acoustic vowel space (from Lindblom 1986)

Not all vowels are simple ones, however. Wright 1986 considers nasalized vowels, and concludes that their separation from oral vowels must be recognized as an additional dimension. The total nasal vowel space is found to be smaller than that of oral vowels—a contraction of the oral vowel space. This corroborates the observable phenomenon that nasal vowels are perceived as shifted inward and (for high vowels) downward from their oral counterparts. We thus have an xplanation for the typological observation that languages generally contain more oral vowels than asal ones, since the size of the nasal subspace is smaller than the oral subspace and maximization f distance across both subspaces will necessarily mean putting fewer segments into the nasal ıbspace.

With the introduction of nasal vowels, the vowel subspace starts to become less dimensionally omogenous. It would become even less so if we wish to include diphthongs and vowels with ther manners of articulation. This same problem is encountered more quickly in consonant ıbsystems.

5. A METRIC MODEL FOR THE CONSONANT SUBSPACE. We can apply reasoning to onsonant inventories similar to that which was used for vowel subsystems and construct a onsonant subspace. As a first approximation consider the traditional organization of a consonant ıventory as a crude representation of a phonetic space, organized within the dimensions of PLACE, cross the top, and MANNER, down the side (cf. Fig. 3).

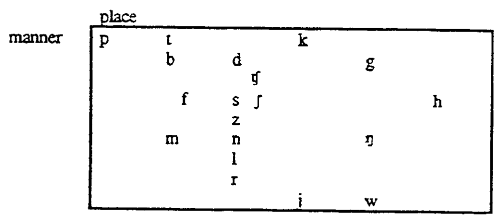


FIGURE 3. A modal consonant inventory.

Clearly, consonants are inherently ıore discrete than vowels and ıbtypes are more important, so a lobal measure of distance is more roblematical. We are used to ıinking of consonants as discrete oints. In actuality, they are henomena which occur over time nd involve the entire vocal tract. ariations in the combination and ming of articulatory movements rovide continua of acoustic outputs ong many of the parameters of the phonetic space. A subset of points from these continua are nployed by a language as the prototypical acoustic forms from which the language's segment ventory will be constructed.

By way of illustration, consider the English bilabial stops /p/ and /b/. Both are made with the ıme articulatory gesture, the significant difference between them being whether or not voicing curs during the closure. Since voicing transition and articulator closure are independent actions, eir relative timing is variable. In fact, as we know, voicing does not begin immediately on lease of closure for /p/ in English, but is systematically delayed in most environments, making a oser transcription of this sound /p^h/. While /p/ and /p^h/ do not contrast in English, Voice Onset me (VOT) is an example of a continuous parameter that is exploited by many languages to oduce distinctive phonemes, such as those transcribed as /b/, /p/, and /p^h/. We could readily fine a distance measure for the continuous parameter VOT and determine the phonetic separation segments lying along this dimension (Neary & Hogan 1986). Unfortunately, the VOT mension can only be applied locally and is only one of many parameters which together fall ıder the cover term MANNER.

Is there a way to introduce articulatory dimensions which will subsume other manners of stop ticulation? Maddieson documents thirteen different types of stop articulation for consonants ıich do not involve secondary or double articulation (shown in Fig. 4). (I am disregarding cks, as I will throughout this paper, in order to simplify analysis, and I am using 'stop' to refer oral, unaffricated stops.)

We have shown that (a) voiceless, (b) voiced, and (c) aspirated segments vary along the nension of VOT. To this dimension we can add (d) and (e). The other articulations involve ntinuous parameters describing an articulatory gesture other than the primary closure. There are ree sites for these gestures, involving (i) the velum, (ii) the glottis, and (iii) the larynx.

(a) (plain) voiceless
(b) (plain) voiced
(c) (post-) aspirated voiceless
(d) pre-aspirated voiceless
(e) long (of any of 1-3 only)
(f) breathy voiced (murmured)
(g) voiceless with breathy release } Laryngeal sounds
(h) laryngealized (creaky) }
(i) pre – glottalized }
(j) implosive } Glottalic sounds
(k) ejective }
(l) pre-nasalized
(m) post-nasalized.

FIGURE 4. Types of simple stop articulation.

(i) MANNERS INVOLVING THE VELUM. Stops that are (l) pre-nasalized or (m) post-nasalized involve lowering the velum. It is the relative timing of a lowering of the velum and the period of primary closure that is significant. If the velum is lowered before release of the primary closure we get (l), and if the velum is raised before the release of the primary closure we get (m). (Note that lowering the velum throughout the primary closure gives a nasal stop, and if the velum remains raised we have a plain stop.) I will call this dimension Relative Release of Velum (RRV).

(ii) MANNERS INVOLVING THE GLOTTIS. Manners (f-i) involve varying degrees of glottal closure, affecting the mode of vocal cord vibration. Loosening the vocal cords produces (f) breathiness, tightening them produces (h) laryngealization, and complete closure of the glottis gives (i) pre-glottalization. Maddieson does not distinguish between (h) and (i), lumping them both under 'laryngealized.' I am not sure whether this is predominantly a problem of language description or whether the difference between (h) and (i) is never distinctive. The dimension along which (f), (h), and (i) vary will be called Degree of Glottalic Closure (DGC). Note that (g), in addition to having a DGC component, varies from (f) along the VOT dimension.

(iii) MANNERS INVOLVING THE LARYNX. Finally, (j) implosive and (k) ejective stops involve movement of the larynx to effect an airstream mechanism other than pulmonic egressive. Moving the larynx up produces (k), and moving it down produces (j). This dimension will be referred to as Relative Laryngeal Movement (RLM).

The continuous subspaces defined by the above dimensions are interlinked. Figs. 5 & 6 depict the possible relationships between the four dimensions of Voice Onset Time (VOT), Degree of Glottalic Closure (DGC), Relative Laryngeal Movement (RLM), and Relative Release of Velum (RRV). (The segments in parentheses are not attested in UPSID, but if they did occur, they would be positioned as indicated.)

There are some questions about the representations in Figs. 4 & 5. I am at this time uncertain how ejectives vary in VOT; they may be positioned at either head of the arrow separating them. The sounds classified as implosives seem to vary in the degree of their glottalic closure; hence the position of implosives is represented as a range along DGC (cf. Pinkerton 1986). (Though the two positions below the top /ɓ/ would be more accurately represented respectively as /ɓ̰/—a laryngealized implosive—and /ˀɓ/—a preglottalized implosive.[6]

Notice that dimensions DGC, RLM, and VOT are depicted (and currently conceptualized) as bounded and linear. This is not the case for RRV, in which we can move from /m/ to /b/ and back to /m/ again without changing "direction." The diachronic implication of the circular structure of RRV is that we should not see sound changes directly from /b/ to /m/. That is, as the feature [+nasal] or [-nasal] assimilates to a homorganic stop of opposite value, there should be an intermediate stage of pre- or post-nasalization. It may turn out that various other connections exist

https://scholar.colorado.edu/cril/vol12/iss1/1
DOI: https://doi.org/10.25810/8d78-9s92

between points in the same or different subspaces. These too might be revealed by diachronic pathways.

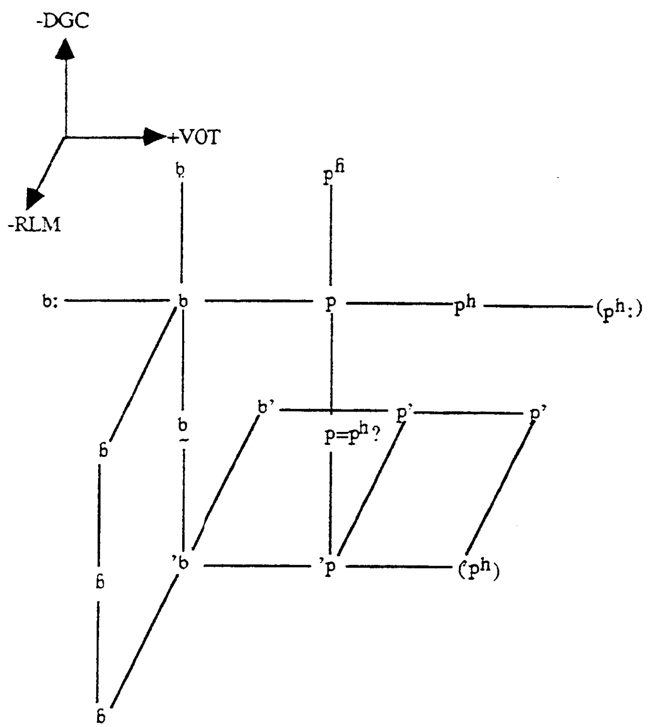


FIGURE 5. A view of the consonant space for bilabials using dimensions Degree of Glottalic Closure (DGC), Relative Laryngeal Movement (RLM), and Voice Onset Time (VOT).

Relative distances between some pairs of segments from different dimensions can be determined synchronically by looking at inclusion restrictions on the pairs. For example, Maddieson notes that no language distinguishes between glottalized and ejective stops. He also states that implosives are frequently not distinguished in the literature from the laryngeal sounds (h) and (i), and I can find no examples in UPSID of a language containing both implosives and a laryngeal sound. The significance of inclusion restrictions along a measurable dimension is that they define a lower bound on the minimum separation that must be maintained in that dimension in accordance with P1.

6. AN ABSTRACTION OF THE MODEL: A TYPOLOGY ON PHONETIC SPACE. We are comfortable with the notion of distance because we live in a Euclidean space in which distance is readily measured; however, it is not necessary for a space to have a distance measure—a metric—defined on it for it to have internal structure. We can borrow the mathematical notion of topology

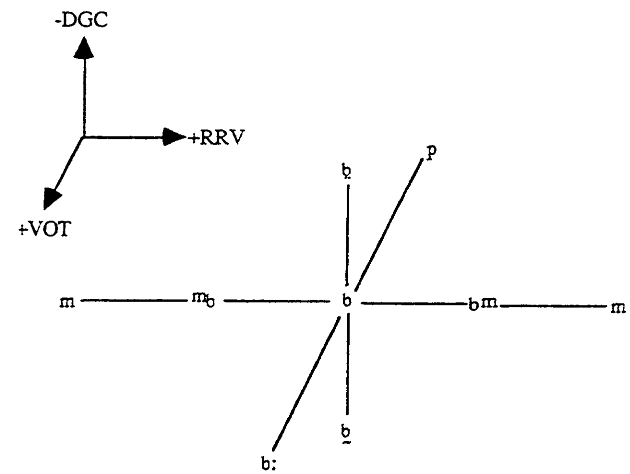


FIGURE 6. A view of the consonant space for bilabials using dimensions Degree of Glottalic Closure (DGC), Voice Onset Time (VOT), and Relative Release of Velum (RRV).

to give structure to a space: a topology is a way to describe which subsets of a space are, in some sense, close together without being able to quantify their nearness.

It is at this point not clear how one could rigorously define a topology on phonetic space, but topological spaces can be characterized by two important properties. One is connectedness: either the space is a single 'piece,' or can it be divided into two or more separate pieces. The interlocking dimensions of the subspace described in §5 form a connected subspace. There is, however, plenty of evidence that phonetic space is not connected globally. One example of this is easily seen along the dimension PLACE. Since there seems to be no way to pass continuously from a bilabial stop /p/ to an alveolar stop /t/ by choosing possible segments, because different articulators are involved, there is a discontinuity in phonetic space along this dimension.

The second topological property of importance is a measure of how many 'holes,' if any, there are in a space. In §5 we saw that the RRV subspace was circular. This means there is a hole in the middle of this subspace. We cannot go directly from /b/ to /m/, but there is a way to get between these two segments through pre- or post-nasalization, so the space is still connected.

Establishing these topological properties depends largely on diachronic processes that should be investigated, but this lies beyond the scope of this paper. Mention should be made, however, of at least one implication of these properties for P1. Across discontinuities, P1 does not obtain. For example, if there is a discontinuity between labial place and alveolar/dental place as postulated above, there should be no restriction on co-occurrence of /θ/ and /f/ or of /ɸ/ and /θ/, as there is between /f/ and /ɸ/ (Maddieson 1984:14), even though all three pairs are similar acoustically.

The above claim can be tested. The following counts of occurrences are found in UPSID:

(1)				(i)	(ii)	(iii)
	/f/	/ɸ/	/θ/	/f/ & /θ/	/f/ & /ɸ/	/ɸ/ & /θ/
	135	21	18	9	4	1

The Fisher test gives the following values for the pairs (i)-(iii):

(2)	(i)	(ii)	(iii)
	/f/ & /θ/	/f/ & /ϕ/	/ϕ/ & /θ/
	432	1563	60

Of these only case (ii) will fail the null hypothesis that the pair is independent for $\alpha = .005$.

We should note that disconnectedness like that between labial and dental/alveolar places could merely a local phenomenon. For example, adding a secondary articulation like velarization to a ial segment could provide a connection to the velar place, and from there paths exist to an eolar articulation (such as palatalization). This does not change the consequences for P1 cussed above.

7. RETURN TO A DISTRIBUTIONAL MODEL: A PROBABILITY SPACE. We can quantify the ion of distribution of segments as the probability that a segment, or class of segments, occurs in articular region of phonetic space, such as the partitions of §3. A space in which probabilities be measured—a probability space—assumes an underlying metric space, though we need not cify the measure. A probability space could tell us a good deal about the composition of onetic inventories. It would allow us to predict the composition of an inventory given an entory size, within the degree of resolution to which the probability function is defined.

I have discussed Lindblom's (1987) conclusion that the proportion of sonorants to obstruents nains fairly constant—at a ratio of 70% obstruents to 30% sonorants—as the number of nsonants increases. This is the starting point for defining the probability function for phonetic ce as a function of inventory size. I have continued this line of inquiry by extending it to other ospaces.

As a check of Lindblom, Table 1 shows the relation between number of consonants and the nber of stops plus affricates plus fricatives (i.e. obstruents) in the inventories of the languages UPSID. The solid line (on this and all similar graphs) indicates a least square linear regression of all data points. The dashed lines enveloping the solid line show a 5% confidence interval for fit. Each dot on these graphs represents at least one language, but possibly many.[7] The linear of the UPSID data for obstruents is in good agreement with Lindblom, giving an obstruent ratio .657 or about 66%.

When we remove the fricatives from the obstruents (cf. Table 2), we find that the ratio of affricated oral) stops plus affricates to total consonants is 44%. Removing the affricates, the portion of stops to total consonants is 35% (cf. Table 3). (The r-square value for this last fit is 68 for a probability of .0001, indicating a good linear fit. Corresponding values for Tables 1 2 are similarly acceptable.)

We can conclude that for a given consonant inventory size, 35% of the consonants, on erage, will be stops, 44-35 = 9% will be affricates, and 66-44 = 22% will be fricatives. Since linear fits are good approximations for the data, we can make an initial assumption that the portion of segments from each of these subspaces is independent of inventory size. These ospaces are thus candidates for the subspaces of P3 which are orthogonal to the nested classes l partition them.

The three orthogonal slices examined above are all along the MANNER dimension, though we equally well take a slice along the dimension of PLACE. Table 4 shows a linear fit of total nsonants to number of labially articulated segments. (The r-square value for the fit is .9050 for obability .0001.) The slope of the linear fit indicates that the proportion of labials in a consonant entory should be 17%.

If we assume a normal distribution for the number of segments in each of these classes, then closeness of any one prediction will depend on the shape of the normal curve, characterized by standard deviation. If the standard deviations are large, we should not expect an exact diction for a set as small as labial stops, and indeed there seems to be considerable variation in number of labials in any individual language, though a measure of the variance was not made.

Nevertheless, labials—and, it may turn out other subspaces along the PLACE dimension—could be considered as orthogonal subspaces.

Are the orthogonal subspaces along the MANNER dimension and the PLACE dimension independent? The probability of two independent events co-occurring is the product of their individual probabilities. If the events 'stop' and 'labial' are independent, then the probability of a labial stop should be

(3) P(labial stop) = P(labial) x P(stop) = .17x.35 = 6%.

To test this, I calculated the average for a subsample of UPSID consisting of every twelfth language (still an approximate quota sample). The average number of labials for these 25 languages turned out to be 7.3%, a little higher than predicted, but close enough to encourage the approach.

Lindblom's data for elaborated and complex segments supported P2. His observations regarding these classes also fit into a probability space. If an inventory has less than 10 consonants it will not likely have elaborated or complex segments, and elaborated segments will only be introduced when the total consonant inventory reaches 25-30. But as these classes are introduced they will continue to be present in constant proportion of the total. (I will take this as a first approximation from Lindblom's graphs—Lindblom 1987:3—though he does comment on the relationship.)

Tables 5 and 6 are plots of affricates and fricatives, respectively, versus total consonant inventory size. Notice that they differ in shape from the plots of Tables 1-4. In particular, affricates and fricatives continue to be frequently unrepresented in inventories of 20 or more and affricates in inventories of 30 or more. This suggests that these classes behave as elaborated and complex segments do, viz. they come into use as inventory size increases. (In fact, affricates are a type of Lindblom's elaborated segments.) In this sense, fricatives and affricates behave like the nested classes of P2.

But in addition to languages that continue to have no affricates or fricatives for inventories of twenty or more, there are languages of the same size that do have these segment types. And there is a linearly increasing upper bound on the number of affricates and fricatives in these latter languages as inventory size increases. This may represent a variation in the hierarchy of P2, with some languages treating fricatives and affricates as higher on the markedness hierarchy and others including them as orthogonal subspaces and representing them proportionally regardless of inventory size. Further investigation is necessary to determine which scenario is justified.

If an inventory's proportion of obstruents is constant and some subsets of obstruents are not represented until some minimum inventory size, the proportions of the subparts cannot be linear. Tables 7-9 show quadratic fits of stops, affricates, and fricatives. The fits for affricates and fricatives curve upward, while that for stops curves slightly down. This is as we would expect: as affricates and fricatives come into use, the proportion of stops begins to decrease, keeping the proportion of obstruents constant. It is difficult to quantify these functions, however, since the amount of data for large consonant inventories—where the difference becomes significant—is meager.

8. A METRIC IN THE AGGREGATE: STOP DENSITY. The constant distribution of segments in some broad classes suggested P3 on rather imprecise evidence. Is there a way to look inside any of the classes to see if segments are maximizing their separation from one another? We assumed maximization of separation is accomplished by even distribution over the partitions created by the intersection of the classes of P2 and P3, but the partitions we have identified do not yet cover the phonetic space finely or evenly. However, if maximization of separation creates a uniform distribution, then the number of segments per unit volume—segment density—will be constant. This is like looking at segment distribution in each unit of volume as a substitute for the partitions. We are once again assuming an underlying metric space, though we do not specify the metric.

How do we measure the volume of phonetic space? Let's confine ourselves to stops to make tabulation of the data manageable. This class should be an adequate substitute for an orthogonal

https://scholar.colorado.edu/cril/vol12/iss1/1
DOI: https://doi.org/10.25810/8d78-9s92

subspace, since for smaller inventories the proportion of stops is approximately constant. Stops can be characterized by the dimensions PLACE and MANNER. As noted earlier, manners are often represented at each place of articulation, and stops with the same manner can be termed a series. For an inventory with S stops at P places in Σ series, the size of the stop subspace—the number of logically possible stops—can be assumed to be

(4) $$S_{poss} = P \times \Sigma,$$

and the density of the stop subspace is

(5) $$D = S \,/\, S_{poss}.$$

If all series employed by a language tend to be filled, the density of the space can be said to remain constant and phonetic distance is maximized with increasing inventory size. If, on the other hand, elaborated series are only sparsely occupied, then the mean phonetic distance is increasing as inventory size increases.

Table 10 plots the number of stops versus the number of logically possible stops. Density remains fairly close to 1, but varies from saturation by a constant amount as the number of stops increases. The proportion of stops to possible stops from Table 10 is about .87, indicating that stop density does not remain constant with increasing inventory size, but also does not vary appreciably.

The ratio of stops to possible stops might be an artifact of the method used to determine the size of the stop subspace S_{poss}. The method assumes it is possible to have complex segments involving all elaboration types at every place of articulation. This seems less likely as the number of series and places increases. It might make sense to limit the possible complex segments to those attested in Maddieson as an approximation of the set of cross-linguistically possible elaborations. As an example, there are 14 attested palatalized stop segments in UPSID over four places of articulation, as shown in Fig. 7.[6]

In the space containing these six series (plain voiceless, aspirated, plain voiced, breathy, prenasalized, and ejective) and four places of articulation (bilabial, dental, alveolar, and velar), we would calculate 24 possible palatalized segments using (f). Making the adjustment to attested segments would shift the points for languages with large numbers of stops closer to saturation, allowing us to conclude more forcefully that phonetic density remains constant and phonetic distance is maximized. This adjustment, or lower bound, was not calculated, as the amount of work involved did not seem justified.

pʲ	t̪ʲ	tʲ	kʲ
pʲʰ		tʲʰ	kʲʰ
bʲ	d̪ʲ	dʲ	gʲ
b̤ʲ			
		ndʲ	
			k'ʲ

FIGURE 7. Palatalized stop segments by series in UPSID (from Maddieson 1987).

On the other hand, eliminating elaborations because they are not attested may not be acceptable, since the fact that it is not attested may only mean it is rare, not typologically impossible. In this case the variation in the proportion of stops to logically possible stops may be real, and density may indeed decrease as inventory size increases. Adjustments could be made to density measures based on articulatory, auditory, or perceptual grounds to maintain constant density, or a reformulation of P3 might be necessary.

9. CONCLUSION. Consonant space is complex. We can at this point glimpse only outlines of its gross structure, but some conclusions can be drawn.

There are some indications that phonetic distance plays a factor in determining the content of consonant inventories. Minimum separation plays a role mostly at a very local level in the composition of inventories. The formulation in P1 prohibits the inclusion of some pairs of very similar segments. At the global level, the effect of P1 is felt in ensuring that as inventory size

increases the portion of phonetic space that an inventory exploits will increase. There is also some indication that maximization of separation occurs. In the stop subspace the density of stops over manner series and place of articulation remains fairly uniform as the size of the stop subspace (i.e. the number of manners and places exploited in an inventory) increases. Further precision in the investigation of phonetic distance must await selection of appropriate dimensions over which a metric can be defined. The dimensions suggested herein for a portion of the consonant subspace need to be quantified to determine whether the physical measures which can be defined on them have explanatory power. In lieu of specific measures, topological properties can give insight into the outlines of the spatial structure.

There is support for the existence of a hierarchy of nested classes as defined in P2, though this principle may need to be amended to allow the hierarchy to vary cross-linguistically. The nested classes identified in this paper as universal are simple, elaborated, and complex consonants, as defined by Lindblom 1987. Some languages may also treat fricatives and affricates as nested classes ranked in the hierarchy above simple consonants within the class of elaborated consonants.

There is also support for a set of subspaces orthogonal to the nested classes, as described P3. Obstruents are the best candidates, since they are represented proportionally in segment inventories independently of inventory size. There is evidence that labials are also represented in constant proportion to inventory size, and thus also exemplify P3. Some languages also seem to treat fricatives and affricates as orthogonal subspaces by including them proportionally in their consonant inventories. Fricatives and affricates thus may represent either the nested subspaces of P2 or the orthogonal subspaces of P3, suggesting that the classes of P2 and P3 can delineate typological variation.

Taken together, principles P1-3 require an orderly change in inventory composition with increasing inventory size. Though these principles are only first approximations, if they can be refined and yet remain applicable, then composition could be shown to be predictable and explainable, in a probabilistic way, as a function of inventory size.

NOTES

[1] 70% of the sample in Maddieson 1984 have between 20 and 37 segments.

[2] Maddieson 1984 lists about 750 different segment symbols used for describing the inventories of the 317 languages in his UPSID database. This is a conservative lower bound for the total. A quick count of the number of IPA symbols in Pullum & Ladusaw 1986 gives about 100 vowels and 100 basic consonants. The basic consonants can be modified by perhaps 12 different methods (secondary articulations and variations in phonation, e.g.), though these do not apply uniformly to the entire set of consonants and can apply in combination. A generous upper limit seems to be 1500 possible segments.

[3] The probability of two sets of size m selected randomly from a population of size n being disjoint is given by

$$P(n, m) = \binom{n-m}{m} \Big/ \binom{n}{m}$$

where the binomial coefficient $\binom{i}{j}$ gives the number of ways i items can be partitioned into sets of j elements. As an example, if the number of theoretically possible segments is taken as 1000 and the inventory size for two languages is 30, then the probability that these two inventories is disjoint—assuming they are random samples of segments—is P(1000, 30) = .396

$$P(1000, 30) = \binom{1000-30}{30} \Big/ \binom{1000}{30} = .396$$

or about 40%. We would, then, expect random selection to produce disjoint inventories about 40% of the time, but as noted earlier, there are no two disjoint inventories.

[4] Of the 198 languages in UPSID that contain /b/, 21 (10.6%) do not have /p/ or any other voiceless counterpart, though all these languages have a voicing distinction at some place. Having /b/ implies having /p/ for 177/198 = 89.4%.

[5] I count 87 languages in UPSID with an aspirated stop series. Of these only 19/87 = 21.8% re missing an aspirated stop at some place where a non-aspirated stop occurs.

[6] I am not sure larygealization per se is possible using an ingressive airstream, but some istinction can be made.

[7] No good way was found within SAS to clearly show the number of languages at each osition.

[8] I count a secondary articulation at each singly-articulated place as a separate place, since they eem to behave more like places than series. Thus 'palatalized bilabials' is a single place, rather an a place and a manner.

REFERENCES

ELL, ALAN. 1978. Language samples. In Universals of human language, Vol. 1, Method and theory, ed. by Joseph H. Greenberg et al., 123-156. Stanford, CA: Stanford University Press.

_____. 1984. Review of Maddieson 1984. Lg. 62.899-902.

HOMSKY, NOAM, & MORRIS HALLE. 1968. The sound pattern of English. New York: Harper & Row.

ROFT, WILLIAM. 1990. Typology and universals. Cambridge: Cambridge University Press.

UGUNDJI, JAMES. 1966. Topology. Boston, MA: Allyn and Bacon.

ILJENCRANTS, JOHAN, & BJÖRN LINDBLOM. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. Lg. 48:839-62.

INDBLAD, PER. 1980. Svenskans *sje-* och *tje-*ljud i ett allmanfonetisk perspektiv. (Travaux de l'Institut de Linguistique de Lund, 16.) Lund: Gleerup.

INDBLOM, BJÖRN. 1986. Phonetic universals in vowel systems. In Ohala & Jaeger (eds.), 13-44. Orlando, FL: Academic Press.

_____. 1987. A typological study of consonant systems: The role of inventory size. (Reports from Uppsala University, Department of Linguistics, 17.) Uppsala.

IADDIESON, IAN. 1984. Patterns of sounds. Cambridge: Cambridge University Press.

IILLER, GEORGE A., & PATRICIA E. NICELY. 1955. An analysis of perceptual confusion among some English consonants. Journal of the Acoustical Society of America 27:338-352.

EARY, TERRANCE M., & JOHN T. HOGAN. 1986. Phonological contrast in experimental phonetics: Relating distributions of production data to perceptual categorization curves. In Ohala & Jaeger (eds.), 141-61.

HALA, JOHN. 1980. Chairman's introduction to Symposium on phonetic universals in phonological systems and their explanation. Proceedings of the Ninth International Congress of Phonetic Sciences, 1979, Copenhagen: Institute of Phonetics, University of Copenhagen.

HALA, JOHN J., & JERI J. JAEGER, eds. 1986. Experimental phonology. Orlando: Academic Press.

NKERTON, SANDRA. 1986. Quichean (Mayan) glottalized and nonglottalized stops: A phonetic study with implications for phonological universals. In Ohala & Jaeger (eds.), 125-39.

ULLUM, GEOFFREY K., & WILLIAM A. LADUSAW. 1986. Phonetic symbol guide. Chicago: University of Chicago Press.

TEVENS, KENNETH N. 1989. On the quantal nature of speech. Journal of Phonetics 17.3-45.

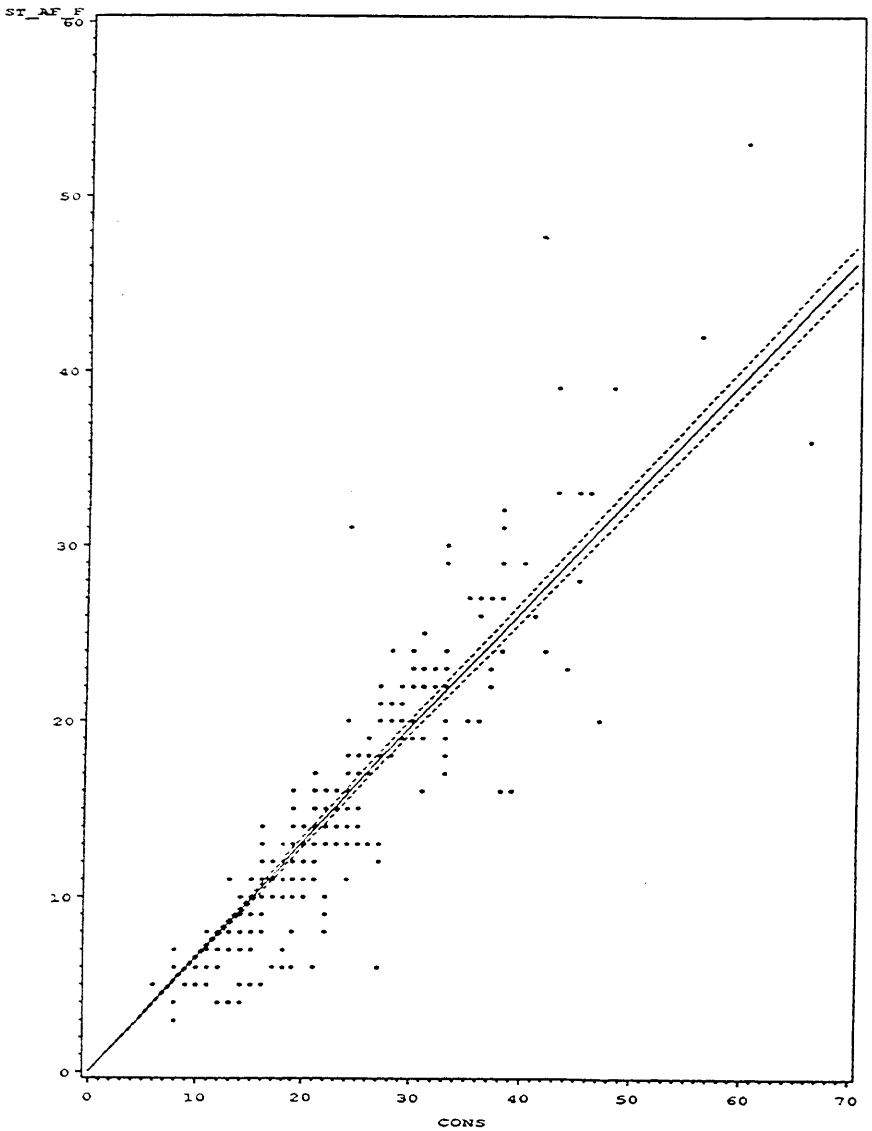
EN BOSCH, LOUIS F. M., & LOUIS C. W. POLS. 1989. On the necessity of quantal assumptions: Questions to the quantal theory. Journal of Phonetics 17.63-70.

RIGHT, JAMES T. 1986. The behavior of nasalized vowels in the perceptual vowel space. In Ohala & Jaeger (eds.), 45-67.

APPENDIX

Tables I–IX are the statistical plots used in the typological study and are discussed in §7. They were produced using the statistical package SAS at the Statistical Computing Center of the University of Colorado at Boulder and represent linear and quadratic least square fits of the data. Each point may represent more than one language.

https://scholar.colorado.edu/cril/vol12/iss1/1
DOI: https://doi.org/10.25810/8d78-9892



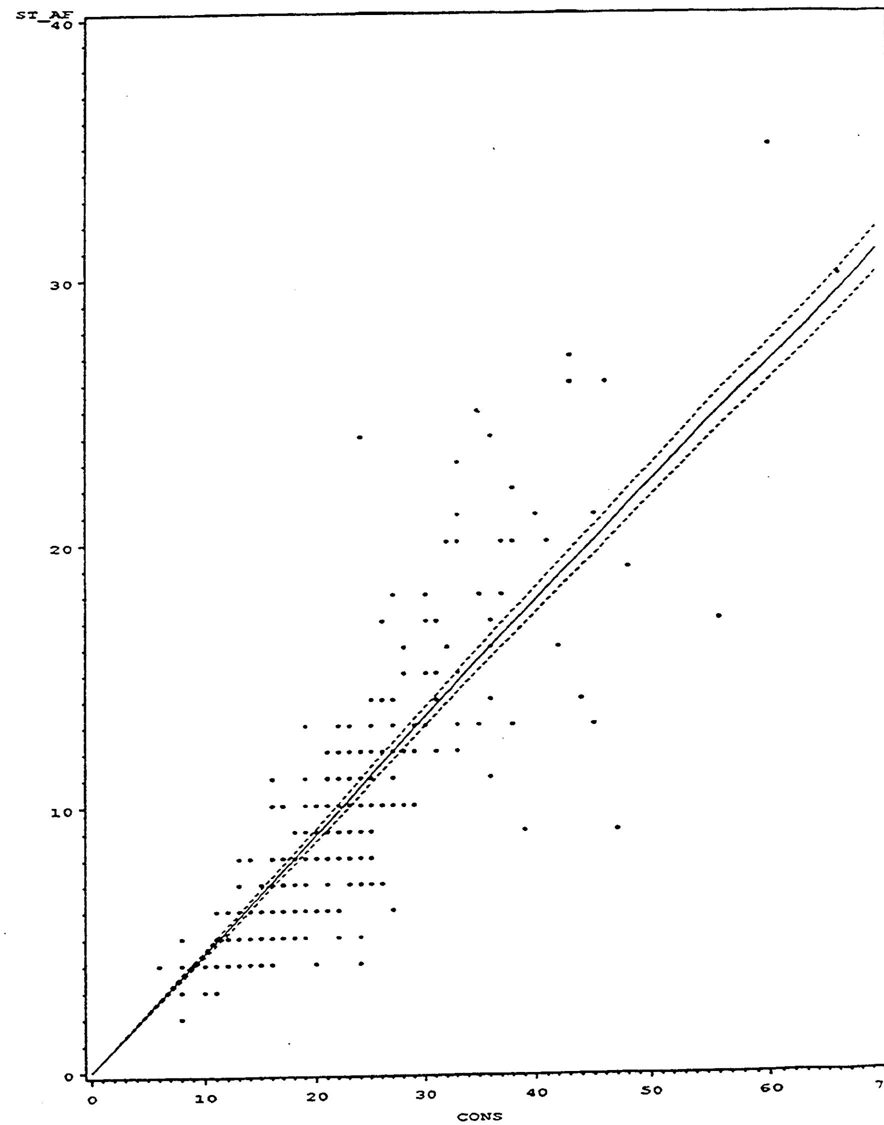
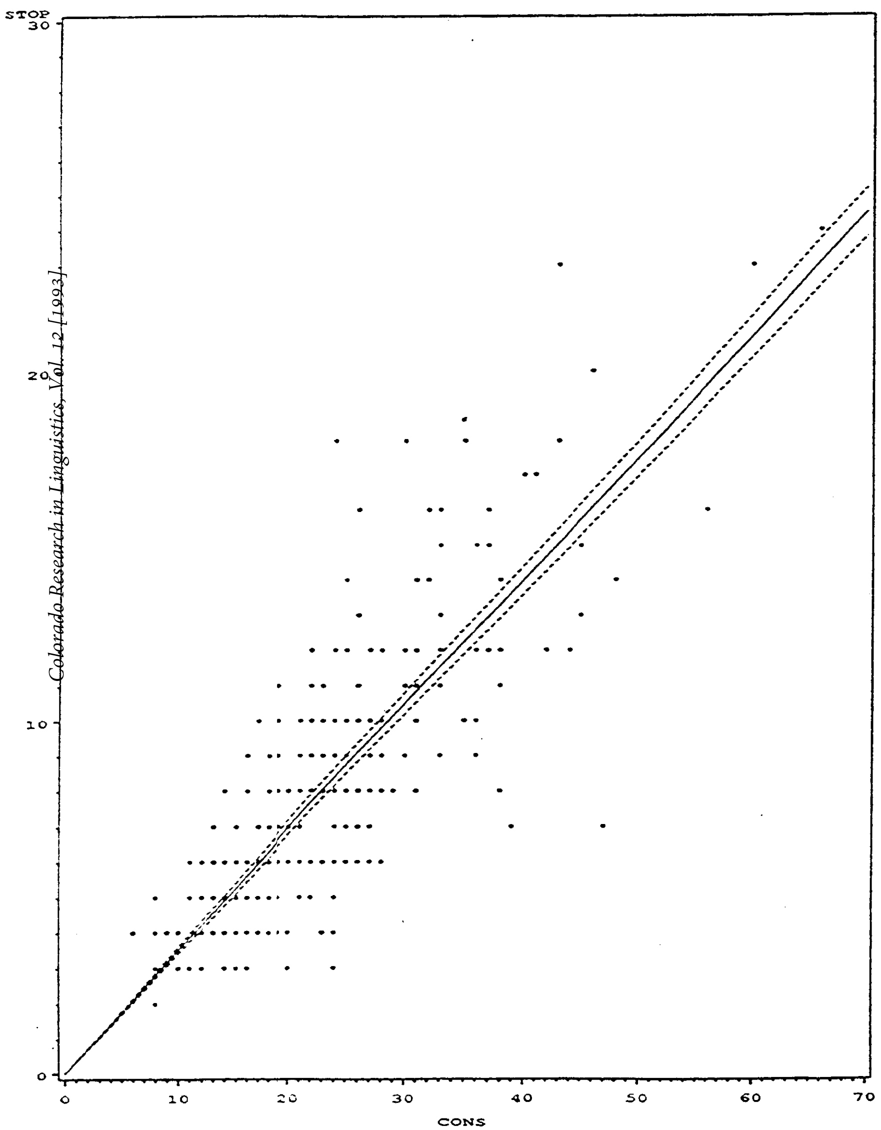


TABLE 2. Linear fit of stops and affricates to total consonants.



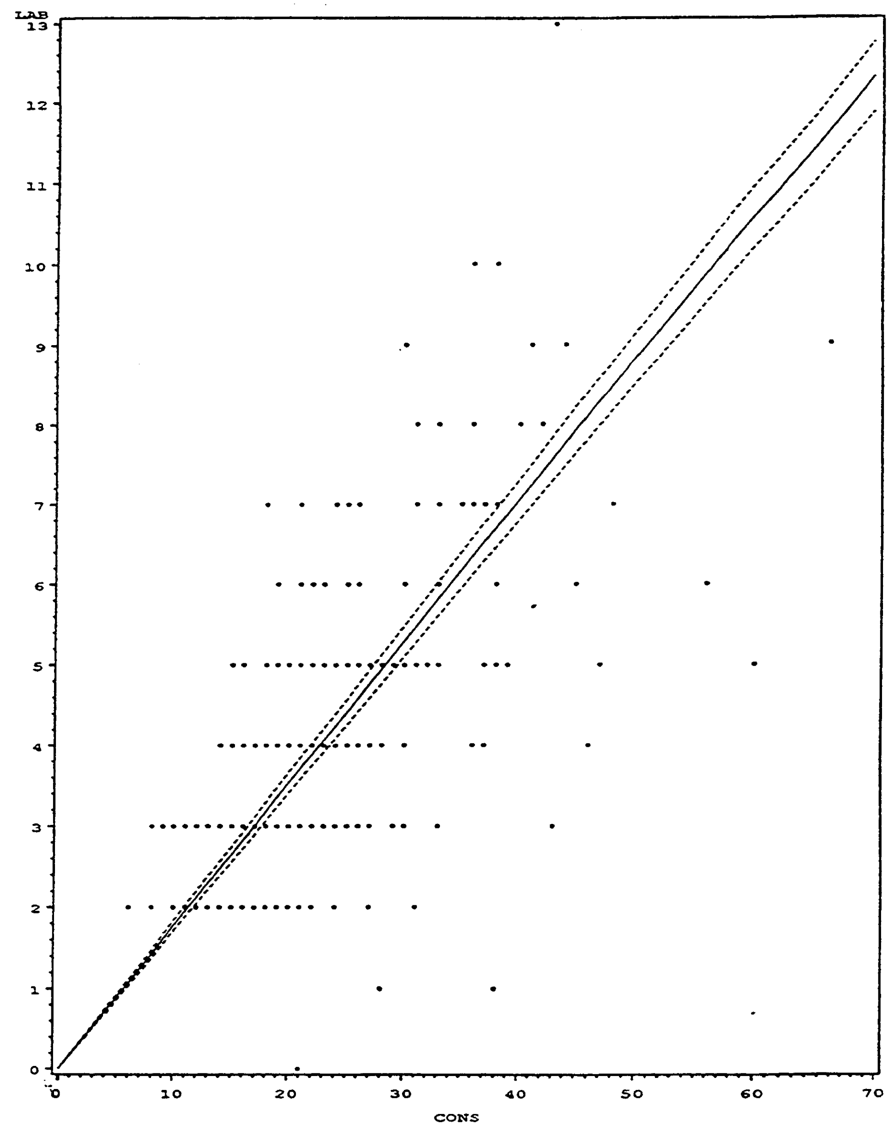


TABLE 4. Linear fit of labials to total consonants.

https://scholar.colorado.edu/cril/vol12/iss1/1
DOI: https://doi.org/10.25810/8d78-9s92

20

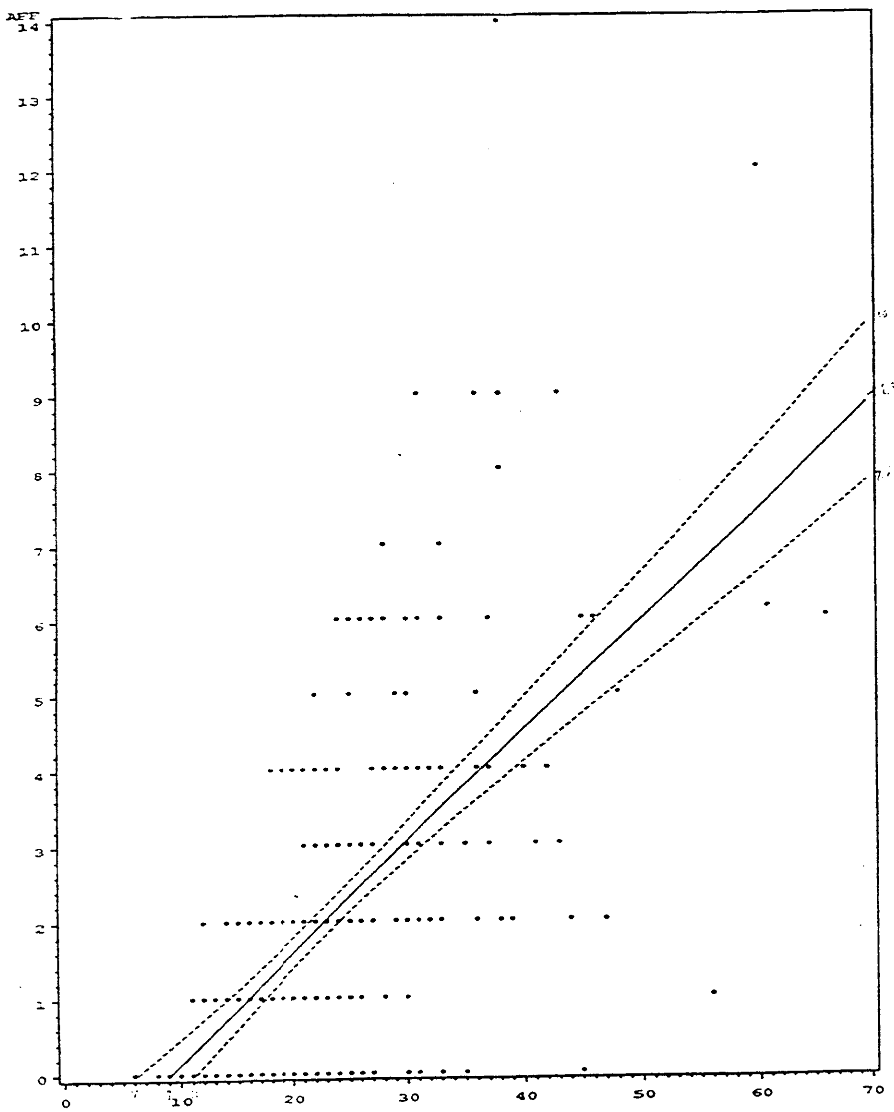


TABLE 5. Linear fit of affricates to total consonants.

21

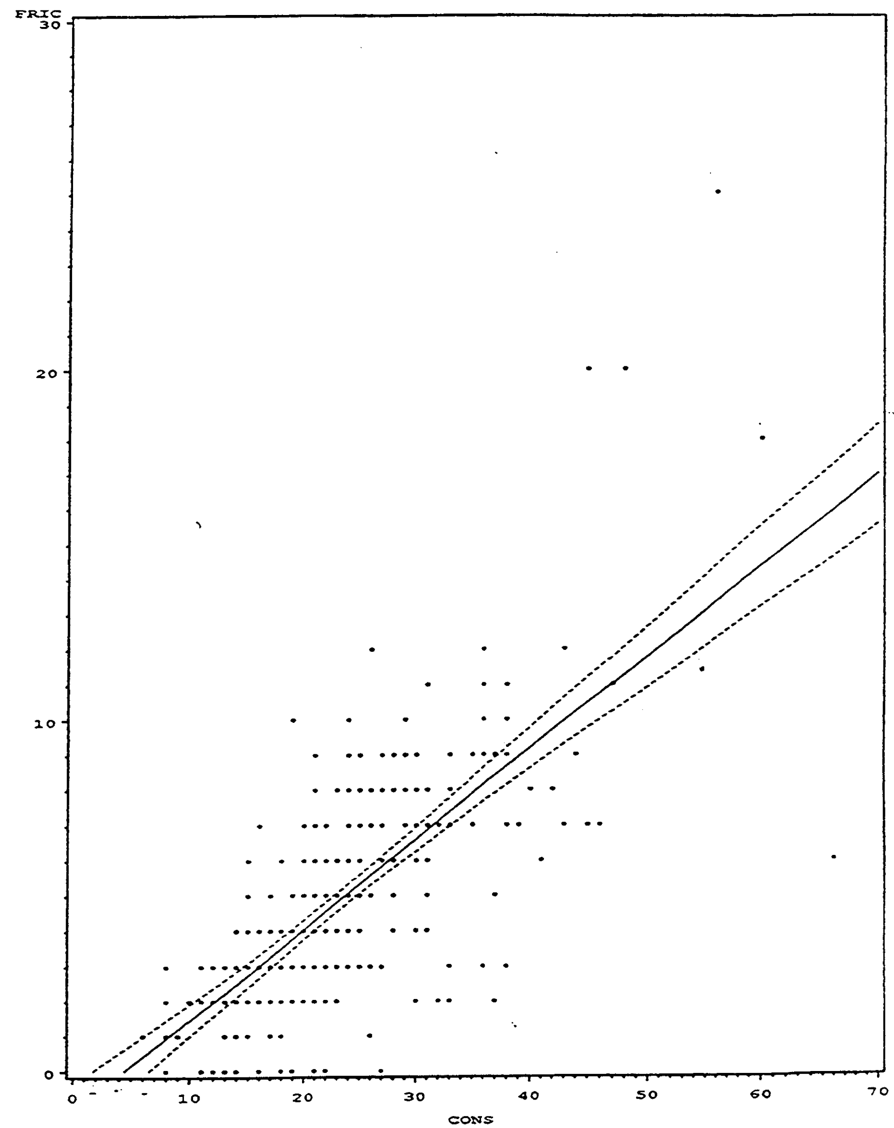


TABLE 6. Linear fit of fricatives to total consonants.

TABLE 7. Quadratic fit of stops to total consonants.

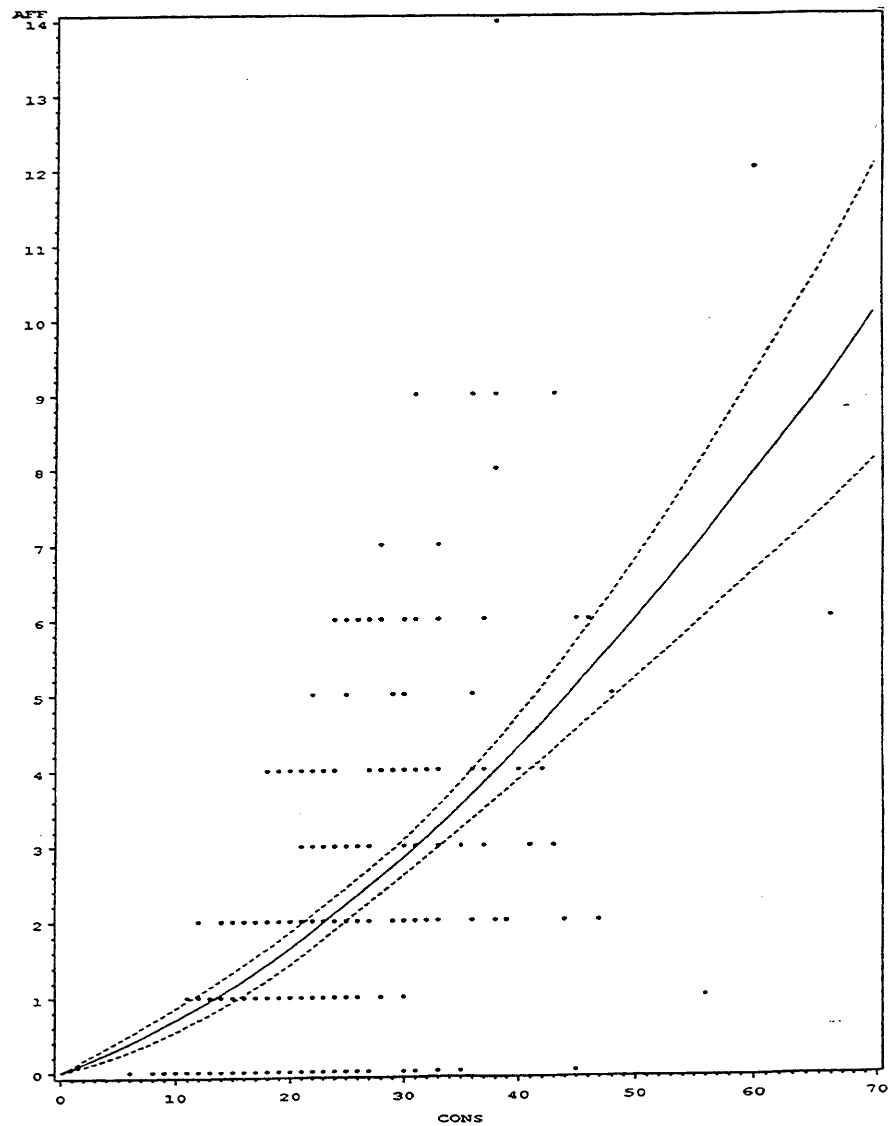


TABLE 8. Quadratic fit of affricates to total consonants.

https://scholar.colorado.edu/cril/vol12/iss1/1
DOI: https://doi.org/10.25810/8d78-9s92

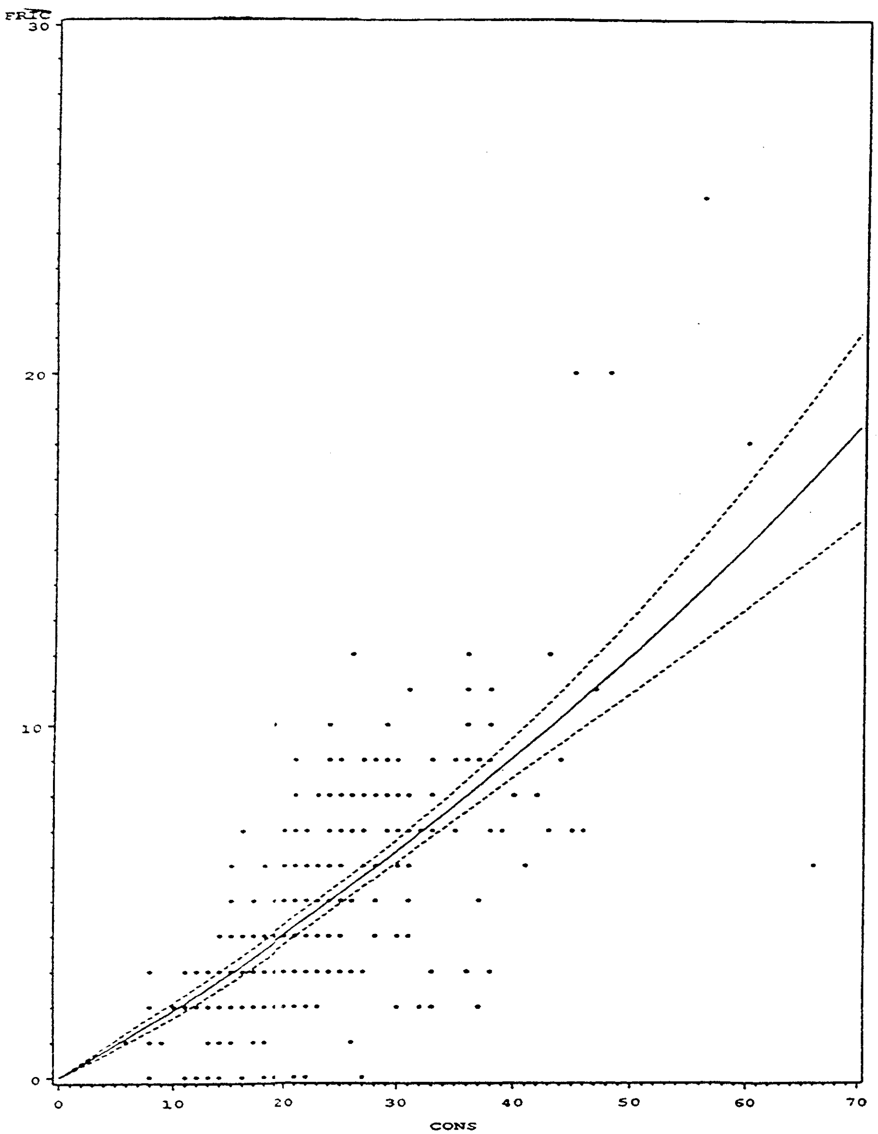


TABLE 9. Quadratic fit of fricatives to total consonants.

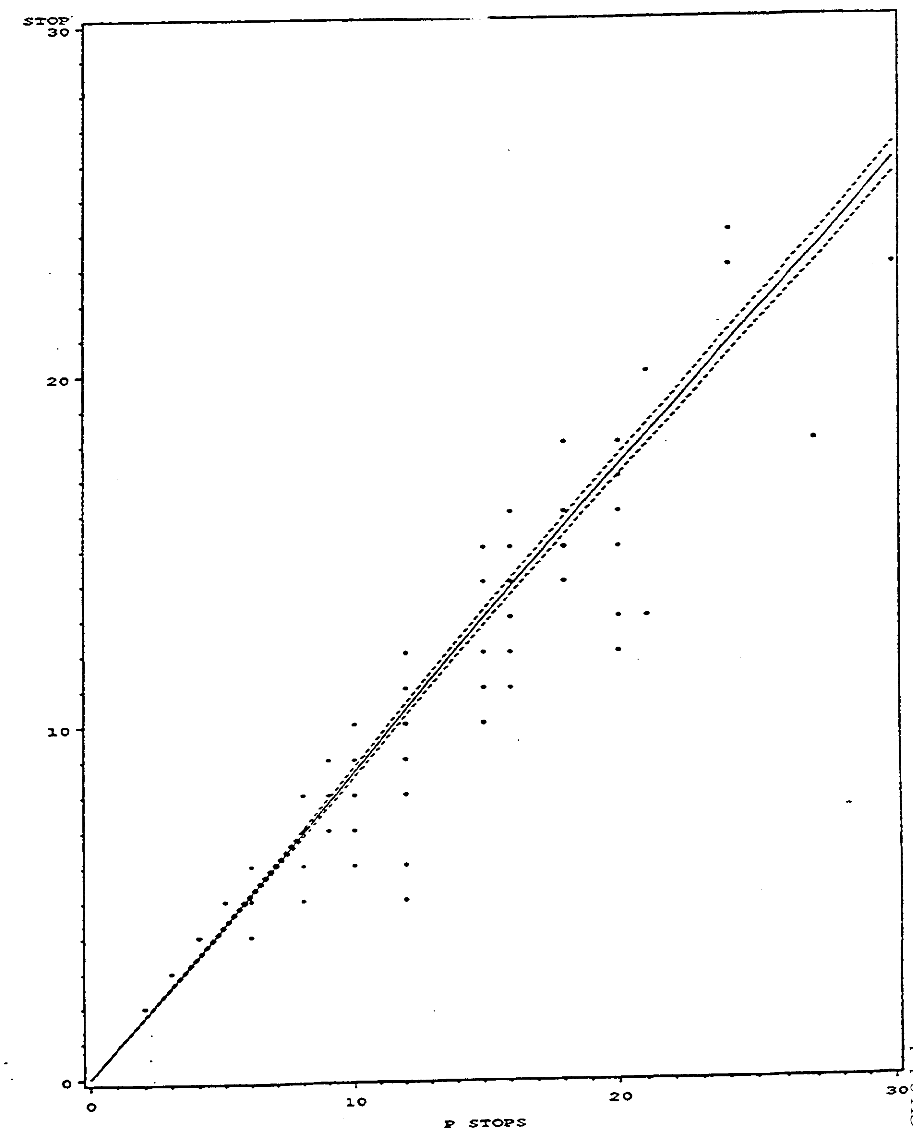


TABLE 10. Linear fit of stops to possible stops.