UPSID

U C L A PHONOLOGICAL SEGMENT INVENTORY DATABASE

WPP

U C L A WORKING PAPERS IN PHONETICS

FIFTY

AUGUST

1980

UCLA Working Papers in Phonetics - Volume 50

Contents

A U G U S T 1 9 8 0

The UCLA Phonetics Laboratory Group

Diane Ainsworth
Ron Anderson
Steven R. Anderson
Pat Coady
Sandy F. Disner
James F. Fordyce
Vicki Fromkin
Manuel Godinez
Steven Greenberg
Kathleen Houhilan
Richard Janda
Hector Javkin
Peter Ladefoged
Mona Lindau-Webb
Wendy Linker
Ian Maddieson
Priscilla McCoy
Jonas Nartey
George Papcun
Lloyd Rice
Diane Ridley
Mary Smith
Anne Wingate
Andreas Wittenstein

As on previous occasions, the material which is presented here is simply a record for our own use, a report as required by the funding agencies, and a preliminary account of work in progress.

Funds for the UCLA Phonetics Laboratory are provided through:

USPHS grant NS09780
NSF grant BNS78-07680
NSF grant BNS79-10051
and UCLA Department of Linguistics.

Correspondence concerning this series should be addressed to:

Phonetics Laboratory,
Department of Linguistics,
UCLA,
Los Angeles, CA 90024.

Prefatory Note

This volume of working papers reports on the work done at UCLA to create and exploit a computer-accessible database containing the phonological segment inventories of a sample of the world's languages. We have named this project UPSID - the UCLA Phonological Segment Inventory Database. It was conceived in the course of a class on the structure of segment inventories taught by Ian Maddieson, but expanded in magnitude far beyond the original plans. It has now become a significant resource for the Phonetics Laboratory's work on phonetic differences between languages as well as for research on phonological universals and related issues.

The principal work of establishing the computer file was done by Sandra F. Disner, Vivian Flores, James F. Fordyce, Ian Maddieson, Jonas N.A. Nartey, Diane G. Ridley and Vincent van Heuven. Help in collecting data was also provided by Steve Franks, Bonnie Glover, Peter Ladefoged, Mona Lindau-Webb, Robert Thurman, Anne Wingate, Andreas Wittenstein and Eric Zee. Additional assistance has come from other linguists at UCLA and elsewhere. Mel Widawsky of UCLA's Academic Computing Services provided valuable help in persuading the computer to accept the indigestible bulk of our input. A library of the sources from which data was drawn has been compiled thanks mainly to Hector Javkin and Diane G. Ridley. John Crothers supplied us with an early copy of the final report of the Stanford Phonology Archive report, enabling us to benefit from the experience accrued at Stanford. Obviously the Language Universals Project at Stanford served as an inspiration for our work and we are grateful for their example.

In this volume the first paper presents a relatively full account of the UPSID project and the form of the resulting database. The remaining papers report on some of the analysis of the data which has been completed to date. Besides these papers, an M.A. thesis by Jonas N.A. Nartey dealing with stops, nasals and fricatives has been completed and has already appeared as <u>UCLA Working Papers in Phonetics</u> 46 (1979).

The work represented here has been funded in part by the National Science Foundation through grant no. BNS78-07680.

Ian Maddieson.

UPSID : The UCLA Phonological Segment Inventory Database

Ian Maddieson

1. Introduction

The discovery of generalizations concerning the structure of phonological inventories has been a significant objective of recent work in linguistics. Such generalizations have been taken into account, explicitly or implicitly, in the formulation of phonological theories, in evaluating competing historical reconstructions, in constructing models of language change and language acquisition, and they have stimulated important linguistically-oriented phonetic research. Their value is evidently considerable, but what is their origin?

There seem to be three types of sources for these generalizations. The type with the longest tradition is an essentially impressionistic account based on a linguist's experience of a number of languages. Statements by Trubetskoy (1939), Jakobson and Halle (1956), and Ladefoged (1971) as well as incidental remarks in the papers of numerous authors are examples of this category. Although they may be based on familiarity with a very large number of languages, because the list of languages represented in this experience is not given and there is no quantification attached to the statements made, there is some doubt about the scope of the conclusions reached. The second type consists of explicit samples of languages compiled for the purpose of a single study, such as Ferguson (1963), Greenberg (1970) and Hyman (1977) on nasals, glottalic consonants and stress respectively. In these cases the quality of the sample (cf. Bell, 1978) and the significance of the conclusions reached can be independently assessed by the reader (cf. Hurford, 1977). The third kind of data source is a standardized multi-purpose survey, epitomised by the Stanford Phonology Archive (SPA), compiled at Stanford University as part of the broad Language Universals Project under the direction of J.H.Greenberg and C.A.Ferguson. A large proportion of recent work on phonological universals is either directly based on the SPA or owes an indirect debt to it. The UCLA Phonological Segment Inventory Database (UPSID) is a source of this third kind.

2. Why a survey?

There are several reasons for the superiority of this third kind of data source and it may be useful to rehearse them here. They arise from the nature of the purposes for which the survey was designed. These purposes are, first, to generate observations - e.g observations concerning the frequency of segments of different types and of the phonetic attributes of segments, as well as their co-occurrence in phonological inventories. Secondly, to subject hypotheses concerning such matters as segment frequency to the test of comparison with empirical observations. These may range from simple hypotheses of the form that there are significant differences in the frequency of segments of different types to more elaborate ones positing contingent relationships between the occurrence of (sets of) different segments, or limitations on the distribution of phonetic attributes within inventories. The third, and perhaps most significant, purpose behind compilation of such data sources is as a stimulus to the generation of hypotheses which relate to other fields of the study of language but for which such matters as segment frequencies, inventory size, and so on, may be the point of departure. Such hypotheses may relate to matters of production, perception, or acquisition, linguistic change or languages in contact, but extend to data other than observations concerning segments and inventories.

Now, most observations and hypotheses alluded to above necessarily relate to relative frequencies, relatively restricted distribution, etc. Experience has shown that few interesting things are to be said about phonological inventories that are truly universal; i.e. exceptionless. Apart from observations such as "all languages have a contrast between consonants and vowels" most of the substantive generalizations concerning segments and inventories are or can be expected to be of the form "a situation X occurs more (or less) frequently than chance leads us to predict." That is, in layman's terms, they are statistical observations. They can therefore only be meaningful if they are drawn from, or tested with respect to a body of data appropriately designed for statistical analysis. In other words, one which is representative, extensive and uniform in analysis as far as possible. This requires establishment of a large and appropriately selected sample of languages and a standardized procedure for interpreting their phonologies. Once established, numerous commensurate studies on the same data can be made.

3. Why UPSID?

Given the establishment of the Stanford Phonology Archive (also known as CALM--Computer Archive of Language Materials), which set out to satisfy the criteria outlined above, is there a need for a new archive? The Stanford Universals Project has terminated and hence SPA is no longer readily computer-accessible to outsiders. The final report prepared lists for each language the segment inventory and phonotactic constraints, etc., and supplies extensive comments citing details and obscurities in the original sources. There are also indexes referring to languages having given segments, features and allophonic rules. There is a great deal of information in this report, but the unaided user is taking on a commitment to a great deal of manual labor if he/she aims to use it. Besides this consideration of convenience, there are a number of other drawbacks with SPA. In aiming to construct a properly structured sample and to encode complete phonologies, SPA may have been taking on more than it was possible to achieve. The progressive adoption of smaller target numbers of language perhaps indicates a growing awareness at Stanford of the ambitiousness of the original plan. Although at one time aiming to include as many as 277 languages, subsequent reductions limited the number first to 225, then 209, with the final size of the inventory being 196. A principal reason for exclusion was the scarcity of adequately documented phonologies. The variability in detail of the sources which were used also necessarily produced entries which vary in their completeness--from those which cover little more than a list of phonemes to those which are able to include a lot of allophonic details and information on phonological alternations. Thus, for retrieval of certain information, the true sample size is smaller than 196 languages. With each reduction the liklihood that the sample is no longer representative and properly balanced increases. With Stanford's experience in mind, UPSID was designed to be less ambitious in the scope of information about each language filed, but to be more comprehensive in the number of languages covered. Users of SPA have also commented that there is a certain inflexibility inherent in the format chosen for data entry. This is basically a text-oriented system (for some description, see Vihman 1974). For example, each segment is entered as an alphabetic character (p,b,m, etc.) or string (|-retroflex, epsilon, o-open-long-nasalized, etc.), followed by a rather open-ended set of features attributed to the segment (e.g. obstruent, bilabial, voiceless, etc.), comments from the source and so on. With this format, retrieval of essentially numerical data is not the most readily achieved operation and manipulation of the data can be complex and cumbersome. Stanford's aim was to maximize "accountability" i.e. limiting their information to that available in the source grammar. In designing UPSID, we have aimed to maximize the ease and flexibility with which numerical data can be manipulated, and have decided that the compilers may on occasion need to adopt a more active role in interpreting the source to guarantee a consistent treatment of similar sounds. We also differ somewhat from SPA on the principles that should govern selection of languages, e.g. in rejecting the number of speakers as a criterion for inclusion (or exclusion) of a language. We have however profitted greatly from the example of SPA, and from the excellent hard work that went into its compilation by using the final report as a secondary source of data. The following sections describe the plan of UPSID.

4. Selection of languages for UPSID

The ideal sample for purposes of statistical evaluation is a random sample, drawn from the total population under study. In the case of language data, the "population" is all the world's extant languages. It is impossible to draw a random sample from this population for two reasons. First, there are areas of the world about whose languages we have no data or wholly inadequate data. Second, a 'language' is not a clearly demarcated object. The common criterion of "mutual comprehensibility" used to define linguistic similarity yields a gradient, and besides, is often not reflexive in its characteristics (cf. Ladefoged et al. 1972). Thus, unlike a population of, say, registered voters, it is not possible to enumerate and individualize the members of the population. Hence, no basis exists for drawing a random sample.

A proper sample of languages must therefore be constructed by some other sampling procedure. The chosen one is a variety of quota sample (in fact, the usual linguistic sampling procedure is to draw a quota sample based on typological, genetic or areal groupings). The principle on which UPSID is based is to select one and only one language from each moderately distant genetic grouping, so that the selected languages represent in proper proportion the internal genetic diversity of various groupings. The obvious difficulties in the way of such a scheme are the lack of sound genetic classifications in certain areas (e.g. South America, New Guinea), the difficulty of comparing genetic distances in different language families, and lack of requisite data from some known groups. The advantages of this procedure are that it precludes, in principle, selection of data which represents arguably the same language in several varieties (unlike SPA which includes for example both Moroccan and Egyptian dialects of Arabic, and Maltese); also it directs a principled search for the data to fulfill the quota design and avoids undue reliance on descriptions that happen to be at hand (the 'bibliographic convenience' factor mentioned by Bell 1978). A genetic basis for the sample is selected in preference to any other since it is the only classification which is, in principle, not arbitrarily determined by the criteria chosen for the classification, but instead aims to represent real historical relationships. In addition, it is appropriately independent of the phonological characteristics of which it is desired to find the frequencies. (Note that phonological diversity per se is not a basis for inclusion.) Because each language included is relatively distinct genetically from all others in the sample, each represents the outcome of the opportunity for independent operation of historical processes. Similarities between languages in the sample are therefore not due solely to the effect of shared historical origin. As noted before, the number of speakers is considered a quite inappropriate basis for including (or excluding) a language. The size of extant populations of speakers of languages is an accident of political and social history that is quite irrelevant to questions relating to the structure of human languages.

No thorough-going attempt to determine a single criterion for the degree of genetic separation required for inclusion of a language has been made. The procedure has been to assemble the most comprehensive genetic classifications available based on the best available data and to produce, by synthesis of several classifications where necessary, an overall classification for each of eleven major groupings of languages, plus several smaller groups. These groups are shown in Table 1 together with the number of languages included in UPSID from each at the time of writing.

Indo-European	21
Ural-Altaic	22
Niger-Kordofanian	31
Nilo-Saharan	21
Afro-Asiatic	21
Austro-Asiatic	6
Australian	19
Austro-Tai	25
Sino-Tibetan	18
Indo-Pacific	27
Amerindian	89
Others (Dravidian, Caucasian, Khoisan, Eskimo-Aleut, etc.)	18

Table 1

Intermediate levels of classification were then sampled to select the languages for inclusion in UPSID. The density of this sampling might be thought of as representing an intention to include no pair of languages which had not developed within their own independent speech communities for at least some 1000-1500 years, but to include one language from each group of languages which shared a closer history than that. An example might help to clarify the process. The existence of a Nilo-Saharan language family was proposed by Greenberg (1966). A composite classification was drawn up using Greenberg (1966,1971), Tucker and Bryan (1956, 1966), Bender (1976) and other minor sources. This included ten major groupings of languages whose relationship of each other is clearly relatively remote. Seven of these (Songhai, Saharan, Maban, Fur, Berta, Kunama and Koman) were judged to have relatively little internal diversity* and hence were represented by one language each. The tenth group, Gumuz, is insufficiently known and no phonological data is available. This leaves two groupings with considerable internal diversity, Eastern Sudanic and Central Sudanic, consisting of eleven and seven sub-groupings respectively. In one case, namely the Nilotic subgroup of Eastern Sudanic, the degree of internal genetic diversity justified inclusion of a language of each of the Western, Eastern and Southern subgroups. The target was therefore thirteen Eastern Sudanic languages. Eleven were actually obtained as two groupings lacking adequate data. Only three of the targeted seven Central Sudanic languages were included; in this case partly because sources known to exist were unavailable to us. Thus an ideal figure of twenty-eight languages to represent all of Nilo-Saharan is reflected with some imperfection in a total of twenty-one actually sampled, a seventy-five percent 'response rate.' In certain instances, notably the Southern Amerindian groups of languages and the Indo-Pacific family, it proved much more difficult to obtain complete classifications and to resolve conflicting groupings. In these instances it is not really possible to assess how adequately the whole family has been sampled. However, an educated guess is that overall the present sample contains between seventy and eighty percent of the languages that it should include in order to completely fulfill its design specifications. A complete list of the languages currently included together with a skeleton classification is given as Appendix A.** Languages included in the SPA are indicated by (S). The sources consulted, in some cases indirectly via Stanford's report, are listed in Appendix B.

5. Determining the inventories

For each language a list of phonologically contrastive segments was drawn up. This procedure of course presupposes that

* This conclusion may be unjustified in the case of Saharan, which consists of four languages which <u>may</u> be rather diverse.

** Data on salient omissions is welcome from readers. Please write to Ian Maddieson, Dept. of Linguistics, UCLA, Los Angeles, CA 90024.

such an analysis represents significant and interesting facts about a language. We do not propose to take space to argue for this proposition here--merely to point out that in practice nearly all linguists use such an analysis. Our own view goes further than some in asserting that the phonological segments can (and should) be characterized by certain phonetic attributes. Linguists who believe that phonology is necessarily involved with purely abstract constructs will perforce part company with us at this point. (However, since such abstractions cannot be compared, they presumably are not interested in the kind of language universal properties this archive is designed to investigate.)

Determining the phonological inventory for each language involves two principal aspects--determining how many contrastive units there are and determining what phonetic characteristics should be attributed to each one. The first aspect concerns defining what is contrast and resolving questions about the unity or otherwise of 'suspect' complex phonetic events, such as affricates, geminated consonants, diphthongs, prenasalized stops, etc., which may be open to interpretation as unitary segments or as sequences of some simpler segments in the language.

'Contrasts' are sound differences capable of distinguishing lexemes or morphemes in the language involved, given that data on relevant factors such as boundaries (word, morpheme, etc.), stress placement, syllabification, etc., can be used to predict variants but that diacritical features, arbitrary rule types, morpheme classes, etc., cannot be used. These principles are applied to evaluate critically the information in the source and the resulting analysis may differ from the phonological inventory assumed in the source. The suspect units/clusters have been examined as carefully as available information permits to determine their status as units or sequences. If, for example, they can be split by a morpheme boundary or form a part of a more general set of permitted clusters, they are treated as sequences. If, on the other hand, there are no similar clusters to those that would be created by a sequence interpretation, this is taken as favoring a unit interpretation.

In addition to the above considerations, certain types of contrasts have been interpreted as suprasegmental, i.e. as not relevant to setting up an inventory of phonological segments. These have included, by definition, tone and stress phenomena. Nasalization and properties involved in vowel harmony systems have always been treated as segmental. Length, however, has been treated as suprasegmental if it applies to a whole class of segments, such as all vowels. Otherwise it is treated as a segmental property: for example, Yay (a Northern Tai language) has long and short low vowels /a, a:/ but the other vowels do not contrast in length. So there is considered to be no basis in this language for regarding length as suprasegmental, it is an attribute of a particular segment only.

The remaining issue concerning the size of the inventory of segments has to do with inclusion or exclusion of segments with more or less marginal status. Certain segments which occur only, say, in interjections or in foreign words that are clearly not established as loans have been excluded from the inventory altogether. In certain other cases, segments are included without distinction, for example, if they are restricted to loanwords but these loans appear to be fully assimilated. A third option is to include the segments but to indicate the nature of their marginal status. How this is done will be discussed below in the section on variables in the data base (see 6 (j)).

In determining the phonetic properties to be attributed to each segment, the principle was to select the major allophone which was determined to be the most representative. Where information was available, this selection was based on three principal considerations: 1) What allophone has the widest distribution (i.e. appears in the widest range of and/or most frequently occurring environments); 2) What allophone most fully represents the phonetic range of variation of all allophones; 3) What allophone is the one from which other allophone can be most simply and naturally derived. There are cases in which answers to these questions produce conflicting results; in these cases an attempt was made to resolve the conflict by considering how badly overall the conditions would be violated if the answer given by, say, consideration (2) was preferred over that given by (1), and to select the answer which did least violence to all three considerations taken together. The level of detail aimed at is one that is approximately equal to that attained by the traditional 3-item label of phonetics, specifying voicing, place and manner for consonants, height, frontness and lip-position for vowels, plus additional labels required for features such as secondary articulations, etc. A set of coding variables to represent the required phonetic attributes with a minimum of appeal to redundancy to interpret their meaning was designed and is discussed in full below. This set was also designed to accommodate some of the major indeterminancies found in the phonological sources consulted. Overall, the effort was to design a set of variables which was as little prone to errors of interpretation and as little likely to generate impressions of explicitness when unjustified as could reasonably be accommodated within practical limits of convenience and economy. However, the data base cannot be used without danger of misinterpretation unless the system of coding variables is fully understood. For this reason, the following section should be read carefully by any user and, indeed, would advisedly be read by those who are interested in interpreting the results of UPSID represented in the following articles in this volume.

6. Indices and Variables

UPSID represents each segment as a separate record in the form of a (notional) 80-column card image. The first seven columns are devoted to identifying indices, 10-70 contain variables for phonetic attributes, 80 contains the ANOMALY feature. Each of the variables referring to phonetic attributes takes a value of 1 or 0 with 1 indicating that the allophone chosen to represent the segment possesses the attribute named and 0 indicating that it lacks the attribute. (There are a few minor exceptions to this rule which are explained below.) In the exposition which follows, each variable is referenced by its (notional) column location, given a definition and followed by the form of the variable name in the actual computer data base (parenthesized and in capitals). For convenience, the variables are grouped into classes. Notes and comments on the use of the variables are interspersed where appropriate.

a) Indices.

1-4. Language identification number (LANGNO). This number serves to identify the language to which a segment belongs. It currently consists of three digits, the first of which indicates affiliation to one of the major groupings used as the genetic basis of the sampling.

5-7. Segment identification number (SEGNO). Each segment within a language is numbered sequentially. The combination of LANGO and SEGNO thus identifies one and only one record in the data base.

b) Alphanumeric code.

8-9. Segment code (SEGCODE). For convenience in assembling the data base, a mainly alphabetic code was devised to represent commonly occurring segments. Variable values were then generated from this code. As these symbols have a useful mnemonic function, they have been left in.

c) Variables indicating place of articulation.

All consonants except /h/ and /ɦ/ are specified for at least one place of articulation. These are listed in the conventional front to back sequence. Double articulations are indicated by specifying two places of articulations, but secondary articulations are indicated by use of a separate set of variables. /h/ and /ɦ/, having place determined by environment, do not receive any place specification (see further below under fricative).

10. Bilabial (BILABIAL).

11. Labiodental (LABODENT).

12. Dental (DENTAL). This variable has the value 1 if it can be ascertained that a true dental is intended, e.g. if the source says "the tongue touches the teeth" or "/t/ is like French t̪, not English t̪." The description 'dental' is often applied to a segment which is more accurately described as alveolar, hence, if no added evidence is available, segments are not assumed to be true dentals.

13. Unspecified dental or alveolar (UNSPDENT). Segments simply indicated by transcriptions using /t,d,s/ etc. may be either dental or alveolar. In order not to falsify the data this variable is used to indicate such segments with an incompletely specified place of articulation. Segments described simply as 'dentals' are also included here.

14. Alveolar (ALVEOLAR).

15. Palato-alveolar (PALATALV). Palato-alveolar and alveo-palatal are not distinguished.

16. Retroflex (RETROFLX).

17. Palatal (PALATAL).

18. Velar (VELAR).

19. Uvular (UVULAR).

20. Pharyngeal (PHARYNGL).

21. Glottal (GLOTTAL). This variable is only used for glottal stops; the characterization of /h/ as a "glottal fricative" is rejected.

d) Variables indicating manner of articulation.

Conventional phonetic labels, such as 'plosive,' 'click,' 'vowel' generally combine information on aperture and airstream. This set of variables does likewise. All segments must be specified for at least one of these, but rarely a segment may be specified for more than one, e.g. a fricative trill.

22. Plosive (PLOSIVE). For pulmonic egressive stops including glottal stops. Does not include nasals.

23. Implosive (IMPLOSIVE). For glottalic ingressive stops whether voiced or voiceless.

24. Ejective stop (EJECTSTP).

25. Click (CLICK). For non-affricated clicks.

26. Fricative (FRICATIV). For pulmonic eggressive fricatives, /h/ and /ɦ/ are not considered fricatives.

27. Ejective fricative (EJECTFRC).

28. Affricate (AFFRICAT). For pulmonic eggressive affricates.

29. Ejective affricate (EJECTAFF).

30. Affricated click (AFFCLICK).

31. Unspecified 'r-sound' (UNSPECR). This variable takes the value 1 for segments which are simply identified as some kind of 'r-sound' (e.g. by being transcribed /r/ or called a 'vibrant') but which cannot be further classified as a trill, tap, flap, approximant, etc.

32. Tap (TAP).

33. Flap (FLAP).

34. Trill (TRILL).

35. Approximant (APPROXMT). This variable identifies 'semi-vowels,' r-glides, nonfricative laterals, etc.

36. Nasal (NASAL). For nasal consonants (with complete oral closure).

37. Simple vowel (SIMPVOWL). For monophthongs. A segment which is not identified as either a simple vowel or a diphthong below is thereby classified as a consonant.

38. Dipthong (DIFTHONG). This variable takes the value for unit dipthongs. Phonetic diphthongs which are phonologically analyzed as the result of a juxtaposition of simple vowels or a vowel and an approximant are, of course, not units but sequences.

e) Other consonant features.

39. Lateral (LATERAL). Takes the value 1 for all lateral segments, e.g. it indicates lateral release in lateral affricates.

40. Sibilant (SIBILANT). This variable serves to identify the class of sibilants within the fricative/affricate group. In many languages this is functionally redundant as place of articulation distinguishes the same class. However, it has occasional distinctive function in distinguishing between fricatives and/or affricates with similar place of articulation but distinct acoustic characteristics (e.g. /s/ vs. /θ/).

f) Secondary articulations.

41. Labialized (LABLZED).

42. Palatalized (PALTLZED). This variable takes the value 1 only for true palatalized consonants, i.e. those with a secondary palatal articulation. Thus a segment /c/ which occurs in a language as part of a "palatalized" series of stops / p^j, t^j,/ etc., will not be coded with this feature. Instead it will be reported as a palatal stop.

43. Velarized (VELRZED). Similar comments apply as to the above variable.

44. Pharyngealized (PHARGZED). May take the value 1 for either consonants and vowels.

45. Nasalized (NASLZED). This variable takes the value 1 for nasalized consonants and vowels, i.e. those with simultaneous nasal and oral escape. It is also used to characterize prenasalized stops (when these are clearly units). Thus the value 1 for this variable in combination with the value 1 for any stop variable (except click or affricated click) indicates a nasal onset to the stop.

46. Nasal release (NASRELSE). Takes the value 1 for post-nasalized segments only.

g) Vowel features.

All simple vowels are specified by a value of 1 on one vowel height, one vowel backness and one lip position variable. (They may also have the value 1 for other variables to indicate other distinctions.) Dipthongs segments are specified by assigning the value 1 to all the vowel quality variables needed to describe both their beginning and end points. A set of dipthong variables, discussed below, indicates the order of conflicting specifications.

47. High (HIGH).

48. Higher mid (HIGHMID).

49. Mid (MID). Used with systematic ambiguity for vowels which are indicated as 'mid' without further particularization and those which are true mid vowels (i.e. lie between higher mid and lower mid on a height scale).

50. Lower mid (LOWMID). Note that /ɛ/ and /ɔ/ are considered lower mid vowels not low vowels.

51. Low (LOW).

52. Front (FRONT).

53. Central (CENTRAL). /a/ in most languages is considered a central vowel.

54. Back (BACK).

55. Nonperipheral (NONPERIF). This variable takes the value 1 for 'laxed' noncentral vowels which are produced away from the periphery of the vowel space, for example /ɩ/ and /ɷ/. It may on occasion serve a mainly diacritical function where other features fail to distinguish vowels.

56. Rounded (ROUNDED).

57. Unrounded (UNROUNDD).

58. Lip-compressed (LIPCOMP). Takes the value 1 for 'labial' vowels that are produced with vertical compression of the lips but no drawing in and forward of the corners of the mouth ('rounding').

59. R-colored (RCOLORED). Used for retroflexed or r-colored vowels.

h) Diphthong variables.

The three variables for characterizing movement in diphthongs differ from most other variables in the inventory, which indicate only presence or absence of the attribute named by the variable. Here, a zero specification may indicate a specific property of a diphthong, namely movement opposite to that indicated by the variable name. Also, unlike most of the variables, they require reference to the values of other variables for their interpretation. Their function is to indicate which value precedes when conflicting specifications of vowel height, backness or lip position are given to a single segment. This method of coding diphthongs was adopted in order to avoid a very large number of variables.

60. Backing (BACKING). This variable takes the value 1 when the end point of a diphthong is more back than the beginning, as in /iə/, /eu/, /əu/, etc. It takes the value 0 when the end point is either more front than or has the same degree of backness as the beginning, as in /oi/, /ae/, /ou/, etc. Note that only three degrees of backness are considered (front, central, back).

61. Lowering (LOWERING). This variable takes the value 1 for diphthongs that have an end point lower than their beginning, such as /iə/, /ea/, etc. It takes the value 0 when the endpoint is higher than or equal to the beginning on the five-point vowel height scale used in UPSID, for example /oi/, /ou/, /ai/ etc.

62. Rounding (ROUNDING). This variable takes the value 1 when the endpoint of a diphthong is rounded but the beginning is unrounded, as in /eu/, /ao/, etc. It takes the value 0 when the endpoint is unrounded or both the beginning and endpoints are rounded, as in /oi/, /ai/ or /ou/. Note that a pair of diphthongs such as /oi/ and /io/ receive the same values on the vowel variables. This will involve having conflicting values for height, backness and rounded; they will be specified as being high <u>and</u> mid, front <u>and</u> back, rounded <u>and</u> unrounded. The diphthong features interpret these conflicts. /oi/ will have the value 0 for all three diphthong variables since there is no backing, lowering or rounding movement in this diphthong. Yet because the segment has conflicting values, there must be movement; it has to be in the opposite direction from the variable names, i.e. fronting raising and unrounding. On the other hand, /io/ receives the value 1 for all diphthong features and this indicates that in this case the mid portion of the diphthong follows the high portion, the back portion follows the front portion and the rounded portion follows the unrounded portion.

i) Variables for phonation types, etc.

63. Voiceless (VOICELES). Takes the value **1** for all segments in which the vocal cords do not vibrate. /h/ has the value 1 for this variable but for no others.

64. Voiced (VOICED). Takes the value 1 for all segments in which the vocal cords vibrate, whether as 'regular' voicing or in some other mode (creaky, breathy etc.).

65. Aspirated (ASPIRATD). Takes the value 1 for all voiceless aspirated segments.

66. Laryngealized, creaky, glottalized (LARYNGD). Takes the value 1 for all segments with a laryngeal constriction in which that constriction is not serving as an airstream initiator or primary place of articulation. Thus, 'glottalized' consonants, laryngealized vowels etc. have the value 1 for this variable. Various distinctions which have been made in the literature between 'pre-glottalized,' 'postglottalized,' 'creaky' etc. were judged to be inconsistently applied and hence it was decided not to represent such distinctions in the inventory.

67. Long (LONG). Takes the value 1 for contrastively long vowels or germinate consonants which are single units but are not generated as part of a general series of such long segments either from a suprasegmental length feature or because adjacent identical segments occur. (See 'overshort' below.)

68. Voiced aspirated, breathy, murmured (BREATHY). Takes the value 1 for all segments characterized by breathiness.

69. Overshort (OVRSHORT). In certain languages, the basic series of vowels is longer than a subset which may occur contrastively short. This situation led to proposing the variable 'overshort' which takes the value 1 for such contrastively short segments.

70. Preaspirated (PREASPRT).

j) The anomaly variable.

80. Anomaly (ANOMALY). This variable indicates segments with a doubtful or marginal status in an inventory. Unlike other variables, 'anomaly' may take values greater than 1. A value of 0 indicates a segment is a normal member of the inventory of the language. The other values have the meanings described below:

1--indicates a segment of extremely low frequency (e.g. it only occurs in a handful of words or certain morphological markers).

2--indicates a segment that occurs only in foreign words or unassimilated loans but these are frequent enough to consider including the segment in the inventory.

3--indicates a segment which is posited in underlying forms but which is neutralized in surface forms.

4--indicates a segment which is sometimes treated as phonemic but which may be regarded as derived from other underlying segments.

5--indicates a segment which although apparently a genuine member of the inventory, is described in particularly obscure or contradictory fashion (one example being a segment in Ashuslay which is described as simultaneously a (velar) stop and a lateral).

The values 3 and 4 were provided in order to provide m an escape from unresolved questions concerning the size of inventories. In practice, very little use has been made of these indications as a resolution has been achieved in nearly all cases.

7. Using UPSID

UPSID, containing 317 languages at the time of writing, consists of 9957 records of the form described in the preceding section. These constitute a SAS (Helwig and Council, 1979) system file on a resident disk pack at UCLA's Central Computing facility. SAS is a very flexible data manipulation and statistical analysis system. It conveniently permits selection of subsets of records, computation of language-by-language totals or frequency counts on the complete data set, creation of new variables for special purposes and many other operations to be performed. Of course, it will also straight-forwardly print out the information in the file. A printout of the information in the inventory for Spanish is given in Figure 1.

SAMPLE OF DATA FROM THE PHONOLOGICAL SEGMENT INVENTORY DATABASE (Language 11 - Spanish)

S T A T I S T I C A L   A N A L Y S I S   S Y S T E M

LANGNO=11

SEGNO	SEGCODE	BILABIAL	LABDENT	DENTAL	UNSPDENT	ALVEOLAR	PALATALV	RETROFLX	PALATAL	VELAR	UVULAR	PHARYNGL	GLOTTAL	PLOSIVE	IMPLOSIV	EJECTSTP	CLICK	FRICATIV	EJECTFRC	AFFRICAT	EJECTAFF	AFFCLICK
1	P	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
2	TD	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
3	K	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
4	CH	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
5	F	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
6	BF	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
7	TF	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
8	DF	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
9	SD	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
10	X	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
11	GF	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
12	M	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ND	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	NY	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
15	LD	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	LY	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
17	RT	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18		0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	J	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
23	O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	W	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0

SEGNO	SEGCODE	UNSPECR	TAP	FLAP	TRILL	APPROXMT	NASAL	SIMPVOWL	DIPTHONG	LATERAL	SIBILANT	LABLZED	PALTLZED	VELRZED	PHARGZED	NASLZED	NASRELSE	HIGH	HIGHMID	MID	LOWMID	LOW
1	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	TD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	CH	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
5	F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	BF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	TF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	DF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	SD	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
10	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	GF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	M	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ND	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	NY	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	LD	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
16	LY	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
17	RT	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18		0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	I	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
20	E	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
21	A	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
22	J	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	O	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
24	U	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
25	W	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

SEGNO	SEGCODE	FRONT	CENTRAL	BACK	NONPERIF	ROUNDED	UNROUNDD	LIPCOMP	RCOLORED	BACKING	LOWERING	ROUNDING	VOICELES	VOICED	ASPIRATD	LARYNGD	LONG	BREATHY	OVRSHORT	PREASPRT	ANOMALY
1	P	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
2	TD	0	0	0	0	0	0	0	0	0	0	0	1	0	C	0	0	0	0	0	0
3	K	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
4	CH	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
5	F	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
6	BF	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
7	TF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	DF	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
9	SD	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
10	X	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
11	GF	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
12	M	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
13	ND	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	NY	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
15	LD	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
16	LY	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
17	RT	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
18		0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
19	I	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
20	E	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
21	A	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
22	J	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
23	O	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
24	U	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
25	W	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0

Most segments are characterised by the value 1 for a **small** number of features **which amount** to the equivalent of a conventional phonetic label for the segment; for example segment 1 has the value 1 for the features BILABIAL PLOSIVE VOICELESS, Segment 11 for VELAR FRICATIVE VOICED, Segment 18 for ALVEOLAR TRILL VOICED, and segment 24 for SIMPLE-VOWEL HIGH BACK ROUNDED VOICED.

Figure 1.

Without duplicating procedural information, which is more properly sought in the SAS manual, an example of the use of UPSID will be given. This it is hoped, will help to clarify the conceptual steps involved in getting answers from the data.

Suppose that one wishes to use the UPSID data to check the hypothesis that no nasal consonant appears in a language unless a stop occurs at the same place of articulation. As a first step it is necessary to define more precisely what is intended by the terms such as 'stop,' 'nasal,' 'place of articulation' in this hypothesis and then to translate this into the appropriate set of variable values in UPSID. Let us assume that what is intended by 'stops' are plosives and central affricates, but not lateral affricates, implosives, ejectives, clicks etc., and 'nasals' are voiced nasal consonants. Let us also assume that 'place of articulation' refers to the usual set of primary places of articulations (See Section 7 (c) above), that double articulations count as separate 'places' but that secondary articulations will not be taken into account (e.g. /ŋ/ and /ŋʷ/ will both count as velar). We may also decide to disregard any of the 'anomalous' segments identified by a non-zero value for the anomaly variable. The data to be selected for examination from the main file is now clear. It includes any segment with the value zero on the anomaly variable and meeting one of the following specifications: (i) it has the value 1 for the variable 'plosive'; (ii) it has the value 1 for the variable 'affricate' and the value 0 for 'lateral'; (iii) it has the value 1 for the variables 'nasal' and 'voiced'. All other segments can be excluded and all variables except for 'nasal' and those relating to places of articulation can be dropped from the new, more compact, data file which is to be tested. Among places of articulation, pharyngeal and glottal can also be dropped as nasals cannot be produced at these places.

Now, we may restate the hypothesis as it relates to this reduced data set. The question to be answered is: for each segment with a given language number and the value 1 for 'nasal', is there a segment with a matching language number, matching values for the place of articulation features and the value 0 for the variable 'nasal'?

There are a number of ways one might proceed in order to obtain the answer to this question, but one simple way is to create an index which represents in a single number the information that one is interested in. There are ten relevant place of articulation variables, plus one double articulation (labial-velar) that is found for nasals. One relatively straight-forward procedure would be to generate eleven additional variables, one for each place at which nasals occur, set these to some conveniently large value, say 10, in case a segment is a nasal at the given place. The value

1 can then be changed to 0 for the original place variables for those nasals. The next step is to generate an additional variable for labial-velar stops and change the 1 values for 'labial' and 'velar' to 0 in these segments. One can then sum the totals for each variable on a language-by-language basis. This will produce a single line of data for each language in which each non-zero value on one of the eleven new nasal place variables indicates that a nasal appears at that place, and each non-zero value for a given original place of articulation variable indicates that a given plosive or affricate appears at that place in the language. The next step would be to add the value for the corresponding original place variable to the nasal place variable for each case where there was a non-zero nasal place value. This will produce a two-digit number which indicates whether or not each nasal is matched by a plosive at the same place. Any nasals remaining unmatched will be indicated by a value of 10 (or a multiple of 10) for the relevant nasal place variable. Any such cases can then be listed and theirm frequency, distribution by place, etc., examined. For example, a language with the inventory of stops, affricates and nasals shown in (1) would end up with the value 12 (1 nasal, 2 plosives) at the bilabial and dental places of articulation and with value 24 (2 nasals, 4 plosives) at the velar

(1)

p	b	t̪	d̪	tʃ		k	g	k^w	g^w
	m		n		ɲ		ŋ		$ŋ^w$

place of articulation. The palatal nasal place variable would have the value 10 , indicating there was no palatal plosive or affricate matching the palatal nasal. This language would then contain one of the listed exceptions to the hypothesis under investigation. (However, in this case one might have decided it was desirable to match a palatal nasal with a palato-alveolar affricate since these are more frequent than palatal plosives. This could be done by adding the values for the original 'palato-alveolar' and 'palatal' variables together for each record to produce a more broadly defined palatal variable, producing for the data in (1) an index value of 11 (1 nasal, 1 affricate) and no exception to the hypothesis.)

The remaining papers in this volume illustrate some of the analysis of the UPSID file which has been done so far. We believe that they contain valuable new insights into the structure of phonological inventories and useful verification or correction of several previously proposed hypotheses.

REFERENCES

Bell, A. (1978) "Language samples" in Universals of Human Language 1 Method and Theory, (ed.) J.H. Greenberg et al., Stanford University Press: 123-156.

Bender, M.L. (1976) "Nilo-Saharan overview," The Non-Semitic Languages of Ethiopia, (ed. M.L. Bender), Monograph 5, Committee on Ethiopian Studies, African Studies Center, Michigan State University, East Lansing: 439-483.

Ferguson, C.A. (1963) "Some assumptions about nasals" Universals of Language (ed. J.H. Greenberg), MIT Press, Cambridge: 42-47.

Greenberg, J.H. (1966) The Languages of Africa. University of Indiana, Bloomington.

Greenberg, J.H. (1970) "Some generalizations concerning glottalic consonants, especially implosives" IJAL 36: 123-145.

Greenberg, J.H. (1971) "Nilo-Saharan and Meroitic" Current Trends in Linguistics 7: Sub-Saharan Africa (ed. T. Sebeok), Mouton, The Hague: 421-442.

Helwig, J.T. and K.A. Council (1979) SAS User's Guide, 1979 Edition (Statistical Analysis System), SAS Institute, Raleigh.

Hurford, J.R. (1977) "The Significance of linguistic generalizations" Language 53: 574-620.

Hyman, L.M. (1977) "On the nature of linguistic stress" Studies on Stress and Accent (Southern California Occasional Papers in Linguistics 4) University of Southern California, Los Angeles.

Jakobson, R. & M. Halle (1956) "Phonology and Phonetics" (Part 1 of) Fundamentals of Language , Mouton, The Hague: 3-51.

Ladefoged, P., R. Glick and C. Criper (1972) "Mutual intelligibility" Language in Uganda, Oxford University Press, London: 65-68.

Ladefoged, P. (1971) Preliminaries to Linguistic Phonetics. University of Chicago Press, Chicago and London.

Troubetskoy, N. (1939) Grundzuge der Phonologie (Travaux du Cercle Linguistique de Prague 9) Prague.

Tucker, A.N. and M.A. Bryan (1956) The Non-Bantu Languages of Northeast Africa (Handbook of African Languages, Part 3) Oxford University Press for International African Institute, London.

Tucker A.N. and M.A. Bryan (1966) Linguistic Analyses: The Non-Bantu Languages of Northeastern Africa. Oxford University Press for International African Institute, London.

Vihman, M.M. (1974) "Excerpts from the Phonology Archive coding manual" Working Papers on Language Universals 15:141-153.

Appendix A

List of languages included and outline classification.

The language identification number assigned in UPSID is given after each language name and (S) indicates that the language is included in the Stanford Phonology Archive report used as a secondary data source.

Indo-European (000-049)
- Greek: Greek 000 (S)
- Celtic: Irish 001 (S), Breton 002 (S)
- Germanic: German 004 (S), Norwegian 006 (S)
- Baltic: Lithuanian 007 (S)
- Slavic: Russian 008 (S), Bulgarian 009 (S)
- Romance: French 010 (S), Spanish 011 (S), Romanian 012 (S)
- Iranian: Farsi 013 (S), Pashto 014 (S), Kurdish 015
- Indic: Hindi-Urdu 016 (S), Bengali 017 (S), Kashmiri 018 (S), Punjabi 019 (S), Sinhalese 020 (S)
- Albanian: Albanian 021 (S)
- Armenian: Eastern Armenian 022 (S)

Ural-Altaic (050-099)
- Finno-Ugric: Ostyak 050 (S), Cheremis 051 (S), Komi 052 (S), Finnish 053 (S), Hungarian 054 (S), Lappish 055
- Samoyed: Yurak 056 (S), Tavgy 057
- Turkic: Osmanli (Turkish) 058 (S), Azerbaijani 059 (S), Chuvash 060 (S), Yakut 061 (S), Kirghiz 062 (S), Bashkir 063, Khalaj 064, Tuva 065.
- Mongolian: Mongolian 066 (S)
- Tungus: Evenki 067 (S), Goldi 068, Manchu 069
- Korean: Korean 070 (S)
- Japanese: Japanese 071 (S)

Niger-Kordofanian (100-199)
- Kordofanian: Katcha 100 (S), Moro 101, Kadugli 102
- Mande: Kpelle 103 (S), Bisa 104, Bambara 105, Dan 106
- West Atlantic: Wolof 107 (S), Diola 108, Temne 109
- Voltaic: Dagbani 110 (S), Senadi 111, Tampulma 112, Bariba 113
- Kwa: Ewe 114 (S), Akan 115 (S), Igbo 116 (S), Ga 117 (S)
- Togo Remnant: Lelemi 118
- Cross River: Efik 119
- Plateau: Birom 120, Tarok (Yergam) 121, Amo 122
- Bantoid: Beembe 123 (S), Swahili 124 (S), Luvale 125 (S), Zulu 126 (S), Teke (Kukuya) 127
- Adamawa: Doayo 128
- Eastern: Gbeya 129 (S), Zande (130)

Nilo-Saharan (200-249)
- Songhai: Songhai 200 (S)
- Saharan: Kanuri 201 (S)
- Maban: Maba 202
- Fur: Fur 203

Eastern Sudanic: Maasai 204 (S), Luo 205 (S), Nubian 206 (S), Nyangi 207, Ik 208, Sebei 209, Tama 210, Temein 211, Nera 212, Tabi 213, Mursi 214
Central Sudanic: Logbara 215 (S), Yulu 216, Sara 217
Berta: Berta 218
Kunama: Kunama 219
Koman: Koma 220

Afro-Asiatic (250-299)
Semitic: Arabic 250 (S), Tigre 251 (S), Amharic 252 (S), Hebrew 253 (S), Socotri 254, Neo-Aramaic 255
Berber: Shilha 256 (S), Tuareg 257
Cushitic: Somali 258 (S), Awiya 259 (S), Iraqw 260 (S), Beja 261
Omotic: Kullo 262, Dizi 263, Kefa 264, Hamer 265
Chadic: Hausa 266 (S), Angas 267 (S), Margi 268 (S), Ngizim 269, Kanakuru 270

Austro-Asiatic (300-349)
Munda: Mundari 300 (S), Kharia 301 (S)
Khasi: Khasi 302 (S)
Vietmuong: Vietnamese 303 (S)
Bahnaric: Sedang 304 (S)
Khmer: Khmer 306 (S)

Australian (350-399)
Iwaidjan: Maung 350 (S)
Bureran: Burera 352
Tiwian: Tiwi 351
Nunggubuyan: Nunggubuyu 353 (S)
Maran: Alawa 354 (S)
Daly: Maranungku 355 (S), Malakmalak 356
Nyulnyulan: Bardi 357
Pama-Nyungan: Wik-Munkan 358 (S), Kunjen 359 (S), Western Desert 360 (S), Nyangumata 361 (S), Aranda 362, Kariera-Ngarluma 363, Gugu-Yalanji 364, Mabuiag 365, Arabana-Wanganura 366, Diyari 367, Bandjalang 368

Austro-Tai (400-499)
Kam-Tai: Standard Thai 400 (S), Lakkia 401 (S), Yay 402 (S), Sui 403, Saek 404, Po-ai 405, Lungchow 406
Atayalic: Atayal 407 (S)
West Indonesian: Sundanese 408 (S), Javanese 409 (S), Malagasy 410 (S), Cham 411 (S), Malay 412 (S), Batak 413 (S)
Philippine: Tagalog 414 (S), Sa'ban 415 (S), Chamorro 416 (S), Rukai 417
Formosan: Tsou 418
N.E. New Guinea: Adzera 419 (S), Roro 420
New Britain: Kaliai 421 (S)
Loyalty Is: Iai 422 (S)
Polynesian: Maori 423 (S), Hawaiian 424 (S)

Sino-Tibetan (500-599)

Sinitic: Mandarin 500 (S), Taishan 501 (S), Hakka 502 (S), Chang-chow (Wu) 503 (S), Amoy 504, Fuchow 505, Kan 506

Himalayish: Tamang 507

Mirish: Dafla 508 (S)

Lolo-Burmese: Burmese 509 (S), Lahu 510 (S)

Kachin: Jingpho 511

Kuki-Chin: Ao 512, Tiddim Chin 513

Baric: Garo 514 (S), Boro 515

Karenic: Karen 516 (S)

Miao-Yao: Yao 517 (S)

Indo-Pacific (600-699)

Andamanese: Andamanese 600

West New Guinea: Asmat 601 (S)

North New Guinea: Washkuk 602 (S), Sentani 603 (S), Nimboran 604, Iwam 605

South-East New Guinea: Telefol 606 (S)

Central New Guinea: Selepet 607 (S), Gadsup 608 (S), Yagaria 609, Kewa 610, Chuave 611, Pawaian 612, Dani 613, Wantoat 615, Daribi 616, Fasu 617

South New Guinea: Suena 618

North-East New Guinea: Dera 619

East New Guinea: Kunimaipa 620 (S), Yareba 621, Koiari 622, Taoripi 623

Bougainville: Nasioi 624 (S), Rotokas 625

Central Melanesian: Nambakaengo 626

Amerindian I (Northern) (700-799)

Haida: Haida 700 (S)

Tlingit: Tlingit 701

Athapaskan: Navaho 702 (S), Chipewyan 703 (S), Tolowa 704 (S), Hupa 705 (S),

Northern Penutian: Nez Perce 706 (S), Klamath 707

California Penutian: Maidu 708 (S), Wintu 709

Mexican Penutian: Chontal 710 (S), Zoque 711 (S), Tzeltal 712 (S), Totonac 713 (S), K'ekchi 714, Mixe 715

Oto-Manguean: Otomi 716 (S), Mazahua 717 (S), Mazatec 727 (S), Mixtec 728 (S), Chatino 729

Wakashan: Nootka 730 (S), Kwakw'ala 731 (S)

Chemakuan: Quileute 732

Salish: Squamish 733 (S), Puget Sound 734 (S)

Uto-Aztecan: Papago 736 (S), Luiseño 737 (S), Hopi 738 (S), Yaqui 739

Kiowa-Tanoan: Tiwa (Picuris) 740 (S)

Hokan: Karok 741 (S), Pomo 742 (S), Diegueño 743 (S), Achumawi 744, Yana 745, Shasta 746

Tarascan: Tarascan 747 (S)

Zuni: Zuni 748 (S)

Keres: Acoma 749

Macro-Algonkian: Ojibwa 750 (S), Delaware 751 (S), Tonkawa 752, Wiyot 753
Macro-Siouan: Seneca 754 (S), Wichita 755 (S), Dakota 756 (S), Yuchi 757 (S), Tunica 758 (S), Alabama 759 (S)
Wappo: Wappo 760

Amerindian II (Southern) (800-899)
Chibchan: Itonama 800 (S), Bribri 801, Mura 802
Paezan: Cayapa 803 (S), Paez 804 (S)
Witotoan: Ocaina 805 (S), Muinane 806
Carib: Carib 807 (S)
Macro-Ge: Apinaye 809 (S)
Pano-Tacanan: Amahuaca 810 (S), Chacobo 811 (S), Tacana 812, Cashinahua 813
Mataco: Ashuslay 814
Guaycuru: Abipon 815
Nambiquara: Southern Nambiquara 816
Zaparoan: Arabela 817, Auca 818
Quechumaran: Quechua 819 (S), Jaqaru 820 (S)
Chon: Gununa-Kena (Puelche) 821
Arawakan: Wapishana 822 (S), Island Carib 823 (S), Amuesha 824 (S), Campa 825 (S), Guajiro 826 (S), Moxo 827 (S)
Tupi: Guarani 828 (S), Sirionо 829 (S)
Guahibo-Pamigua: Guahibo 830
Tucanoan: Ticuna 831 (S), Barasano 832 (S), Siona 833, Tucano 834
Jivaroan: Jivaro 835 (S), Cofan 836
Penutian: Araucanian 837 (S)

Other Families (900-999)
Eskimo-Aleut: Greenlandic 900 (S), Aleut 901 (S)
Dravidian: Telugu 902 (S), Kota 903 (S), Kurukh 904 (S), Malayalam 905, Brahui 917
Paleo-Siberian: Ket 906 (S), Yukaghir 907 (S), Chukchi 908 (S), Gilyak 909 (S)
Caucasian: Georgian 910 (S), Kabardian 911 (S), Lak 912 (S)
Khoisan: Nama 913 (S), !xũ 918
Basque: Basque 914 (S)
Burushaski: Burushaski 915 (S)
Ainu: Ainu 916 (S)

Note The classifications given are not intended to be complete. Main subfamilies not listed do not have a representative in UPSID. In certain cases affiliation of a language (e.g. Cofan 836) is uncertain. Corrections and improvements are welcome.

Appendix B
Bibliography of Sources Used in Compiling UPSID

Part I: Alphabetic Index by Languages

In this index, languages are listed alphabetically by the names we have used throughout the UPSID project. In some cases a cross-reference from another well-known name is given to assist those who wish to verify if a given language is included. The language name is followed by the language identification number (see Appendix A). If a language is included in the Stanford Phonology Archive this is indicated by (S) and it may be assumed that the SPA report was used in determining the inventory for that language. For most of these languages, the sources used by SPA have also been directly consulted, and, for a few, additional or different sources were used. The remaining languages were contributed solely by UCLA. Sources are identified by author and date and a full list of references follows as Part II of this Appendix.

Note Special thanks to Hector Javkin and Diane Ridley for assistance in collecting copies of the sources.

Abipon 815 Najlis (1966)
Adzera 419 (S) Holzknecht (1973)
Achumani 744 Olmsted (1964, 1966)
Acoma 749 Miller (1966)
Ainu 916 (S) Simeon (1969)
Akan 115 (S) Welmers (1946), Schachter and Fromkin (1968), Stewart (1967)
Alabama 759 (S) Rand (1968)
Alawa 354 (S) Sharpe (1972)
Albanian 021 (S) Newmark (1957)
Aleut 901 (S) Bergsland (1956)
Amahuaca 810 (S) Osborn (1948)
Amharic 252 (S) Leslau (1968), Klingenheben (1966), Sumner (1957)
Amo 122 di Luzio (1972)
Amoy Chinese 504 [Hanyu Fangyan Gaiyao] (1960)
Amuesha 824 (S) Fast (1953)
Andamanese 600 Voegelin and Voegelin (1966), Radcliffe-Brown (1914)
Angas 267 (S) Burquest (1971)
Ao 512 Gowda (1972)
Apinayé 809 (S) Burgess and Ham (1968)
Arabela 817 Furne (1963)
Arabana-Wanganura 366 Hercus (1973)
Arabic 250 (S) Mitchell (1962), Tomiche (1964), Kennedy (1960)
Aranda 362 O'Grady, Voeglin and Voegelin (1966)
Araucanian 837 (S) Echeverria and Contreras (1965)
Armenian (Eastern) 022 (S) Allen (1950)
Ashuslay 814 Stell (1972)
Asmat 601 (S) Voorhoeve (1965)

Atayal 407 (S) Egerod (1966)
Auca 818 Saint and Pike (1962)
Awiya 259 (S) Hetzron (1969)
Aymara, see Jaqaru
Azande, see Zande
Azerbaijani 059 (S) Householder (1965)
Bambara 104 Bird, Hutchinson and Kante (1977)
Bandjalang 368 Cunningham (1969)
Barasano 832 (S) Stolte and Stolte (1971)
Bardi 357 Metcalfe (1971)
Bariba 113 Welmers (1952)
Bashkir 063 Poppe (1964)
Basque 914 (S) N'diaye (1970)
Batak 413 (S) van der Tuuk (1971)
Beembe 123 (S) Jacquot (1962)
Beja 261 Hudson (1976)
Bengali 017 (S) Ferguson and Choudhury (1960)
Berta 218 Triulzi, Dafallah and Bender (1976)
Birom 120 Wolff (1959)
Bisa 104 Naden (1973)
Boro 515 Bhat (1968)
Brahui 917 Emeneau (1935-37), De Armond (1975)
Breton 002 (S) Ternes (1970)
Bribri 801 Arroyo (1972)
Bulgarian 009 (S) Klagstad (1958), Aronson (1968)
Burera 352 Glasgow and Glasgow (1967)
Burmese 509 (S) Okell (1969)
Burushaski 915 (S) Morgenstierne (1945)
Campa 825 (S) Dirks (1953)
Carib 807 (S) Peasgood (1972), Hoff (1968)
Cashinahua 813 Kensinger (1963)
Cayapa 803 (S) Lindskoog and Brend (1962)
Chacobo 811 (S) Prost (1967)
Cham 411 (S) Blood (1967)
Chamorro 416 (S) Topping (1973, 1969), Seiden (1960)
Changchow Chinese (Wu) 503 (S) Chao (1970)
Chasta Costa, see Tolowa
Chatino 729 Pride (1965)
Cheremis 051 (S) Ristinen (1960)
Chipewyan 703 (S) Li (1946, 1933, 1932)
Chontal 710 (S) Keller (1959)
Chuave 611 Thurman (1970)
Chukchi 908 (S) Skorik (1968, 1961)
Chuvash 060 (S) Andreev (1966), Krueger (1961)
Cofan 836 Borman (1962)
Dafla 508 (S) Ray (1967)
Dagbani 110 (S) Wilson and Bendor-Samuel (1969)
Dakota 756 (S) Boas and Deloria (1939)
Dan 106 Béarth and Zemp (1967)
Dani 613 Bromley (1961), van der Stap (1966)
Daribi 616 Macdonald (1973)

Delaware 751 (S) Voegelin (1946)
Dera 619 Voorhoeve (1971)
Diegueño 743 (S) Langdon (1970)
Diyari 367 Austin (1978)
Dizi 263 Allen (1976b)
Diola 108 Sapir (1965)
Doayo 128 Wiering (1974)
Efik 119 Cook (1969)
Evenki 067 (S) Novikova (1960)
Ewe 114 (S) Berry (n.d. a), Stahlke (1971), Ladefoged (1968)
Fante, see Akan
Fasu 617 Loeweke and May (1964)
Farsi 013 (S) Obolensky, Panah and Nouri (1963)
Finnish 053 (S) Lehtinen (1964), Harms (1964)
French 010 (S) Sten (1963)
Fuchow Chinese 505 [Hanyu Fangyan Gaiyao] (1960)
Fur 203 Beaton (1968), Tucker and Bryan (1966)
Gã 117 (S) Berry (n.d. b), J.N.A. Nartey (p.c.)
Gadsup 608 (S) Frantz and Frantz (1966)
Garo 514 (S) Burling (1961)
Gbeya 129 (S) Samarin (1966)
Georgian 910 (S) Robins and Waterson (1952), Tschenkeli (1958), Vogt (1938, 1958)
German 004 (S) Moulton (1962), Philipp (1974)
Gilyak (Nivkh) 909 (S) Panfilov (1962, 1968)
Goldi 068 Avrorin (1968)
Greek 000 (S) Householder, Kazazis and Koutsoudas (1964)
Greenlandic 900 (S) Rischel (1974), Thalbitzer (1904), Kleinschmidt (1851)
Guahibo 830 Kondo and Kondo (1967)
Guajiro 826 (S) Holmer (1949)
Guarani 828 (S) Gregores and Suarez (1967), Uldall (1956), Lunt (1973)
Gugu-Yalanji 364 Oates and Oates (1964), Wurm (1972), Oates (1964)
Gununa-Kena 821 Gerzenstein (1968)
Haida 700 (S) Sapir (1923)
Hakka Chinese 502 (S) Hashimoto (1973)
Hamer 265 Lydall (1976)
Hausa 266 (S) Hodge (1947), Kraft and Kraft (1973), Hodge and Umaru (1963)
Hawaiian 424 (S) Pukui and Elbert (1965)
Hebrew 253 (S) Chayen (1973), Cohen and Zafrani (1968)
Hindi-Urdu 016 (S) Kelkar (1968), Vermeer and Sharma (1966)
Hopi 738 (S) Whorf (1946), Voegelin (1956)
Huambisa, see Jivaro
Hungarian 054 (S) Kalman (1972), Banhidi, Jokay and Szabo (1965), Hall (1938, 1944)
Hupa 705 (S) Woodward (1964), Golla (1970)
Iai 422 (S) Ozanne-Rivierre (1976), Tryon (1968)
Igbo 116 (S) Williamson (1969), Carnochan (1948), Swift, Ahaghota and Ugorji (1962)

Ignaciano, see Moxo
Ik 208 Heine (1975 b)
Inuit, see Greenlandic
Iraqw 260 (S) Whiteley (1958)
Irish 001 (S) Mhac an Fhailigh (1968)
Island Carib 823 (S) Taylor (1955)
Itonama 800 (S) Liccardi and Grimes (1968)
Iwam 605 Laycock (1965)
Japanese 071 (S) Bloch (1950), Martin (1952), Jorden (1963)
Jaqaru 820 (S) Hardman (1966)
Javanese 409 (S) Horne (1961)
Jingpho 511 Liu (1964)
Jivaro 835 (S) Beasley and Pike (1957)
Kabardian 911 (S) Kuipers (1960)
Kadugli 102 Abdalla (1973)
Kaliai 421 (S) Counts (1969)
Kan Chinese 506 [Hanyu Fangyan Gaijao] (1960)
Kanakuru 270 Newman (1974)
Kanuri 201 (S) Lukas (1937)
Karen (Sgaw) 516 (S) Jones (1961)
Kariera-Ngarluma 363 O'Grady, Voegelin and Voegelin (1966), Wurm (1972)
Karok 741 (S) Bright (1957)
Kashmiri 018 (S) Kelkar and Trisal (1964)
Katcha 100 (S) Stevenson (1957), Tucker and Bryan (1966)
K'ekchi 714 Haeseriju (1966), Freeze (1975)
Kefa 264 Fleming (1976)
Ket 906 (S) Dul'zon (1968), Krejnovich (1968)
Kewa 610 Franklin and Franklin (1962)
Khalaj 064 Doerfer (1971)
Khalka, see Mongolian
Kharia 301 (S) Biligiri (1965), Pinnow (1959)
Khasi 302 (S) Rabel (1961)
Khmer 306 (S) Huffman (1970 a, 1970 b), Jacob (1968)
Kirghiz 062 (S) Hebert and Poppe (1963)
Klamath 707 Barker (1964)
Koiari 622 Dutton (1969)
Koma 220 Tucker and Bryan (1966)
Komi 052 (S) Lytkin (1966), Bubrikh (1949)
Korean 070 (S) Martin (1951), Cho (1967), Martin and Lee (1969)
Kota 903 (S) Emeneau (1944)
Kpelle 103 (S) Welmers (1962), Hyman (1973)
Kullo 262 Allen (1976 a)
Kunama 219 Tucker and Bryan (1966)
Kunimaipa 620 (S) Pence (1966)
Kunjen 359 (S) Sommer (1969)
Kurdish 015 Abdulla and McCarus (1967)
Kurukh 904 (S) Pinnow (1964), Pfeiffer (1972)
Kwakw'ala 731 (S) S.R. Anderson (p.c.), Boas (1947)
Lahu 510 (S) Matisoff (1973)
Lak 912 (S) Murkelinskij (1967), Zhirkov (1955), Khaidakov (1966)
Lakkia 401 (S) Haudricourt (1967)

Lappish 055 Hasselbrink (1965)
Lelemi 118 Hoffmann (1971)
Lithuanian 007 (S) Senn (1966), Augustitis (1964), Ambrazas, Vajtkavichjute et al. (1966)
Logbara 215 (S) Crazzolara (1960), Tucker and Bryan (1966)
Luiseño 737 (S) Malécot (1963), Bright (1965, 1968), Kroeber and Grace (1960)
Lungchow 406 Li (1977)
Luo 205 (S) Gregersen (1961)
Luvale 125 (S) Horton (1949)
Maasai 204 (S) Tucker and Mpaayei (1955), Tucker and Bryan (1966)
Maba 202 Tucker and Bryan (1966)
Mabuias 365 Wurm (1972)
Maidu 708 (S) Shipley (1956, 1964)
Malagasy 410 (S) Dahl (1952), Dyen (1971)
Malakmalak 356 Tryon (1974), Birk (1975)
Malay 412 (S) Verguin (1967), Macdonald and Soenyono (1967)
Malayalam 905 Kumari (19720, McAlpin (1975), Velayudhan (1971)
Manchu 069 Austin (1962)
Mandarin Chinese 500 (S) Dow (1972), Chao (1968), Cheng (1973)
Maori 423 (S) Biggs (1961), Hohepa (1967)
Maranungku 355 (S) Tryon (1970)
Margi 268 (S) Hoffmann (1963)
Maung 350 (S) Capell and Hinch (1970)
Mazahua 717 (S) Spotts (1953)
Mazatec 727 (S) Pike and Pike (1947)
Mixe (Totontepec) 715 Crawford (1963), Schoenhals and Schoenhals (1965)
Mixtec 728 (S) Hunter and Pike (1969)
Mongolian (Khalka) 066 (S) Hangin (1968), Street (1963), Luvšanvandan (1964)
Moro 101 Black (1971)
Moxo 827 (S) Ott and Ott (1967)
Muinane 806 Walton and Walton (1967)
Mundari 300 (S) Gumperz and Biligiri (1957)
Mura 802 Sheldon (1974)
Mursi 214 Turton and Bender (1976)
Nama 913 (S) Beach (1938)
Nambakaengo 626 Wurm (1972 b)
Nambiquara (Southern) 816 Price (1976)
Nasioi 624 (S) Hurd and Hurd (1966)
Navaho 702 (S) Sapir and Hoijer (1967)
Nenets, see Yurak
Neo-Aramaic 255 Garbell (1965)
Nera 212 Thompson (1976)
Nez Perce 706 (S) Aoki (1970, 1966)
Ngizim 269 Schuh (1972)
Nimboran 604 Anceaux (1965)
Nivkh, see Gilyak
Nootka (Tseshaht) 730 (S) Sapir and Swadesh (1939, 1955)
Norwegian 006 (S) Vanvik (1972)

Nubiah (Mahas) 206 (S) Bell (1971)
Nunggubuyu 353 (S) Hughes and Leeding (1971)
Nyangi 207 Heine (1975 a)
Nyangumata 361 (S) O'Grady (1964)
Ocaina 805 (S) Agnew and Pike (1957)
Ojibwa 750 (S) Bloomfield (1956)
Osmanli (Turkish) 058 (S) Swift (1963), Lees (1961)
Ostjak 050 (S) Gulya (1966)
Otomi 716 (S) Blight and Pike (1976)
Paez 804 (S) Gerdel (1973)
Papago 736 (S) Hale (1959), Saxton (1963), I. Maddieson (p.c.)
Pashto 014 (S) Shafeev (1964)
Pawaian 612 Trefry (1972)
Persian, see Farsi
Po-ai 405 Li (1977)
Pomo 742 (S) Moshinsky (1974)
Puget Sound Salish 734 (S) Snyder (1968)
Punjabi 019 (S) Gill and Gleason (1963)
Quechua 819 (S) Bills, Vallejo and Troike (1969), Lastra (1968)
Quileute 732 Powell (1974)
Romanian 012 (S) Agard (1958), Ruhlen (1973)
Roro 420 Bluhme (1970), Davis (1974)
Rotokas 625 Firchow and Firchow (1969)
Rukai 417 Li (1973)
Russian 008 (S) Jones and Ward (1969), Halle (1959)
Sa'ban 415 (S) Clayre (1973)
Saek 404 Gedney (1970)
Sara 217 Caprile (1968), Thayer and Thayer (1971)
Sebei 209 Montgomery (1970)
Sedang 304 (S) Smith (1968)
Selepet 607 (S) McElhanon (1970)
Senadi 111 Welmers (1950)
Seneca 754 (S) Chafe (1967)
Sentani 603 (S) Cowan (1965)
Shasta 746 Silver (1964)
Shilha 256 (S) Applegate (1958)
Sinhalese 020 (S) Coates and da Silva (1960)
Siona 833 Wheeler and Wheeler (1962)
Siriono 829 (S) Priest (1968)
Somali 258 (S) Armstrong (1934), Andrzejewsky (1955)
Socotri 254 Johnstone (1975), Leslau (1938)
Songhai 200 (S) Prost (1956), Williamson (1967)
Spanish 011 (S) Navarro (1961), Harris (1969), Saporta and Contreras (1962)
Squamish 733 (S) Kuipers (1967)
Suena 618 Wilson (1969)
Sui 403 Li (1948)
Sundanese 408 (S) Van Syoc (1959), Robins (1953)
Swahili 124 (S) Polomé (1967)
Tabi 213 Tucker and Bryan (1966)
Tacana 812 Key (1968), Van Wynen and Van Wynen (1962)
Tagalog 414 (S) Bloomfield (1917), Schachter and Otanes (1972)

Tahaggart, see Tuareg
Taishan Chinese 501 (S) Cheng (1973)
Tama 210 Tucker and Bryan (1966)
Tamang 507 Mazaudon (1973)
Tampulma 112 Bergman, Gray and Gray (1969)
Taoripi 623 Brown (1973)
Tarascan 747 (S) Foster (1969)
Tarsk 121 Robinson (1974)
Tavgy 057 Castrén (1966), Tereščenko (1966)
Teke (Kukuya) 127 Paulian (1975)
Telefol 606 (S) Healey 1964
Telugu 902 (S) Lisker (1963), Krishnamurti (1961), Kelley (1963)
Temein 211 Tucker and Bryan (1966)
Temne 109 Dalby (1966), Wilson (1961)
Thai (Standard) 400 (S) Abramson (1962), Noss (1954, 1964)
Ticuna 831 (S) Anderson (1959), Anderson (1962)
Tiddim Chin 513 Henderson (1965)
Tigre 251 (S) Palmer (1962)
Tiwa (Picuris) 740 (S) Trager (1971)
Tiwi 351 Osborne (1974)
Tlingit 701 Story and Naish (1973), Swanton (1909)
Tolowa 704 (S) Bright (1964)
Tonkawa 752 Hoijer (1972, 1949)
Totonac 713 (S) Aschmann (1946)
Tseshaht, see Nootka
Tsou 418 Tung (1964)
Tuareg 257 Prasse (1972)
Tucano 834 West and Welch (1967)
Tunica 758 (S) Haas (1941)
Turkish, see Osmali
Tuva 065 Sat (1966)
Tzeltal 712 (S) Kaufman (1971)
Vietnamese 303 (S) Thompson (1965)
Wantoat 615 Davis (1969)
Wapishana 822 (S) Tracy (1972)
Wappo 760 Sawyer (1965)
Washkuk 602 (S) Kooyers, Kooyers and Bee (1971)
Western Desert 360 (S) Douglas (1955, 1964)
Wichita 755 (S) Garvin (1950), Rood (1975)
Wik-Munkan 358 Sayers and Godfrey (1964)
Wintu 709 Broadbent and Pitkin (1964)
Wiyot 753 Teeter (1964)
Wolof 107 (S) Sauvageot (1965)
Yagana 609 Renck (1967, 1975)
Yakut 061 (S) Krueger (1962), Böhtlingk (1964)
Yana 745 Sapir and Swadesh (1960)
Yao 517 (S) Purnell (1965)
Yaqui 739 Johnson (1962), Crumrine (1961)
Yareba 621 Weimer and Weimer (1972)
Yay 402 (S) Gedney (1965)
Yuchi 757 (S) Crawford (1973), Ballard (1975)

Yukaghir 907 (S) Krejnovich (1958, 1968)
Yulu 216 Thayer (1969), Santandrea (1970)
Yurak 056 (S) Hajdu (1963), Decsy (1966), Ristinen (1965, 1968)
Ịxũ 918 Snyman (1969, 1975)
Zande 130 Tucker and Hackett (1959)
Zoque 711 (S) Wonderly (1951)
Zulu 126 (S) Doke (1926, 1961)
Zuni 748 (S) Newman (1965)

Part II: List of References.

Note: The following abbreviations are used in the references:

BSOAS Bulletin of the School of Oriental and African Studies, London University.

BSOS Bulletin of the School of Oriental Studies, London University (earlier title of BSOAS).

IUPUAS Indiana University Publications, Uralic and Altaic Series, Indiana University, Bloomington.

JAOS Journal of the American Oriental Society.

JNSSSR Jazyki Narodov SSSR [Languages of the Peoples of the USSR]. Series under general editorship of V.V. Vinogradov, Nauka, Leningrad.

MIFAN Memoires de l'Institut français d'Afrique Noire (later, ... Institut fondamental), Dakar.

LSNA Linguistic Structures of Native America (Viking Fund Publications in Anthropology, 6) (ed. H. Hoijer) Wenner-Gren Foundation, New York.

NSLE The Non-Semitic Languages of Ethiopia (Monograph 5, Occasional Papers Series, Committee on Ethiopian Studies) (ed. M.L. Bender) African Studies Center, Michigan State University, East Lansing.

SILP Publications of the Summer Institute of Linguistics, Linguistic Series, (later titled Publications in Linguistics and Related Fields) Summer Institute of Linguistics, University of Oklahoma, Norman.

UCPL University of California Publications in Linguistics, University of California Press, Berkeley and Los Angeles.

References.

ABDALLA, A.I. (1973) Kadugli Language and Language Usage (Salambi Prizes Series 3), Khartoum University Press, Khartoum.

ABDULLA, J. and E.N. McCarus (1967) Kurdish Basic Course--Dialect of Sulaimania, Iraq, University of Michigan Press, Ann Arbor.

ABRAMSON, A.S. (1962) The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments (IJAL 28/2 Part II) (Indiana University Publications in Anthropology, Folklore and Linguistics 20), Indiana University, Bloomington.

AGARD, F.B. (1958) Structural Sketch of Rumanian (Language monograph 26), Linguistic Society of America, Baltimore.

AGNEW, A. and E.G. Pike (1957) "Phonemes of Ocaina (Huitoto)", IJAL 23, 24-27.

ALLEN, E.J. (1976 a) "Kullo", NSLE, 324-50.

ALLEN, E.J. (1976 b) "Dizi", NSLE, 377-92.

ALLEN, W.S. (1950) "Notes on the phonetics of an Eastern Armenian speaker", Transactions of the Philological Society, 180-206.

AMBRAZAS, V., V. Vajtkavichjute, A. Valjatskene, K. Morkunas, A. Sabaljauskas, and K. Ul'vidas (1966) "Litovskij jazyk", JNSSSR 1, 500-27.

ANCEAUX, J.C. (1965) The Nimboran Language: Phonology and Morphology, Nijhoff, The Hague.

ANDREEV, I.A. (1966) "Chuvashskij jazyk", Jazyki Narodov SSSR Vol. 2 (ed. V.V. Vinogradov), Nauka, Moscow, 43-65.

ANDERSON, D. (1962) Conversational Ticuna, Summer Institute of Linguistics, University of Oklahoma, Norman.

ANDERSON, L. (1959) "Ticuna vowels: with special regard to the system of five tonemes", Publicações do Museu Nacional, Série Lingüistica Especial I (Museu Nacional, Rio de Janeiro), 76-119.

ANDRZEJEWSKY, B.W. (1955) "The problem of vowel representation in the Isaaq dialect of Somali", BSOAS 17, 567-80.

AOKI, Haruo (1966) "Nez Perce vowel harmony and Proto-Sahaptian vowels", Language 42, 759-67.

AOKI, Haruo (1970) Nez Perce Grammar, UCPL 62.

APPLEGATE, J.R. (1958) An outline of the structure of Shilha (Program in Oriental Languages Publications Series B, No. 11), American Council of Learned Societies, New York.

ARONSON, H.I. (1968) Bulgarian Inflectional Morphophonology (Slavistic Printings and Reprintings 70), Mouton, The Hague.

ARMSTRONG, L.E. (1934) "The phonetic structure of Somali", Mitteilungen des Seminars für Orientalische Sprachen (Berlin) 37/3, 116-161.

ARROYO, V.M. (1972) Lenguas Indigenas Costarricenses (2nd ed.), Editorial Universitaria Centroamericana (EDUCA), San Jose.

ASCHMANN, H.P. (1946) "Totonaco phonemes", IJAL 12, 34-43.

AUGUSTITIS, D. (1964) Das Litauische Phonationssystem (Slavistische Beiträge, 12), Sagner, Munich.

AUSTIN, P.A. (1978) A Grammar of the Diyari Language of North-East South Australia, Ph. D. Thesis, Australian National University, Canberra.

AUSTIN, W. (1962) "The phonemics and morphophonemics of Manchu", American Studies in Altaic Linguistics, ed. N. Poppe, IUPUAS 13, 15-22.

AVRORIN, V.A. (1968) "Nanajski jazyk", JNSSSR 5, 129-48.

BALLARD, W.L. (1975) "Aspects of Yuchi morphonology", Studies in South-eastern Indian Languages (ed. J.M. Crawford), University of Georgia Press, Athens, GA, 164-87.

BANHIDI, Z., A. Jokay and D. Szabo (1965) Lehrbuch der Ungarischen Sprache, Publishing House for Textbooks, Budapest.

BARKER, M.A.R. (1964) Klamath Grammar, UCPL 32.

BEACH, D.M. (1938) The Phonetics of the Hottentot Language, Heffer, Cambridge.

BÉARTH, T. and H. Zemp (1967) "The phonology of Dan (Santa)", Journal of African Languages 6, 9-29.

BEASLEY, D. and K.L. Pike (1957) "Notes on Huambisa phonemics", Lingua Posnaniensis 6, 1-8.

BEATON, A.C. (1968) A Grammar of the Fur Language (Linguistic Monograph Series 1), Sudan Research Unit, University of Khartoum, Khartoum.

BELL, H. (1971) "The phonology of Nobiin Nubian", African Language Review 9, 115-39.

BERGMAN, R., I. Gray and C. Gray (1969) Collected Field Reports on the Phonology of Tampulma (Collected Language Notes 9), Institute of African Studies, University of Ghana, Legon.

BERGSLAND, K. (1956) "Some problems of Aleut phonology", For Roman Jakobson, ed. M. Halle, 38-43.

BERRY, J. (no date a) The pronunciation of Ewe, Heffer, Cambridge.

BERRY, J. (no date b) The pronunciation of Gã, Heffer, Cambridge.

BHAT, D.N.S. (1968) Boro Vocabulary (Deccan College Building Centenary and Silver Jubilee Series 59), Deccan College, Poona.

BIGGS, B. (1961) "The structure of New Zealand Maori", Anthropological Linguistics 3/3, 1-54.

BILIGIRI, H.S. (1965) Kharia: Phonology, Grammar and Vocabulary (Building Centenary and Silver Jubilee Series 3), Deccan College, Poona.

BILLS, G.D., B. Vallejo C. and R.C. Troike (1969) An Introduction to Spoken Bolivian Quechua, University of Texas Press for Institute of Latin Merican Studies, Austin.

BIRD, C., J. Hutchinson and M. Kante (1977) An Ka Bamanankan Kalan: Beginning Bambara, Indiana University Press, Bloomington.

BIRK, D.B.W. (1975) "The phonology of Malakmalak", Papers in Australian Linguistics 8 (Pacific Linguistics Series A, 39), 59-78.

BLACK, Mr. and Mrs. K. (1971) The Moro Language: Grammar and Dictionary (Linguistic Monograph Series, 6), Sudan Research Unit, University of Khartoum, Khartoum.

BLIGHT, R.C. and E.V. Pike (1976) "The phonology of Tenango Otomi", IJAL 42, 51-57.

BLOCH, B. (1950) "Studies in colloquial Japanese, IV: Phonemics", Language 26, 86-125.

BLOOD, D.L. (1967) "Phonological units in Cham", Anthropological Linguistics 9, 15-32.

BLOOMFIELD, L. (1956) Eastern Ojibwa, University of Michigan Press, Ann Arbor.

BLOOMFIELD, L. (1917) Tagalog Texts with Grammatical Analysis (University of Illinois Studies in Language and Literature 3/2-3), University of Illinois, Urbana.

BLUHME, H. (1970)"The phoneme system and its distribution in Roro", Pacific Linguistic Studies in Honor of Arthur Capell (ed. S.A. Wurm and D.C. Laycock) (Pacific Linguistics, Series C, 13), 867-77.

BOAS, F. and E. Deloria (1939) Dakota Grammar (Memoirs of the National Academy of Sciences 23/2).

BOAS, F. (1947) "Kwakiutl grammar", Transactions of the American Philosophical Society 37, 201-377.

BÖHTLINGK, O. (1964) Über die Sprache der Jakuten (repr. of Dr. A. Th. v. Middendorff's Reise in der äussersten Norden und Osten Sibiriens, Band III, St. Petersburg, 1851), IUPUAS 35.

BORMAN, M.B. (1962) "Cofan phonemes", Studies in Ecuadorian Indian Languages 1 (SILP 7), 45-49.

BRIGHT, J.O. (1964) "The phonology of Smith River Athapaskan (Tolowa)", IJAL 30, 101-07.

BRIGHT, W. (1957) The Karok Language, UCPL 13.

BRIGHT, W. (1965) "Luiseño phonemics", IJAL 31, 342-45.

BRIGHT, W. (1968) A Luiseño Dictionary, UCPL 51.

BROADBENT, S.M. and H. Pitkin (1964) "A comparison of Miwok and Wintu", Studies in Californian Linguistics (ed. W. Bright), UCPL 34, 19-45.

BROMLEY, H.M. (1961) The Phonology of Lower Grand Valley Dani (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 34), Nijhoff, The Hague.

BROWN, A.H. (1973) "The Eleman language family", The Linguistic Situation in the Gulf District and Adjacent Areas, Papua New Guinea (ed. K. Franklin) (Pacific Linguistics, Series C, 26), Australian National University, Canberra, 279-375.

BUBRIKH, D.V. (1949) Grammatika Literaturnogo Komi jazyka, Leningrad State University, Leningrad.

BURGESS, E. and P. Ham (1968) "Multilevel conditioning of phoneme variants in Apinayé", Linguistics 41, 5-18.

BURLING, R. (1961) A Garo Grammar (Deccan College Monograph Series, 25), Deccan College, Poona.

BURQUEST, D.A. (1971) A Preliminary Study of Angas Phonology (Studies in Nigerian Languages, 1), Institute of Linguistics, Zaria.

CAPELL, C. and H. E. Hinch (1970) Maung Grammar (Janua Linguarum Series Practica, 98), Mouton, The Hague.

CAPRILE, J.-P. (1968) "Essai de phonologie d'un parler mbay", Bulletin de la SELAF 8, 1-40.

CARNOCHAN, J. (1948) "A study on the phonology of an Igbo speaker", BSOAS 22, 416-27.

CASTRÉN, M.A. (1966) Grammatik der Samojedischen Sprachen IUPUAS 53. [Reprint of 1854 edition, St. Petersburg.]

CHAFE, W.L. (1967) Seneca Morphology and Dictionary (Smithsonian Contributions to Anthropology, 4), Smithsonian Institution, Washington D.C.

CHAO, Y.-R. (1968) A Grammar of Spoken Chinese, University of California Press, Berkeley.

CHAO, Y.-R. (1970) "The Changchow dialect", JAOS 90, 45-59.

CHAYEN, M.J. (1973) The Phonetics of Modern Hebrew (Janua Linguarum, Series Practica, 162), Mouton, The Hague.

CHENG, C.-C. (1973) A Synchronic Phonology of Mandarin Chinese (Monographs on Linguistic Analysis, 4), Mouton, The Hague.

CHENG, T. M. (1973) "The phonology of Taishan", Journal of Chinese Linguistics 1, 256-322.

CHO, S.-B. (1967) A Phonological Study of Korean (Acta Universitatis Upsaliensis, Studia Uralica et Altaica Upsaliensia, 2), Almqvist and Wiksells, Uppsala.

CLAYRE, I.F.C.S. (1973) "The phonemes of Sa'ban: a language of Highland Borneo", Linguistics 100, 26-46.

COATES, W.A. and M.W.S. da Silva (1960) "The segmental phonemes of Sinhalese", University of Ceylon Review 18, 163-75.

COHEN, D. and H. Zafrani (1968) Grammaire de l'Hebreu Vivant, Presses Universitaires de France, Paris.

COOK, T.L. (1969) The Pronunciation of Efik for Speakers of English, Indiana University, Bloomington.

COUNTS, D.R. (1969) A Grammar of Kaliai-Kove (Oceanic Linguistics Special Publication, 6), University of Hawaii Press, Honolulu.

COWAN, H.K.J. (1965) Grammar of the Sentani Language (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 47), Nijhoff, The Hague.

CRAWFORD, J.C. (1963) Totontepec Mixe Phonotagmemics (SILP 8).

CRAWFORD, J.M. (1973) "Yuchi phonology", IJAL 39, 173-79.

CRAZZOLARA, J.P. (1960) A Study of the Logbara (Ma'di) Language, Oxford University Press, London.

CRUMRINE, L.S. (1961) The Phonology of Arizona Yaqui (Anthropological Papers of the University of Arizona, 5), University of Arizona, Tucson.

CUNNINGHAM, M.C. (1969) A description of the Yugumbir dialect of Bandjalang (University of Queensland Faculty of Arts Papers, 1/8,) University of Queensland Press, Brisband.

DAHL, O.C. (1952) "Étude de phonologie et de phonétique malagache", Norsk Tidsskrift for Sprogvidenskap 16, 148-200.

DALBY, D. (1966) "Lexical analysis in Temne with an illustrative wordlist", Journal of West African Languages 3, 5-26.

DAVIS, D.R. (1969) "The distinctive features of Wantoat phonemes", Linguistics 47, 5-17.

DAVIS, M.M. (1974) "The dialects of the Roro language of Papua: a preliminary survey", Kivung 8, 3-22.

DE ARMOND, R.C. (1975) "Some rules of Brahui conjugation", Dravidian Phonological Systems (ed. H.G. Schiffman and C.M. Eastman), University of Washington Press, Seattle.

DECSY, G. (1966) Yurak Chrestomathy, IUPUAS 50.

DI LUZIO, A. (1972) "Preliminary description of the Amo language", Afrika und Übersee 56, 3-60.

DIRKS, S. (1953) "Campa (Arawak) phonemes", IJAL 19, 302-4.

DOERFER, G. (1971) Khalaj Materials, IUPUAS 115.

DOKE, C.M. (1926) "The phonetics of the Zulu language", Bantu Studies Special number.

DOKE, C.M. (1961) Textbook of Zulu Grammar, Longmans, Cape Town.

DOUGLAS, W.H. (1955) "Phonology of the Australian aboriginal language spoken at Ooldea, South Australia, 1951-1952", Oceania 25, 216-29.

DOUGLAS, W.H. (1964) An Introduction to the Western Desert Language (Oceania Linguistic Monographs, 4, revised ed.), University of Sydney, Sydney.

DOW, F.D.M. (1972) An Outline of Mandarin Phonetics (Oriental Monograph Series, 10), Faculty of Asian Studies, Australian National University, Canberra.

DUL'ZON, A.P. (1968) Ketskij Jazyk, Tomsk University, Tomsk.

DUTTON, T.E. (1969) The Peopling of Central Papua (Pacific Linguistics, Series B, 9), Australian National University, Canberra.

DYEN, I. (1971) "Malagasy", Current Trends in Linguistics 8: Linguistics in Oceania (ed. T.A. Sebeok), Mouton, The Hague, 211-39.

ECHEVERRIA, M.S. and H. Contreras (1965) "Araucanian phonemics", IJAL 31, 132-35.

EGEROD, S. (1966) "A statement on Atayal phonology", Artibus Asiae, Supplement 23, 120-30.

EMENEAU, M.B. (1935-37) "Phonetic observations on the Brahui language", BSOS 8, 981-83.

EMENEAU, M.B. (1944) Kota Texts, I, UCPL 2/1.

FAST, P.W. (1953) "Amuesha (Arawak) phonemes", IJAL 19, 191-94.

FERGUSON, C.A. and M. Chowdhury (1960) "The phonemes of Bengali", Language 36, 22-59.

FIRCHOW, I. and J. Firchow (1969) "An abbreviated phoneme inventory", Anthropological Linguistics 11, 271-76.

FLEMING, H.C. (1976) "Kefa (Gonga) Languages", NSLE, 351-76.

FOSTER, M.L. (1969) The Tarascan Language, UCPL 56.

FRANKLIN, K. and J. Franklin (1962) "Kewa I: phonological assymetry", Anthropological Linguistics 4/7, 29-37.

FRANTZ, C.I. and M.E. Frantz (1966) "Gadsup phoneme and toneme units", Pacific Linguistics Series A, 7, 1-11.

FREEZE, R.A. (1975) A Fragment of an Early Kekchí Vocabulary (University of Missouri Monographs in Anthropology, 2), University of Missouri, Columbia.

FURNE, R. (1963) "Arabela phonemes and high-level phonology", Studies in Peruvian Indian Languages 1 (SILP 9), 193-206.

GARBELL, I. (1965) The Jewish Neo-Aramaic Dialect of Peralan, Azerbaijan, Mouton, The Hague.

GARVIN, P.L. (1950) "Wichita I: phonemics", IJAL 16, 179-84.

GEDNEY, W.J. (1965) "Yay, a Northern Tai language in North Vietnam", Lingua 14, 180-93.

GEDNEY, W.J. (1970) "The Saek language of Nakhon Phanom province", Journal of the Siam Society 58, 67-87.

GERDEL, F. (1973) "Paez phonemics", Linguistics 104, 28-48.

GERZENSTEIN, A. (1968) Fonologia de la lengua Gününa-këna (Cuadernos de Lingüistica Indigena, 5), Centro de Estudios Lingüísticos, University of Buenos Aires, Buenos Aires.

GILL, H.S. and H.A. Gleason (1963) A Reference Grammar of Panjabi (Hartford Studies in Linguistics, 3), Hartford Seminary Foundation, Hartford.

GLASGOW, D. and K. Glasgow (1967) "The phonemes of Burera", Papers in Australian Linguistics I (Pacific Linguistics, Series A, 10), 1-14.

GOLLA, V.K. (1970) Hupa Grammar, Ph. D. Dissertation, University of California, Berkeley.

GOWDA, K.S.G. (1972) Ao-Naga Phonetic Reader, Central Institute of Indian Languages, Mysore.

GREGERSEN, E.A. (1961) Luo: A Grammar, Ph. D. Dissertation, Yale University, New Haven.

GREGORES, E. and J.A. Suarez (1967) A Description of Colloquial Guarani (Janua Linguarum, Series Practica, 27), Mouton, The Hague.

GULYA, J. (1966) Eastern Ostyak Chrestomathy, IUPUAS 51.

GUMPERZ, J.J. and H.S. Biligiri (1957) "Notes on the phonology of Mundari", Indian Linguistics 17, 6-15.

HAAS, M.R. (1941) "Tunica", Handbook of American Indian Languages, Vol. 4, Augustin, New York, 1-143.

HAESERIJU, E.V. (1966) Ensayo de la Gramática del K'ekchí, Suquinay, Purulha.

HAJDÚ, P. (1963) The Samoyed Peoples and Languages, IUPUAS 14.

HALE, K. (1959) A Papago Grammar, Ph. D. Dissertation, Indiana University, Bloomington.

HALL, R.A. (1938) An Analytical Grammar of the Hungarian Language (Language Monograph, 18), Linguistic Society of America, Philadelphia.

HALL, R.A. (1944) Hungarian Grammar (Language Monograph, 21), Linguistic Society of America, Baltimore.

HALLE, M. (1959) The Sound Pattern of Russian, Mouton, The Hague.

HANGIN, J.G. (1968) Basic Course in Mongolian, IUPUAS 73.

HANYU FANGYAN GAIYAO (1960) Wenzi Gaige Chubanshe, Beijing.

HARDMAN, M.J. (1966) Jaqaru: Outline of Phonological and Morphological Structure (Janua Linguarum, Series Practica, 22), Mouton, The Hague.

HARMS, R.T. (1964) Finnish Structural Sketch, IUPUAS 42.

HARRIS, J.T. (1969) Spanish Phonology (MIT Research Monograph, 54), MIT Press, Cambridge, MA.

HASHIMOTO, M.J. (1973) The Hakka Dialect (Princeton-Cambridge Studies in Chinese Linguistics, 5), Cambridge University Press, Cambridge.

HASSELBRINK, G. (1965) Alternative Analyses of the Phonemic System in Central-South Lappish, IUPUAS 49.

HAUDRICOURT, A.G. (1967) "La langue lakkia", Bulletin de la Societé de Linguistique de Paris 62, 165-82.

HEALEY, Alan (1964) Telefol Phonology (Pacific Linguistics, Series B, 3), Australian National University, Canberra.

HEBERT, R.J. and N. Poppe (1963) Kirghiz Manual, IUPUAS 33.

HEINE, B. (1975 a) "Tepes und Nyangi: zwei ostafrikanische Restsprachen", Afrika und Übersee 58, 263-300.

HEINE, B. (1975 b) "Ik--eine ostafrikanische Restsprache", Afrika und Übersee 59, 31-56.

HENDERSON, E.J. (1965) Tiddim Chin: A Descriptive Analysis of Two Texts, Oxford University Press, London.

HERCUS, L.A. (1973) "The prestopped nasal and lateral consonants in Arabana-Wanganura", Anthropological Linguistics 14, 293-305.

HETZRON, R. (1969) The Verbal System of Southern Agaw (University of California Publications in Near Eastern Studies, 12), University of California Press, Berkeley.

HODGE, C.T. and I. Umaru (1963) Hausa basic course, Foreign Service Institute, Washington D.C.

HODGE, C.T. (1947) An Outline of Hausa Grammar (Language Dissertation 41), Linguistic Society of America, Baltimore.

HOFF, B.J. (1968) The Carib Language (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 55), Nijhoff, The Hague.

HOFFMANN, C. (1963) A Grammar of the Margi Language, Oxford University Press for International African Institute, London.

HOFTMANN, H. (1971) The Structure of the Lelemi Language, Enzyklopädie, Leipzig.

HOHEPA, P.W. (1967) A Profile Generative Grammar of Maori (IJAL Memoir 20) (Indiana University Publications in Anthropology and Linguistics), Indiana University, Bloomington.

HOIJER, H. (1949) An Analytic Dictionary of the Tonkawa Language, UCPL 15.

HOIJER, H. (1972) Tonkawa Texts, UCPL 73.

HOLMER, Nils M. (1949) "Goajiro (Arawak) I: Phonology", IJAL 14, 45-56.

HOLZKNECHT, K.G. (1973) "The phonemes of the Adzera language", Pacific Linguistics Series A, 38, 1-11.

HORNE, E.C. (1961) Beginning Javanese, Yale University Press, New Haven.

HORTON, A.E. (1949) A grammar of Luvale, Witwatersrand University Press, Johannesburg.

HOUSEHOLDER, F.W. (1965) Basic course in Azerbaijani, IUPUAS 45.

HOUSEHOLDER, F.W., K. Kazazis and A. Koutsoudas (1964) Reference Grammar of Literary Dhimotiki (IJAL 30/2 Part II) (Publications of the Indiana University Research Center in Anthropology, Folklore and Linguistics, 31), Indiana University, Bloomington.

HUDSON, R.A. (1976) "Beja", NSLE, 97-132.

HUFFMAN, F.E. (1970 a) The Cambodian System of Writing and Beginning Reader, Yale University Press, New Haven.

HUFFMAN, F.E. (1970 b) Modern Spoken Cambodian (Yale Linguistic Series), Yale University Press, New Haven.

HUGHES, E.J. and V.J. Leeding (1971) "The phonemes of Nunggubuyu", Australian Aboriginal Studies 38 (Papers on the Languages of Australian Aboriginals), 72-81.

HUNTER, G.G. and E.V. Pike (1969) "The phonology and tone sandhi of Molinos Mixtec", Linguistics 47, 24-40.

HURD, C. and P. Hurd (1966) Nasioi Language Course, Department of Information and Extension Services, Port Moresby.

HYMAN, L.M. (1973) "Notes on the history of Southwestern Mande", Studies in African Linguistics 4, 183-96.

JACOB, J.M. (1968) Introduction to Cambodian, Oxford University Press, London.

JACQUOT, A. (1962) "Notes sur la phonologie du beembe (Congo)", Journal of African Languages 1, 232-42.

JOHNSON, J.B. (1962) El Idioma Yaqui, Instituto Nacional de Antropologia e Historia, Mexico.

JOHNSTONE, T.M. (1975) "The modern South Arabian languages", Afroasiatic Linguistics 1, 93-121.

JONES, D. and D. Ward (1969) The Phonetics of Russian, Cambridge University Press, Cambridge.

JONES, R.B. (1961) Karen Linguistic Studies, UCPL 25.

JORDEN, E.H. (1963) Beginning Japanese Part I (Yale Linguistic Series, 5), Yale University Press, New Haven.

KALMAN, B. (1972) "Hungarian historical phonology", The Hungarian Language (Janua Linguarum, Series Practica, 134) (ed. L. Benko and S. Imre), Mouton, The Hague.

KAUFMAN, T. (1971) Tzeltal Phonology and Morphology, UCPL 61.

KELKAR, A.R. and P.N. Trisal (1964) "Kashmiri word phonology: a first sketch", Anthropological Linguistics 6/1, 13-22.

KELKAR, A.R. (1968) Studies in Hindi-Urdu I: Introduction and word phonology, Postgraduate and Research Institute, Deccan College, Poona.

KELLEY, G. (1963) "Vowel phonemes and external vocalic sandhi in Telugu", JAOS 83, 67-73.

KELLER, K.C. (1959) "The phonemes of Chontal (Mayan)", IJAL 24, 44-53.

KENNEDY, N.M. (1960) Problems of Americans in Mastering the Pronunciation of Egyptian Arabic, Center for Applied Linguistics, Washington D.C.

KENSINGER, K.M. (1963) "The phonological hierarchy of Cashinahua", Studies in Peruvian Indian Languages 1 (SILP 9), 207-17.

KEY, M.R. (1968) Comparative Tacanan Phonology (Janua Linguarum, Series Practica, 50), Mouton, The Hague.

KHAIDAKOV, S.M. (1966) Ocherki po Lakskoj Dialektologii, Nauka, Moscow.

KLAGSTAD, H. (1958) "The phonemic system of colloquial standard Bulgarian", Slavic and East European Journal 16, 42-54.

KLEINSCHMIDT, S. (1851) Grammatik der Grönländischen Sprache, Reimer, Berlin.

KLINGENHEBEN, A. (1966) Deutsch-Amharischer Sprachführer, Harrassowitz, Wiesbaden.

KOOYERS, O., M. Kooyers and D. Bee (1971) "The phonemes of Washkuk (Kwoma)", Te Reo 14, 36-41.

KONDO, V. and R. Kondo (1967) "Guahibo phonemes", Phonemic Systems of Colombian Languages (SILP 14), 89-98.

KRAFT, C.H. and M.G. Kraft (1973) Introductory Hausa, University of California Press, Berkeley.

KREJNOVICH, E.A. (1958) Jukagirskij jazyk, Academy of Sciences of the USSR, Moscow and Leningrad.

KREJNOVICH, E.A. (1968) "Ketskij jazyk", JNSSSR 5, 453-73.

KREJNOVICH, E.A. (1968) "Jugagirskij jazyk", JNSSSR 5, 435-52.

KRISHNAMURTI, B. (1961) Telugu Verbal Bases, UCPL 24.

KROEBER, A.L. and G.W. Grace (1960) The Sparkman Grammar of Luiseño, UCPL 16.

KRUEGER, J.R. (1961) Chuvash Manual, IUPUAS 7.

KRUEGER, J.R. (1962) Yakut Manual, IUPUAS 21.

KUIPERS, A.H. (1960) Phoneme and Morpheme in Kabardian (Janua Linguarum, Series Minor, 8), Mouton, The Hague.

KUIPERS, A.H. (1967) The Squamish Language (Janua Linguarum, Series Practica, 73), Mouton, The Hague.

KUMARI, B.S. (1972) Malayalam Phonetic Reader, Central Institute of Indian Languages, Mysore.

LADEFOGED, P. (1968) A Phonetic Study of West African Languages (2nd ed.), Cambridge University Press, Cambridge.

LANGDON, M. (1970) A Grammar of Diegueño, Mesa Grande Dialect, UCPL 66.

LASTRA, Y. (1968) Cochabamba Quechua Syntax (Janua Linguarum, Series Practica, 40), Mouton, The Hague.

LAYCOCK, D.C. (1965) "Three Upper Sepik phonologies", Oceanic Linguistics 4, 113-17.

LEES, R.B. (1961) The Phonology of Modern Standard Turkish, IUPUAS 6.

LEHTINEN, M. (1964) Basic Course in Finnish, IUPUAS 27.

LESLAU, W. (1938) Lexique Socotri, Klincksiek, Paris.

LESLAU, W. (1968) Amharic Textbook, University of California Press, Berkeley.

LI, F-K (1932) "A list of Chipewyan stems", IJAL 7, 122-51.

LI, F-K (1933) "Chipewyan consonants", Ts'ai Yuan P'ei Anniversary Volume (Supplementary Volume 1), Bulletin of the Institute of History and Philology, Academia Sinica, 429-67.

LI, F-K (1946) "Chipewyan", Linguistic Structures of Native America (Viking Fund Publications in Anthropology, 6) (ed. H. Hoijer), Wenner-Gren Foundation, New York, 398-423.

LI, F-K (1948) "The distribution of initials and tones in the Sui language", Language 24, 160-67.

LI, F-K (1977) A Handbook of Comparative Tai (Oceanic Linguistics Special Publication, 15), University Press of Hawaii, Honolulu.

LI, P. J-K (1973) Rukai Structure, Special Publication 64, Institute of History and Philology, Academia Sinica, Taipei.

LICCARDI, M. and J. Grimes (1968) "Itonama intonation and phonemes", Linguistics 38, 36-41.

LINDSKOOG, J.N. and R.M. Brend (1962) "Cayapa phonemics", Studies in Ecuadorian Indian Languages 1 (SILP 7) (ed. B. Elson), 31-44.

LISKER, L. (1963) Introduction to Spoken Telugu, American Council of Learned Societies, New York.

LIV, L. (1964) "Ching-p'o-yü Kai-k'uang", Chung Kuo Yu Wen 5, 407-17.

LOEWEKE, E. and J. May (1964) "The phonological hierarchy in Fasu", Anthropological Linguistics 7/5, 89-97.

LUKAS, Johannes (1937) A Study of the Kanuri Language, Oxford University Press, London.

LUNT, H.G. (1973) "Remarks on nasality: the case of Guarani", A Festschrift for Morris Halle (ed. S.R. Anderson), Holt, Rinehart and Winston, New York, 131-39.

LUVSANVANDAN, Š. (1964) "The Khalkha-Mongolian phonemic system", Acta Orientalia (Academiae Scientiarum Hungaricae) 17, 175-85.

LYDALL, J. (1976) "Hamer", NSLE, 393-438.

LYTKIN, V.I. (1966) "Komi-Zyrjanskij jazyk", Jazyki Narodov SSSR, Vol. 3 (ed. V.V. Vinogradov), Nauka, Leningrad, 281-99.

MALÉCOT, A. (1963) "Luiseño, a structural analysis I: Phonology", IJAL 29, 89-95.

MACDONALD, G.E. (1973) "The Teberan Language Family", The Linguistic Situation in the Gulf District and Adjacent Areas, Papua New Guinea (Pacific Linguistics, Series C, 26) (ed. K. Franklin), 113-48.

MACDONALD, R.R. and D. Soenyono (1967) Indonesian Reference Grammar, Georgetown University Press, Washington D.C.

MARTIN, S.E. (1951) "Korean Phonemics", Language 27, 519-33.

MARTIN, S.E. (1952) Morphophonemics of Standard Colloquial Japanese (Language Dissertation, 47), Linguistic Society of America, Baltimore.

MARTIN, S.E. and Y.-S. C. Lee (1969) Beginning Korean, Yale University Press, New Haven.

MATISOFF, J.A. (1973) The Grammar of Lahu, UCPL 75.

MAZAUDON, M. (1973) Phonologie Tamang (Népal) (Collection "Tradition Orale", 4), SELAF, Paris.

MCALPIN, D.W. (1975) "The morphophonology of the Malayalam noun", Dravidian Phonological Systems (ed. H.F. Schiffmann and C. Easton), University of Washington, Seattle, 206-23.

MCELHANON, K.A. (1970) Selepet Phonology (Pacific Linguistics, Series B, 14), Australian National University, Canberra.

METCALFE, C.D. (1971) "A tentative statement of the Bardi aboriginal language", Papers on the Languages of Australian Aboriginals (Australian Aboriginal Studies, 38), 82-92.

MHAC AN FHAILIGH, E. (1968) The Irish of Erris, Co. Mayo, Dublin Institute for Advanced Studies, Dublin.

MILLER, W.R. (1966) Acoma Grammar and Texts, UCPL 40.

MITCHELL, T.F. (1962) Colloquial Arabic, English Universities Press, London.

MONTGOMERY, C. (1970) "Problems in the development of an orthography for the Sebei language of Uganda", Journal of the Language Association of East Africa 1, 48-55.

MORGENSTIERNE, G. (1945) "Notes on Burushaski Phonology", Norsk Tidsskrift for Sprogvidenskap 13, 61-95.

MOSHINSKY, J. (1974) A Grammar of Southeastern Pomo, UCPL 72.

MOULTON, W.G. (1962) The Sounds of English and German, University of Chicago Press, Chicago.

MURKELINSKIJ, G.B. (1967) "Lakskij jazyk", JNSSSR 4, 489-507.

NADEN, A.J. (1973) The Grammar of Bisa, Ph. D. Thesis, School of Oriental and African Studies, London University, London.

NAJLIS, E.L. (1966) Lengua Abipona I, Centro de Estudios Lingüísticos, Universidad de Buenos Aires.

NAVARRO, T.T. (1961) Manual de Pronunciación Española (10th ed.) (Publicaciones de la Revista de Filologia Española, 3), Consejo Superior de Investigaciones Cientificas, Madrid.

N'DIAYE, G. (1970) Structure du Dialecte Basque de Maya, Mouton, The Hague.

NEWMAN, P. (1974) The Kanakuru Language, Cambridge University Press, Cambridge.

NEWMAN, S. (1965) Zuni Grammar (University of New Mexico Publications in Anthropology, 14), University of New Mexico Press, Albuquerque.

NEWMARK, L. (1957) Structural Grammar of Albanian (IJAL 23/4 Part II), Indiana University Research Center in Anthropology, Folklore and Linguistics, Publication 4, Indiana University, Bloomington.

NOSS, R.B. (1954) An Outline of Siamese Grammar, Ph. D. Dissertation, Yale University, New Haven.

NOSS, R.B. (1964) Thai Reference Grammar, Foreign Service Institute, Washington D.C.

NOVIKOVA, K.A. (1960) Ocherki Dialektov Evenskogo Jazyka: Ol'skij Govor I, Academy of Sciences of the USSR, Moscow and Leningrad.

OATES, L.F. (1964) "Distribution of phonemes and syllables in Gugu-Yalanji", Anthropological Linguistics 1, 23-26.

OATES, W. and L.F. Oates (1964) "Gugu-Yalanji and Wik-Munkan Language Studies (Occasional Papers in Aboriginal Studies, 2), Australian Institute of Aboriginal Studies, Canberra.

OBOLENSKY, S., K.Y. Panah and F.K. Nouri (1963) Persian Basic Course, Center for Applied Linguistics, Washington D.C.

O'GRADY, G.N. (1964) Nyangumata Grammar (Oceania Linguistic Monographs, 9), University of Sydney, Sydney.

O'GRADY, G.N., C.F. Voegelin and F.M. Voegelin (1966) "Phonological diversity", Languages of the World: Indo-Pacific Fascicle 6 (Anthropological Linguistics 8/2), 56-67.

OKELL, J. (1969) A Reference Grammar of Colloquial Burmese, Oxford University Press, London.

OLMSTED, D.L. (1964) A History of Palaihnihan Phonology, UCPL 35.

OLMSTED, D.L. (1966) Achumawi Dictionary, UCPL 45.

OSBORN, H. (1948) "Amahuaca phonemes", IJAL 14, 188-90.

OSBORNE, C.R. (1974) The Tiwi Language (Australian Aboriginal Studies, 55), Australian Institute for Aboriginal Studies, Canberra.

OTT, W. and R. Ott (1967) "Phonemes of the Ignaciano Language", Linguistics 35, 56-60.

OZANNE-RIVIERE, F. (1976) Le Iaai, SELAF, Paris.

PALMER, F.R. (1962) The Morphology of the Tigre Noun (London Oriental Series, 13), Oxford University Press, London.

PANFILOV, V.Z. (1962) Grammatika Nivkhskogo Jazyka, I, Akademia Nauk, Moscow and Leningrad.

PANFILOV, V.Z. (1968) "Nivhskij jazyk", JNSSSR 5, 408-34.

PAULIAN, C. (1975) Le Kukya, Langue Teke du Congo (Bibliothèque de la SELAF, 49-50), SELAF, Paris.

PEASGOOD, E.T. (1972) "Carib phonology", Languages of the Guianas (ed. J.E. Grimes) (SILP 35), 35-41.

PENCE, Alan (1966) "Kunimaipa phonology: Hierarchical Levels", Pacific Linguistics, Series A, 7, 49-97.

PFEIFFER, M. (1972) Elements of Kurux Historical Phonology, E.J. Brill, Leiden.

PHILIPP, M. (1974) Phonologie des Deutschen, Kohlhammer, Stuttgart.

PIKE, K.L. and E.V. Pike (1947) "Immediate constituents of Mazateco syllables", IJAL 13, 78-91.

PINNOW, H.-J. (1959) Versuch einer Historischen Lautlehre der Kharia-Sprache, Harrassowitz, Wiesbaden.

PINNOW, H.-J. (1964) "Bemerkungen zur Phonetik und Phonemik des Kurukh", Indo-Iranian Journal 8, 32-59.

POLOMÉ, E.C. (1967) Swahili Language Handbook, Center for Applied Linguistics, Washington D.C.

POPPE, N. (1964) Bashkir Manual, IUPUAS 36.

POWELL, J.V. (1974) Proto-Chimakuan: Materials for a Reconstruction, Ph. D. Dissertation, University of Hawaii, Honolulu.

PRASSE, K.-G. (1972) Manuel de Grammaire Touarègue, Akademisk Forlag, Copenhagen.

PRICE, P.D. (1976) "Southern Nambiquara phonology", IJAL 42, 338-48.

PRIDE, K. (1965) Chatino Syntax, SILP 12.

PRIEST, P. (1968) "Phonemes of the Siriono Language", Linguistics 41, 102-8.

PROST, A. (1956) La Langue Songay et ses Dialectes, MIFAN 47.

PROST, G.R. (1967) "Phonemes of the Chacobo Language", Linguistics 35, 61-65.

PUKUI, M.K. and S.H. Elbert (1965) Hawaiian-English Dictionary (6th ed.), University of Hawaii Press, Honolulu.

PURNELL, H.C. (1965) Phonology of a Yao Dialect (Hartford Studies in Linguistics, 15), Hartford Seminary Foundation, Hartford.

RABEL, L. (1961) Khasi, a Language of Assam (LSU Studies, Humanities Series, 10), Louisiana State University Press, Baton Rouge.

RADCLIFFE-BROWN, A. (1914) "Notes on the languages of the Andaman Islands", Anthropos 9, 36-52.

RAND, E. (1968) "The structural phonology of Alabaman, a Muskogean language", IJAL 34, 94-103.

RAY, P.S. (1967) "Dafla phonology and morphology", Anthropological Linguistics 9, 9-14.

RENCK, G.L. (1967) "A tentative statement of the phonemes of Yagaria", Papers in New Guinea Linguistics 6 (Pacific Linguistics, Series A, 12), 19-48.

RENCK, G.L. (1975) A Grammar of Yagaria (Pacific Linguistics, Series B, 40), Australian National University, Canberra.

RISCHEL, J. (1974) Topics in West Greenlandic Phonology, Akademisk Forlag, Copenhagen.

RISTINEN, E.K. (1960) An East Cheremis Phonology, IUPUAS 1.

RISTINEN, E.K. (1965) "On the phonemes of Neneto", Ural-Altaische Jahrbücher 36, 154-64.

RISTINEN, E.K. (1968) "Problems concerning vowel length in Neneto", Ural-Altaische Jahrbücher 40, 22-44.

ROBINS, R.H. (1953) "The phonology of the nasalised verbal forms in Sundanese", BSOAS 15, 138-45.

ROBINS, R.H. and N. Waterson (1952) "Notes on the phonetics of the Georgian word", BSOAS 15, 55-72.

ROBINSON, J.O.S. (1974) "His and hers morphology: the strange case of the Tarsk possessives", Studies in African Linguistics Supplement 6, 201-9.

ROOD, D. (1975) "The implications of Wichita phonology", Language 51, 315-37.

RUHLEN, M. (1973) Rumanian Phonology, Ph. D. Dissertation, Stanford University, Palo Alto.

SAINT, R. and K.L. Pike (1962) "Auca phonemics", Studies in Ecuadorian Indian Languages 1 (ed. B. Elson) (SILP 7), 2-30.

SAMARIN, W.J. (1966) The Gbeya Language, UCPL 44.

SANTANDREA, S. (1970) Brief Grammar Outlines of the Yulu and Kara Languages (Museum Combonianum, 25), Nigrizia, Bologna.

SAPIR, E. (1923) "The phonetics of Haida", IJAL 3-4, 143-58.

SAPIR, E. and H. Hoijer (1967) The phonology and morphology of the Navaho language, UCPL 50.

SAPIR, E. and M. Swadesh (1939) Nootka Texts (William Dwight Whitney Linguistic Series), Linguistic Society of America, Philadelphia.

SAPIR, E. and M. Swadesh (1955) Native accounts of Nootka ethnography (IJAL 21/4 Part II), Indiana University Publications in Anthropology, Folklore and Linguistics, Indiana University, Bloomington.

SAPIR, E. and M. Swadesh (1960) Yana Dictionary, UCPL 22.

SAPIR, J.D. (1965) A Grammar of Diola-Fogny, Cambridge University Press, Cambridge.

SAPORTA, S. and H. Contreras (1962) A Phonological Grammar of Spanish, University of Washington Press, Seattle.

SAT, Š.Č. (1966) "Tuvinskij jazyk", JNSSSR 2, 387-402.

SAUVAGEOT, S. (1965) Description Synchronique d'un Dialecte Wolof: le Parler du Dyolof, MIFAN 73.

SAWYER, J.O. (1965) English-Wappo Vocabulary, UCPL 43.

SAXTON, D. (1963) "Papago phonemes", IJAL 29, 29-35.

SAYERS, B. and M. Godfrey (1964) Outline Description of the Alphabet and Grammar of a Dialect of Wik-Munkan, spoken at Coen, North Queensland (Occasional Papers in Aboriginal Studies, 2), Australian Institute of Aboriginal Studies, Canberra.

SCHACHTER, P. and V.A. Fromkin (1968) A Phonology of Akan: Akuapem, Asante, Fante (UCLA Working Papers in Phonetics, 9), Phonetics Laboratory, University of California, Los Angeles.

SCHACHTER, P. and F.T. Otanes (1972) Tagalog Reference Grammar, University of California Press, Berkeley.

SCHOENHALS, A. and L.C. Schoenhals (1965) Vocabulario Mixe de Totontepec (Vocabularios Indigenas, 14), Summer Institute of Linguistics, Mexico.

SCHUH, R. (1972) Aspects of Ngizim Syntax, Ph. D. Dissertation, University of California, Los Angeles.

SEIDEN, W. (1960) "Chamorro phonemes", Anthropological Linguistics 2/4, 6-33.

SENN, A. (1966) Handbuch der Litauischen Sprache, Band I: Grammatik, Winter, Heidelberg.

SHAFEEV, D.A. (1964) A Short Grammatical Outline of Pashto (Publications in Anthropology, 33) (trans. and ed. by H.H. Paper) Indiana University, Bloomington.

SHARPE, M. (1972) Alawa Phonology and Grammar (Australian Aboriginal Studies, 37), Australian Institute of Aboriginal Studies, Canberra.

SHELDON, S.N. (1974) "Some morphophonemic and tone perturbation rules in Mura-Pirahá", IJAL 40, 279-82.

SHIPLEY, W.F. (1956) "The phonemes of Northeastern Maidu", IJAL 22, 233-37.

SHIPLEY, WOF. (1964) Maidu Grammar, UCPL 41.

SILVER, S. (1964) "Shasta and Karok: a binary comparison", Studies in Californian Linguistics (ed. W. Bright) (UCPL 34), 170-81.

SIMEON, G. (1969) "Hokkaido Ainu phonemics", JAOS 89, 751-57.

SKORIK, P.I. (1961) Grammatika Chukotskogo Jazyka, I, Akademia Nauk, Moscow and Leningrad.

SKORIK, P.I. (1968) "Chukotskij jazyk", JNSSSR 5, 248-70.

SMITH, K.D. (1968) "Laryngealization and de-laryngealization in Sedang phonemics", Linguistics 38, 52-69.

SNYDER, W. (1968) "Southern Puget Sound Salish: phonology and morphology", Sacramento Anthropology Society Papers 8, 2-22.

SNYMAN, J.W. (1969) An Introduction to the Ịxũ Language, Balkema, Cape Town.

SNYMAN, J.W. (1975) Žu ǀ'hõasi Fonologie en Woordeboek, Balkema, Cape Town.

SOMMER, B.A. (1969) Kunjen Phonology: Synchronic and Diachronic (Pacific Linguistics, Series B, 11), Australian National University, Canberra.

SPOTTS, H. (1953) "Vowel harmony and consonant sequences in Mazahua (Otomi)", IJAL 19, 253-258.

STAHLKE, H. (1971) Topics in Ewe phonology, Ph. D. Dissertation, University of California, Los Angeles.

STELL, N.N. (1972) Fonología de la lengua a tu taj (Cuadernos de Lingüística Indigena, 8) Centro de Estudios Lingüísticos, University of Buenos Aires, Buenos Aires.

STEN, H. (1963) Manuel de Phonétique Française, Munksgaard, Copenhagen.

STEVENSON, R.C. (1957) "A survey of the phonetics and grammatical structure of the Nuba mountain languages IV", Afrika und Übersee 61, 27-65.

STEWART, J.M. (1967) "Tongue root position in Akan vowel harmony", Phonetica 16, 185-204.

STOLTE, J. and N. Stolte (1971) "A description of Northern Barasano phonology", Linguistics 75, 86-92.

STORY, G.L. and C.M. Naish (1973) Tlingit Verb Dictionary, Alaska Native Language Center, University of Alaska, Fairbanks.

STREET, J.C. (1963) Khalkha Structure, IUPUAS 24.

SUMNER, C. (1957) Étude Expérimentale de L'Amharique Moderne, University College Press, Addis-Ababa.

SWANTON, J.T. (1909) Tlingit Myths and Texts (Bureau of American Ethnology Bulletin, 39), Smithsonian Institution, Washington D.C.

SWIFT, L.B. (1963) A Reference Grammar of Modern Turkish, IUPUAS 19.

SWIFT, L.B., A. Ahaghota and E. Ugorji (1962) Igbo Basic Course, Foreign Service Institute, Washington D.C.

TAYLOR, D. (1955) "The phonemes of the Hopkins dialect of Island Carib", IJAL 21, 233-41.

TEETER, K.V. (1964) The Wiyot Language, UCPL 37.

TEREŠČENKO, N.M. (1966) "Nganasanskij jazyk", JNSSR 3, 438-57.

TERNES, E. (1970) Grammaire Structurale du Breton de L'Île de Groix, Winter, Heidelberg.

THALBITZER, W. (1904) A Phonetical Study of the Eskimo Language (Meddelelser om Grønland 31), Copenhagen.

THAYER, L.J. (1969) A Reconstructed History of the Chari Languages: Comparative Bonbo-Bagirmi-Sara Segmental Phonology with Evidence from Arabic Loanwords, Ph. D. Dissertation, University of Illinois, Urbana.

THAYER, L.J. and J.E. Thayer (1971) Fifty Lessons in Sara-Ngambay, University of Indiana, Bloomington.

THOMPSON, E.D. (1976) "Nera", NSLE, 484-94.

THOMPSON, L.C. (1965) A Vietnamese Grammar, University of Washington Press, Seattle.

THURMAN, R.C. (1970) "Chuave phonemic statement", Ms.

TOMICHE, N. (1964) Le Parler Arabe du Caire, Mouton, Paris and the Hague.

TOPPING, D.M. (1969) Spoken Chamorro, University of Hawaii Press, Honolulu.

TOPPING, D.M. (1973) Chamorro Reference Grammar, University of Hawaii Press, Honolulu.

TRACY, F.V. (1972) "Wapishana phonology", Languages of the Guianas (ed. J.E. Grimes) (SILP 35), 78-84.

TRAGER, F.H. (1971) "The phonology of Picuris", IJAL 37, 29-33.

TREFRY, D. (1972) Phonological Considerations of Pawaian (Oceania Linguistic Monographs, 15), University of Sydney, Sydney.

TRIULZI, A. A.A. Dafallah and M.L. Bender (1976) "Berta" NSLE, 513-32.

TRYON, D.T. (1968) Iai Grammar (Pacific Linguistics, Series B, 8), Australian National University, Canberra.

TRYON, D.T. (1970) An Introduction to Maranungku (Northern Australia) (Pacific Linguistics Series B, 15), Australian National University, Canberra.

TRYON, D.T. (1974) Daly Family Languages (Pacific Linguistics, Series C, 32), Australian national University, Canberra.

TSCHENKELI, K. (1958) Einführung in die Georgische Sprache I, Amirani, Zürich.

TUCKER, A.N. and M.A. Bryan (1966) Linguistic Analyses: the Non-Bantu Languages of North-Eastern Africa, Oxford University Press for International African Institute, London.

TUCKER, A.N. and P.E. Hackett (1959) Le Groupe Linguistique Zande (Annales du Musée Royal de l'Afrique Centrale Série in 8°, Sciences Humaines, 22), Tervuren.

TUCKER, A.N. and J.T.O. Mpaayei (1955) A Maasai Grammar (Publications of the African Institute, Leyden 2), Longmans, London.

TUNG, T.-H. (1964) A Descriptive Study of the Tsou Language, Formosa (Special Publication, 48), Institute of History and Philology, Academia Sinica, Taipei.

TURTON, D. and M.L. Bender (1976) "Mursi", NSLE, 533-62.

ULDALL, E. (1956) "Guaraní sound system", IJAL 20, 341-42.

VAN DER STAP, P.A.M. (1966) Outline of Dani Morphology (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 48), Nijhoff, The Hague.

VAN DER TUUK, H.N. (1971) A Grammar of Toba Batak, Nijhoff, The Hague.

VAN SYOC, W.B. (1959) The Phonology and Morphology of the Sundanese Language, Ph. D. Dissertation, University of Michigan, Ann Arbor.

VANVIK, A. (1972) "A phonetic-phonemic analysis of Standard Eastern Norwegian", Norwegian Journal of Linguistics 26, 119-64.

VAN WYNEN, D. and M.G. Van Wynen (1962) Fonemas Tacana y Modelos de Acentuación (Notas Lingüísticas de Bolivia, 6), Summer Institute of Linguistics, Cochabamba.

VELAYUDHAN, S. (1971) Vowel Duration in Malayalam, The Dravidian Linguistic Association of India, Trivandrum.

VERGUIN, J. (1967) Le Malais: Essai d'Analyse Fonctionelle et Structurale (Cahiers de l'Homme, n.s. 7), Mouton, Paris and The Hague.

VERMEER, H.J. and A. Sharma (1966) Hindi-Lautlehre, Groos, Heidelberg.

VOEGELIN, C.F. (1946) "Delaware: an Eastern Algonquian Language", Linguistic Structures of Native America (Viking Fund Publications in Anthropology, 6) (ed. H. Hoijer), Wenner-Gren Foundation, New York, 130-57.

VOEGELIN, C.F. (1956) "Phonemicizing for dialect study with reference to Hopi", Language 32, 116-35.

VOEGELIN, C.F. and F.M. Voegelin (1966) [Andamanese] Languages of the world: Indo-Pacific Fascicle 8 (Anthropological Linguistics 8/4), 10-13.

VOGT, H. (1938) "Esquisse d'une grammaire du georgien moderne", Norsk Tidsskrift for Sprogvidenskap 9, 5-114.

VOGT, H. (1958) "Structure phonemique du georgien", Norsk Tidsskrift for Sprogvidenskap 18, 5-90.

VOORHOEVE, C.L. (1965) The Flamingo Bay Dialect of the Asmat Language, Smits, The Hague.

VOORHOEVE, C.L. (1971) "Miscellaneous notes on languages in West Irian, New Guinea", Papers in New Guinea Linguistics 14 (Pacific Linguistics, Series A, 28), 47-114.

WALTON, J. and J. Walton (1967) "Phonemes of Muinane", Phonemic Systems of Colombian Languages (SILP 14), 37-47.

WEIMER, H. and N. Weimer (1972) "Yareba phonemes", Te Reo 15, 52-57.

WELMERS, W.E. (1946) A Descriptive Grammar of Fanti (Language Dissertation 39), Linguistic Society of America, Philadelphia.

WELMERS, W.E. (1950) "Notes on two languages of the Senufo group I: Senadi", Language 26, 126-46.

WELMERS, W.E. (1952) "Notes on the structure of Bariba", Language 28, 82-103.

WELMERS, W.E. (1962) "The phonology of Kpelle", Journal of African Languages 1, 69-93.

WEST, B. and B. Welch (1967) "Phonemic system of Tucano", Phonemic systems of Colombian Languages (SILP 14), 11-24.

WHEELER, A. and M. Wheeler (1962) "Siona phonemics", Studies in Ecuadorian Indian Languages 1 (SILP 7), 96-111.

WHITELEY, W.H. (1958) A Short Description of Item Categories in Iraqw, East African Institute of Social Research, Kampala.

WHORF, B.L. (1946) "The Hopi language, Toreva dialect", LSNA, 158-83.

WIERING, E. (1974) "The indicative verb in Doowaayãã̃yo", Linguistics 124, 33-56.

WILLIAMSON, K. (1967) "Songhai Word List (Gao Dialect)", Research Notes, Department of Linguistics and Nigerian Languages, University of Ibadan, Vol. 3, 1-34.

WILLIAMSON, K. (1969) "Igbo", Twelve Nigerian Languages (ed. E. Dunstan), Longmans, London, 85-96.

WILSON, D. (1969) "Suena phonology", Papers in New Guinea Linguistics 9 (Pacific Linguistics, Series A, 18), 87-93.

WILSON, W.A.A. (1961) An Outline of the Temne Language, School of Oriental and African Studies, University of London, London.

WILSON, W.A.A. and J.T. Bendor-Samuel (1969) "The phonology of the nominal in Dagbani", Linguistics 52, 56-82.

WOLFF, Hans (1959) "Subsystem typologies and area linguistics", Anthropological Linguistics 1/7, 1-88.

WOODWARD, M.F. (1964) "Hupa phonemics", Studies in Californian Linguistics (ed. W. Bright) UCPL 34, 199-216.

WONDERLY, W.L. (1951) "Zoque II: Phonemes and morphophonemes", IJAL 17, 105-23.

WURM, S. (1972 a) Languages of Australia and Tasmania, Mouton, The Hague.

WURM, S. (1972 b) "Notes on the indication of possession with nouns in Reef and Santa Cruz Islands languages", Papers in Linguistics of Melanesia 3 (Pacific Linguistics, Series A, 35), 85-113.

ZHIRKOV, L.I. (1955) Lakskij jazyk: fonetika i morfologija, Akademia Nauk, Moscow.

Phonological Generalizations from the UCLA Phonological Segment Inventory Database (UPSID)

Ian Maddieson

1. Introduction

A database designed to give more reliable and more readily accessible answers to questions concerning the distribution of phonological segments in the world's languages has been created at the Phonetics Laboratory at UCLA. UPSID (UCLA Phonological Segment Inventory Database) has been used to investigate a number of hypothesized phonological universals and 'universal tendencies'. Principal among these have been certain ideas concerning the overall size and structure of phonological inventories. This paper briefly describes the design of the database and then discusses some of the issues involving inventory size and structure which have been examined with its use.

2. Design of the Database

The languages included in UPSID have been chosen to approximate a properly constructed quota sample on a genetic basis of the world's extant languages. The quota rule is that only one language may be included from each small family grouping (e.g. one from West Germanic, one from North Germanic, etc.), but that each such family should be represented. Availability and quality of phonological descriptions is a factor in determining which language to include from within a group, but such factors as the number of speakers and the phonological peculiarity of the language are not considered. The database currently includes the segmental inventories of 317 languages. (See Maddieson 1980 for more details on UPSID.)

Each segment which is considered phonemic is represented by its most characteristic allophone, specified in terms of a set of 58 phonetic attributes. These are treated as variables which take the value 1 if the segment has the attribute and 0 if it lacks it. The list of attributes with the 1 specification thus provides a phonetic description of the segment. Figure 1 shows how the inventory of Spanish is represented in this system. (See previous article).

For 192 of the 317 languages included, UPSID has profited from the invaluable work done by the compilers of the Stanford Phonology Archive (SPA). Our decisions on phonemic status and phonetic description of segments do not always coincide with those reached in the SPA and we have sometimes examined additional sources, but a great deal of effort was saved by the availability of this source of standardized analyses. It should be noted that UPSID, unlike the SPA, makes no attempt to include information on allophonic variation or phonological rules.

In determining the segment inventories, there are two especially problematical areas. The first involves choosing between a unit or sequence interpretation of, for example, affricates, prenasalized stops, long (geminate) consonants and vowels, diphthongs, labialized consonants, etc. All available evidence which bears on the choice has been examined but with some prejudice in favor of treating phonetically complex events as sequences. The second involves choosing between a segmental and a suprasegmental analysis of certain properties. Stress and tone have always been treated as suprasegmental in themselves, but differences in segments which accompany tone or stress distinctions may be regarded as segmental contrasts if the association does not seem a natural one.

3. Variations in Size

The number of segments in a language may vary widely. The smallest inventories included in the survey have only 11 segments (Rotokas - Firchow and Firchow, 1969, and Mura - Sheldon, 1974) and the largest has 148 (!xu - Snyman, 1975). However it is clear that the typical size of an inventory lies between 20 and 35 segments - 65% of the languages in the survey fall within these limits. The mean number of segments per language is a little over 31; the modal number is 26 with 6.3% of the sampled languages having that number of segments.

Whether this strong tendency to have from twenty to thirty-five segments results from some factor which means that this is an optimum range is an open question. It seems likely that there is an upper limit on the number of segments which can be efficiently distinguished in speech, and a lower limit on the number of segments which are required to build an adequate vocabulary of distinct morphemes. But these limits would appear to lie respectively above and below the numbers 35 and 20. Consider the following. !xũ, with 148 segments, is related to languages which also have unusually large inventories. Diachronic study of these languages (Traill, 1978, Baucom, 1974) indicates that large inventories have been a stable feature which has persisted for a long period of time. If the number of efficiently distinguished segments was substantially smaller, there would be constant pressure to reduce the number of contrasts. There does not seem to be evidence of such pressure.* Similarly, the facts do not seem to show that small inventories (below 20 segments) suffer from problems due to lack of contrastive possibilities at the morphemic level. The symptoms of such difficulties would include unacceptably high incidence of homophony or unmanageably long morphemes. Dictionaries and vocabularies of several languages with small inventories, such as Rotokas (Firchow, Firchow and Akoitai, 1973), Hawaiian (Pukui and Elbert, 1965) and Asmat (Voorhoeve, 1965: 293-361), do not provide evidence that there are symptoms of stress of these kinds in languages with small phoneme inventories. Hawaiian, for example, has been calculated to have an average of just $3\frac{1}{2}$ phonemes per morpheme (Pukui and Elbert, 1965: xix). And again, diachronic evidence indicates that small inventory size may be a

* If such pressure existed on languages with large inventories it should also be manifested in higher error rates in tasks involving phoneme recognition. I know of no relevant experimental data on this point.

persistent phenomenon, as, for example in Polynesian (Grace, 1959).

The restrictions on inventory size may therfore not be theoretical ones relating to message density and channel capacity in language processing. Such considerations are far from being the only ones likely to influence the typical language inventory. Most languages exist in a multi-lingual social context. Limits may be placed on the typical size of an inventory through language contacts, especially the acquisition of new speakers. The mechanism may be one which approximates the following; speakers of an acquired language make substitutions for any segment that is not matched by a closely similar segment in their own language or is not capable of being generated by a simple process of adding familiar features (e.g., acquiring /g/ is easy if you already have /p, b, t, d/ and /k/). The resulting inventory contains only the segments common to both input languages, plus a few 'generated' segments. The smaller the inventory the greater the probability that some segments will be generated in the fashion outlined. The greater the inventory the smaller the probability that similar segments will coincide in the two languages. This proposal predicts not only that upper and lower limits on inventory size will tend to be established, but also that areal-genetic deviations from the central tendency should be expected. Thus, greater than average size inventories in Khoisan or Caucasian languages or smaller than average in Polynesian are expected. This is because local deviations are perpetuated due to contact with other languages tending in the same direction. This proposal also avoids a difficulty, as the need to assume that postulated human processing limitations fail to exert pressure to conform on the 'deviant' cases is avoided.

4. Relationship between Size and Structure

The data in UPSID have been used to address the question of the relationship between the size of an inventory and its membership. The balance between consonants and vowels within an inventory was calculated by dividing the number of vowels into the total number of segments. The resulting ratio varies between 1.76 and 16.0 but is most likely to fall in the range of 3.0 to 5.0 (some 50% of the sampled languages are within these limits). Larger inventories tend to be more consonant-dominated, but there is a tendency also for vowel inventories to be larger in the languages with larger inventories. This is shown by the significant weak correlation obtained between the number of consonants and the number of vowels in the 317 languages ($r = .39$). However, a large consonant inventory with few vowel contrasts is certainly possible, as, for example, in Haida (48C, 3V), Jaqaru (39C, 3V) or Burushaski (43C, 5V). Small consonant inventories with a large number of vowels seem the least likely to occur (cf. the findings of Hockett, 1955). There is thus some relationship between inventory size and consonant/vowel balance.

Study of the overall frequency of segments of each type in UPSID suggested an investigation of hypothesis that a significant amount of the variation between inventories can be captured by explaining structure from frequency. The simplest form of such a hypothesis would propose that languages differ in that the smallest inventories contain the most frequent segments and as the size of the inventory increases, segments are added in descending order of their overall frequency of occurrence.

This is akin to arranging all segments in a single hierarchy. One test of this hypothesis was made by examining the languages in UPSID that have the modal number of consonants, that is, 21, and comparing this with segment frequency and subinventory structure.

Languages are most likely to have between 8 and 10 stops (including affricates) but the scatter is quite wide (some 29% of languages fall within the limits mentioned). For fricatives, 2 to 4 is most likely (48%), and the same range is most likely for nasals (83%). Languages are most likely to have 2 liquids and 2 vocoid approximants (41% and 69% respectively). About 61% of languages have the consonant /h/,which is not included in any of the categories already named.

The most frequently occurring individual consonant segment types would form a 'modal' inventory containing the 20 consonants in (1) plus one

(1)

p, b	t, d	tʃ	k, g	ʔ
f	s	ʃ		
m	n	ɲ	ŋ	
w	l,r	j		h

other, perhaps /z/ or /ts/ which have about equal frequency, or less probably, /x/, /v/ or /dʒ/ which are somewhat less common. Aspirated stops /p^h, t^h, k^h/ are about as frequent as this last group of segments but almost invariably occur as a series, hence they are not separately plausible as the 21st consonant in this inventory.

This tally of the most frequent consonants builds an inventory that conforms to the predominant patterns for numbers of stops, fricatives, etc reported above. This is encouraging. However, none of the 27 languages in UPSID with 21 consonants contains all 20 of the segments in (1). Fur only deviates by lacking /ʔ/ (it has /z/ and /dʒ/). Some of the remaining 26 languages have as few as 10 of the segments in (1). These include two Australian languages, but the 10 in Arabana-Wanganura are not identical to the 10 in Kariera-Ngarluma, and neither are a proper subset of the 15 in Khalaj or the 16 in Thai. These facts are among numerous indications which could be cited showing that this proposal accounts very poorly for the patterns according to which inventories are constructed.

5. Other Proposals on the Structure of Inventories

Although the idea of a single hierarchy cannot be sustained, clearly there are many strong implicational hierarchies between particular types of segments (although few are exceptionless). Some examples of these, validated by data in UPSID though often originally proposed elsewhere, are given in (2).

(2) i) /p/ does not occur without /t/.

ii) Nasal consonants do not occur unless oral obstruents occur at broadly-speaking the same place of articulation.

iii) Voiceless nasals and approximants do not occur unless the language has the voiced counterparts.

iv) Voiced plosives do not occur unless voiceless plosives occur at the same places of articulation.

v) Mid vowels do not occur unless high and low vowels occur.

vi) Rounded front vowels do not occur unless unrounded front vowels of matching height occur.

vii) /ø/ and/or /œ/ do(es) not occur unless /y/ occurs.

Yet, as shown in section 4, such observations cannot be compiled into a single composite hierarchy. At the very least, alternate choices must be built in at certain points. This is because equally valid general prohibitions on the co-occurrence of segments within an inventory can also be found. Some of these are given below in (3).

(3) i) A language does not contain both (voiced) implosives and laryngealized plosives at the same place of articulation.

ii) A language does not contain both /Φ/ and /f/ or both /β/ and /v/ (Tarascan and Ewe are the only exceptions in UPSID).

iii) A language does not contain a voiceless lateral fricative and a voiceless lateral approximant.

iv) A language does not include a dental stop, nasal or lateral and an alveolar stop, nasal or lateral of the same type (There are more exceptions to this observation but significantly fewer than would be anticipated if co-occurrence was free).

The statements in (3) could be subsumed under a general observation that segments do not (usually) function contrastively unless they are sufficiently phonetically distinct. The choices here are all between phonetically similar segments; the types of segments referred to could be collapsed under more inclusive labels (e.g. /β/ and /v/ = voiced labial fricatives) and regarded as among the more extreme cases of noncontrastive phonetic differences. This is a phenomenon that phoneticians have long been aware of but which is currently the subject of focussed research at UCLA (Ladefoged, 1978). That is, there are measureable differences between generally similar segments in different languages along parameters that are not known to serve as the basis for phonemic contrasts in any language. If this approach is adopted the difference between, say, stops at the dental and alveolar places of articulation is a marginal member of a class of distinctions unavailable for meaningful contrast within languages. A more typical member of this class would be, say, a difference in relative timing of the release of the oral and glottal closures in the production of ejectives.

This interpretation of prohibitions on co-occurrence such as those in (3) introduces a concept of phonetic distance or phonetic salience as an explanatory factor in the design of phonological inventories. While such ideas have principally beendiscussed in relation to vowel inventories (cf. Crothers 1978), they can be extended to the whole inventory. Implicational hierarchies can be reinterpreted as involving steps down in phonetic salience and resulting in reduction of phonetic distance as the lower member(s) of the hierarchy are incorporated into an inventory. However, it is probably not necessary to trouble with refinements of the definition of salience as the implicit question here can be answered with only an intuitive approach. To the question: is maximization of distinctiveness the principle on which inventories are constructed? the answer is obviously 'no'. Clicks are highly salient yet few languages (1%) use them, and those that do have multiple series of clicks rather than exploit- them for maximum contrast with, say, plosives. The most frequent vowel inventory is /i e a o u/ not /i ẽ a̤ o̰ uˤ/ where each vowel not only differs in quality but is distinctively plain, nasalized, breathy, laryngealized or pharyngealized. Yet this second set of vowels surely provides for more salient distinctions between them and approaches maximization of contrast more than the first set whose differences are limited to only the primary dimensions conventionally recognised for vowel quality. The world's languages only add the additional parameters of contrast to vowels if they include a fairly wide sample of contrasts on the primary vowel quality dimensions. In other words, these ways of contrasting vowels are involved in an implicational heirarchy which is counter to the one predicted by a principle of maximal salience.

Apart from the above discussion, it must be recognized that phonetic distance cannot explain some of the prohibitions on co-occurrence of segments. There is a class of these that differ from those in (3) in that the distinctiveness of the segments involved is not really in doubt. An example of this type is the co-occurrence restriction which applies to subinventories of laterals. A language with several lateral segments contrasts them <u>either</u> by manner (voiced approximant, voiceless fricative, ejective affricate, etc) or by place (with all laterals being voiced approximants). Only one language in UPSID (Diegueño) clearly violates this rule, although Irish is an arguable exception too. Even two exceptions is significantly less than expected. Thus while all multiple-lateral subsystems must contain an apical or laminal lateral approximant, which is therefore at the top of an implicational hierarchy, at the lower end of this hierarchy there are two branches, one permitting elaboration by place, the other permitting elaboration of laterals sharing the same place of articulation by variation in the mannner of production.

The establishment of the fact that certain types of mutual exclusions occur which do not seem to be based on principles of phonetic distance is reminiscent of the position that there is a principle of 'compensation' controlling the structure of inventories. Martinet (1955), for example suggests that a historical change which simplifies an inventory in one area is counterbalanced by a compensating elaboration elsewhere. If diachronic changes do generally follow this pattern, then the consequence should be measureable relationships between various facets of inventories which follow a pattern of negative correlation. We have already seen, though, one aspect of inventory structure in which compensation does not occur. The tendency for vowel inventories to be larger with larger consonant inventories (section 4) is the opposite of the prediction made by a compensation theory. Several others inventory sectors were investigated for general signs of the operation of a compensation process.

The stop inventories of the languages in UPSID were examined to see if there was a tendency for elaboration of the number of place contrasts to be compensated for by reduction of the number of stop manner contrasts and vice-versa. Such a compensation is suggested by the inventories of Australian languages. These typically have a rich range of places of articulation for stops (and nasals) but no contrasts of manner within the stops (see, e.g., Wurm, 1972). Is this a local aberration or just a particularly striking example of a basic pattern in language?

There are a number of ways in which this comparison of places and manners could be done. In this instance, it was decided to treat doubly-articulated stops (in practise, this means labial-velars) as having a place of articulation distinct from either of their components. Secondary articulations - more likely to appear with a range of primary places of articulation -- seemed more akin to the 'series-generating' nature of differences in initiation and phonation type, and hence were treated as differences in manner. So, of the two inventories given in (4), (4a) is treated as having 4 places of articulation and 2 manners,

(4)	(a)	p	t	k	kp
		b	d	g	gb
	(b)	p	t	k	
		p^w		k^w	
		b	d	g	

whereas that in (4b) is treated as having 3 places of articulation and 3 manners.

The correlation was obtained between the number of places out of a list of ten ('glottal' place was not included in the calculation, as glottal stops do not have contrasting manners) and the number of manners out of a list of fourteen 'series-generating' manner components* used in each language. The numbers of languages involved are shown in Table **1**. Those rows with very sparse representation, i.e. less than 3 and more

		manners				
		1	2	3	4	
places	3	16	83	37	18	154
	4	15	27	32	14	88
	5	3	17	10	13	43
		34	127	79	45	

Table 1.

than 5 places, or more than 4 manners have been eliminated, removing 32 languages from the calculation. The observed counts are significantly different from the expected values under a hypothesis of independence

* Plain voiceless, plain voiced, voiceless aspirated, breathy, preaspirated, laryngealized, implosive, ejective, prenasalized, nasally-released, labialized, palatalized, velarized, pharyngealized.

(p = .0019). There is a very weak positive correlation (r = .15) between the numbers of places and manners, whereas the hypothesis of compensation would predict a strong negative correlation.

A similar computation was performed for fricatives, with cases with over 5 places or over 4 manners dropped (resulting in 38 languages being excluded). The results are given in Table 2. The observed data are again significantly different from expected (p = .0001), and in this

		manners			
		1	2	3	
places	1	36	10	1	47
	2	46	34	1	81
	3	24	46	6	76
	4	7	30	12	49
	5	4	19	3	26
		117	139	23	

Table 2.

case a more substantial positive correlation (r = .45) between the two variables is found. Again this is counter to the predictions of a compensation hypothesis.

6. Segments and Suprasegmentals

Despite the failure to find any confirmation of a compensation hypothesis in several tests involving segmental subinventories, it is possible that the compensation exists at another level. One possibility is evidently in the minds of Firchow and Firchow (1969). In their paper on Rotokas, which has an inventory of only 11 segments, they remark that "As the Rotokas segmental phonemes are simple, the suprasegmentals are complicated".* A similar view of a compensatory relationship between segmental and suprasegmental complexity seems implicit in much of the literature on the historical development of tone. For example, Hombert, Ohala and Ewan (1979) refer to "the development of contrastive tones on vowels because of the loss of a voicing distinction on obstruents". If this phenomenon is part of a pervasive relationship of compensation we would expect that, in general, languages with larger segmental inventories would have relatively simple suprasegmental characteristics and languages with small segmental inventories would tend to have more complex suprasegmental characteristics.

In order to test this prediction, the languages in UPSID which have less than 20 or more than 45 segments were examined to determine if the first group had obviously more complex patterns of stress and tone than the second. Both groups contain 28 languages. The findings on the

* Rotokas is not really very complex in its suprasegmentals. It has a partially predictable stress and a contrast of vowel length that seems only partly independent of stress (Firchow, Firchow and Akoitai, 1973).

suprasegmental properties of these languages, as far as they can be ascertained, are summarized in Table 3. Despite some considerable

		languages with small segment inventory (> 20)	languages with large segment inventory (< 45)
Stress	contrastive stress	6	8
	predictable stress	7	9
	pitch accent (?)	2	2
	no stress	5	4
	inadequate data	8	5
Tone	complex tone system	2	6
	simple tone system	2	4
	no tones	22	15
	inadequate data	2	5

Table 3.

uncertainty of interpretation and incompleteness of the data, the indications are quite clear that these suprasegmental properties are not more elaborate in the languages with simpler segmental inventories: if anything, they tend to be more elaborate in the languages with larger inventories. There are more 'large' languages with contrastive stress and with complex tone systems (more than 2 tones) than 'small' languages. There are more 'small' languages lacking stress and tone. The overall tendency appears once again to be more that complexity of different kinds goes hand in hand, rather than for complexity of one sort to be balanced by simplicity elsewhere.

7. Segment Inventories and Syllable Inventories

Another possibility is that the size of the segment inventory is related to the phonotactics of the language in such a way as to limit the total number of possible syllables that can be constructed from the segments and suprasegmental properties it has. Languages might then have approximately equal numbers of syllables although they differ substantially in number of segments. Rough maintenance of syllable inventory size is envisaged as the function of cyclic historical processes by, for example, Matisoff (1973). He outlines an imaginary language in which "the number of possible syllables is very large since there is a rich system of syllable-initial and final consonants". At a later stage of the same language, these initial and final consonantal systems are found to have been simplified but "the number of vowels has increased and lexically contrastive tones have arisen" maintaining contrasting syllabic possibilities. If tone or vowel contrasts are lost consonant clustering will increase at the syllable margins again.*

* Matisoff also suggests that the morphological typology of the language would undergo evolutionary changes concomitant with the phonological changes.

A brief investigation of the relationship between segmental inventory size and syllable inventory size was carried out by calculating the number of possible syllables in nine languages. The languages are Tsou, Quechua, Thai, Rotokas, Gã, Hawaiian, Vietnamese, Cantonese, Higi and Yoruba (the last three are not in UPSID but detailed data on phonotactics were available). These languages range from those with small segment inventories (Rotokas, Hawaiian) to those with relatively large inventories (Vietnamese, Higi, Quechua) and from those with relatively simple suprasegmental properties (Tsou, Hawaiian, Quechua) to those with complex suprasegmental phenomena (Yoruba, Thai, Cantonese, Vietnamese). In calculating the number of possible syllables, general co-occurrence restrictions affecting classes of consonants, vowels, tones, etc were taken into account, but the failure of a particular combination of elements to be attested if parallel combinations were permitted is taken only as evidence of an 'accidental gap' and it is counted as a possible syllable.

The calculations reveal very different numbers of possible syllables in these languages. The totals are given in Table 4. Even with the

Table 4.

Language	Total Possible Syllables
Hawaiian	162
Rotokas	350
Yoruba	582
Tsou	968
Gã	2,331
Cantonese	3,456
Quechua	4,068
Vietnamese	14,430
Thai	23,638

uncertainties that are involved in this kind of counting, the numbers differ markedly enough for the conclusion to be drawn that languages are not similar in terms of syllable inventory size.

As a follow-up to this calculation, several tests were done to see which of a number of predictors best correlated with syllable inventory size. The predictors used were the number of segments, the number of vowels, the number of consonants, the number of permitted syllable structures (CV, CVC, CCV, etc.), the number of suprasegmental contrasts (e.g. number of stress levels x number of tones), and a number representing a maximal count of segmental differences in which the number of vowels was multiplied by the number of suprasegmentals. Of these, the best predictor is the number of permitted syllable types ($r = .69$), an indication that the phonotactic possibilities of the language are the most important factor contributing to the number of syllables. The next best predictor is the number of suprasegmentals ($r = .59$), with the correlation with the various segmental counts all being lower. Although all the predictors tested show a positive simple correlation with the number of syllables, in a multiple regression analysis only the number of vowels contributes a worthwhile improvement to the analysis (r^2 change = .19) beyond the number of syllable types.

8. Conclusions

Work with UPSID has confirmed that segment inventories have a well-defined central tendency as far as size is concerned. Nonetheless considerable variation in their size and structure occurs. Their structure is subject to a heirarchical organisation in many particulars but cannot be substantially explained in terms of a single hierarchy of segment types. This is partly because segments of certain types are subject to rules of mutual exclusion. The mutual exclusions cannot all be explained as due to avoidance of inadequate phonetic contrasts, as some involve strongly salient distinctions.

A search for evidence that languages maintain a balance by compensation for complexity in one phonological respect by possessing simplicity elsewhere failed to find it in balance between classes of segments, between segments and suprasegmental contrasts, or between segments and phonotactic conditions. These investigations tended rather to suggest that complexity of various kinds occurs together in languages, and that languages really do differ in their phonological complexity.

References

Baucom, K.L. (1974) "Proto-Central Khoisan" Third Annual Conference on African Linguistics (ed. E.Voeltz) Indiana University Publications African Series 7, Indiana University, Bloomington: 3-38.

Crothers, J. (1978) "Typology and universals of vowel systems" Universals of Human Language 2: Phonology (edd. J.H. Greenberg et al.) Stanford University Press: 93-152.

Firchow, I & J. Firchow (1969) "An abbreviated phoneme inventory" Anthropological Linguistics 11: 271-276.

Firchow, I, J. Firchow & D. Akoitai (1973) Vocabulary of Rotokas - Pidgin - English Summer Institute of Linguistics, Papua New Guinea Branch, Ukarumpa.

Grace, G.W. (1959) The Position of the Polynesian Languages within the Austronesian (Malayo-Polynesian) Language Family IJAL Memoir 16, Indiana University, Bloomington.

Hockett, C.F. (1955) A Manual of Phonology IJAL Memoir 11, Indiana University, Bloomington.

Hombert, J.-M., J.J Ohala & W.G. Ewan (1979) " Phonetic explanations for the development of tones" Language 55: 37-58.

Ladefoged, P. (1978) "Phonetic differences within and between languages" UCLA Working Papers in Phonetics 41: 32-40.

Maddieson, I. (1980) "UPSID: the UCLA Phonological Segment Inventory Database" UCLA Working Papers in Phonetics 50: 4-56.

Martinet, A. (1955) Economie des Changements Phonetiques (2nd ed.) Francke, Berne.

Matisoff, J.M. (1973) "Tonogenesis in Southeast Asia" Consonant Types and Tone (ed. L.M. Hyman) Southern California Occasional Papers in Linguistics 1, University of Southern California, Los Angeles.

Pukui, M.K. & S.H. Elbert (1965) Hawaiian - English Dictionary (3rd ed.) University of Hawaii Press, Honolulu.

Sheldon, S.N. (1974) "Some morphophonemic and tone perturbation rules in Mura-Piraha" IJAL 40: 279-282.

Snyman, J.W. (1975) Zu|'hoasi Fonologie en Woordeboek Balkema, Cape Town.

Traill, A. (1978) "Research on the Non-Bantu African Languages" Language and Communication Studies in South Africa (edd. L.W. Lanham & K.P. Prinsloo) Oxford University Press, Cape Town.

Voorhoeve, C.L. (1965) The Flamingo Bay Dialect of the Asmat Language Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 46, Nijhoff, The Hague.

Wurm, S.A. (1972) Languages of Australia and Tasmania Janua Linguarum, Series Critica 1, Mouton, The Hague.

Abstract of M.A. Thesis

A study in Phonemic Universals, especially conerning fricatives and stops.

Jonas N.A. Nartey.

This thesis consists of five chapters. The first briefly describes the database used and its structure and contents. The second reports on the frequency of primary fricatives of different types in the languages surveyed and discusses how far greater frequency of occurrence can be explained by greater relative phonetic salience. A good rank order correlation holds between frequency and salience rankings of voiceless fricatives. Frequency of secondary (modified) fricatives is also reported in this chapter and a number of generalisations on the stucture of fricative inventories are stated.

The third chapter reports in a similar fashion on the frequency of primary (plain voiced and voiceless) stops and secondary stops, and analyses certain aspects of the structure of stop inventories. The frequency of affricates of various types and of the glottal stop is also reported. The fourth chapter is concerned with nasal consonants. It amplifies Ferguson's 1961 observations on nasals and supplements the discussion with an analysis of the structure of nasal inventories. The literature on salience of nasals is reviewed to show that some of the frequency patterns may be related to relative salience.

A common finding in chapters three, four and five is that the most frequent segments in smaller inventories also form the most frequent membership of larger inventories.

The final chapter briefly reviews some of the similarities and differences between the three classes of sounds (fricatives, stops and nasals) with respect to frequency of phonetic attributes such as place of articulation and phonation types. The findings are related to the theory that contrast between segments in an inventory is maximized.

This thesis appeared as UCLA Working Papers in Phonetics 46 in November 1979.

Insights on vowel spacing: results of a language survey

Sandra Ferrari Disner

I. Introduction

This paper presents the results of an analysis of the vowel systems of 317 languages in the UCLA Phonological Segment Inventory Database (UPSID). It shows that deviations from the patterns predicted by a theory which proposes that vowels are dispersed in the available phonetic space are relatively infrequent and, for the most part, confined to matters of small scale, falling into a few definable classes. It will be argued that even in these deviations from the predicted patterns there is evidence that vowels tend toward a balanced and wide dispersion in the available phonetic space.

II. Preliminaries

A few basic vowel inventories and a few basic configurations show up time and again in natural languages, while other no more complex patterns are rare or totally absent. The most prevalent patterns seem to be the so-called 'triangular' systems, particularly those of average size, and notably the five-vowel systems. For example, over a quarter of the 209 languages in the Stanford Phonology Archive have a triangular five-vowel system consisting of /i ɛ a ɔ u/, while less than 5% have any of the other five-vowel configurations; the 'square' four-vowel and six-vowel systems combined total less than 10%. (Crothers 1978)

Several attempts to explain these patterns invoke a principle of vowel dispersion, proposed in slightly differing versions by Liljencrants and Lindblom (1972), Lindblom (1975), Terbeek (1977), and Maddieson (1977).[1] This principle holds that vowels tend to be evenly distributed in the available phonetic space and also widely distributed, within the limitations of the particular system. The proposed models for vowel dispersion predict an optimal arrangement for any given number of vowels in the system; such theoretical systems may then be compared with the vowel systems of natural languages.

Just such a comparison is the starting point for the present paper. The vowel systems of 317 languages are examined for symmetry and dispersion. We take note of those systems in which the vowels are *not* evenly or widely

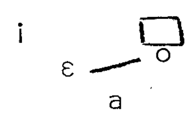
---

[1] The formulations differ somewhat in the degree of dispersion they propose, but no attempt will be made in this paper to choose between them. Except for the absence of vowels at the extreme corners of the vowel space, the data is unsuited for this task. For the most part, we can only look at *areas* within the phonetic vowel space and label the general arrangements according to which areas are filled. In order to investigate whether specific *points* in the space are filled, we would need acoustic measurements drawn from a large number of speakers of each language. (Cf. Disner 1978)

distributed in the available space, and seek to determine whether these vowel configurations can nevertheless be accounted for in a principled way. There may, for example, be straightforward historical or phonetic explanations for these "defective" vowel systems. If, however, there remain a substantial number of vowel systems which seem to obey no apparent rule, we should perhaps reconsider the notion of dispersion. It may be the case that vowel spacing is not at all a principled matter, and the success of, say the Liljencrants and Lindblom model in predicting the balance of the vowel systems may prove to be merely coincidental.[2]

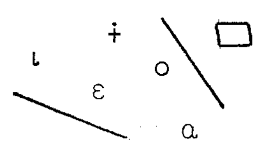
Note that the existence of at least some defective vowel systems in natural languages does not automatically rule out a dispersion theory. The claims made by a dispersion theory may be essentially correct, but languages could nonetheless undergo processes which produce defective vowel systems -- e.g. vowel mergers, shifts, etc. If this is the case then we should expect to see evidence of pressure to 'correct' the vowel spacing by compensatory shifts. This understanding of the interaction of vowel dispersion and other processes predicts that vowel systems, studied synchronically, should include systems which, although well-spaced, include compensation or rotation of the vowels. Some of the ways in which defective vowel systems could assume configurations that are basically consistent with the dispersion hypothesis as understood here are illustrated below.

One possibility is illustrated in Figure 1. This shows a system with a gap in the high back region, but in which the back mid vowel is higher in the phonetic space than the corresponding front mid vowel. The system appears 'skewed' and gives the appearance that one vowel has been drawn higher to compensate for the presence of a gap.



(1)

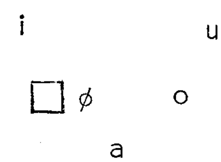
Figure 2 illustrates another possibility, in which the entire system is rotated with respect to the typical, unmarked configuration for a vowel system of a particular size (here, a five-vowel system), thereby achieving maximal dispersion with a slightly different orientation.



(2)

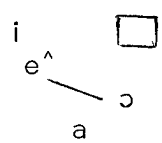
[2] For example, the apparent success of such a model might be attributable to the heavy emphasis on a few language areas in the sample utilized by Liljencrants and Lindblom, or else to a subtle bias in the sources in favor of reporting apparently balanced vowel systems when adequate phonetic detail is lacking.

Figure 3 illustrates a defective system that is complemented with a vowel of unexpected quality in the vicinity of the gap. This, too, serves to balance the system to a certain degree.



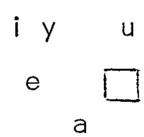
(3)

However, not all conceivable defective systems are open to interpretation as basically consistent with a theory of dispersion in the phonetic space. For example, there may be skewed systems in which the vowel adjacent to the gap is farther away from it than would be expected from comparison with the paired vowel. Compare Figure 4 with Figure 1.



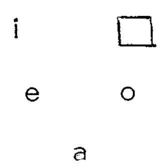
(4)

Or there may be systems in which a vowel of unexpected quality is located well away from the vicinity of the gap, increasing the imbalance even more. (Fig. 5)



(5)

Also, there may be systems which are not open to interpretation either as making compensation for or as conspicuously flaunting the imbalance in the system. These 'stationary' systems simply contain a gap.(Fig. 6)



(6)

These systems, like those in Figures 4 and 5, would be equally counter-examples to a theory of maximal dispersion. However, such cases as Figure 6 are ambiguous in that the vowels might well be phonetically underspecified representations of the vowels in Figures 1 or 3, or in some other such deviation from the most common vowel configurations. Whether or not the vowels of a particular language are represented in sufficient phonetic detail in the archive depends entirely on the phonetic judgments and transcription methods of the field linguist. Some linguists report the auditory quality of vowels in the narrowest detail, while others simply rely on the most common vowel transcriptions -- often those available on any typewriter -- to make all the necessary distinctions.[3a] Unfortunately, not even these two extremes can be reliably differentiated, for although a vowel system reported as /ieaou/ may be concealing a wealth of phonetic detail, it may just as likely be faithfully representing a perfectly balanced system, as in Figure 7.

```
        i         u

           e   o                    (7)

             a
```

In this case what appears to be a broad phonetic transcription is in fact quite narrow.

This typology of deviations establishes that the question of the essential correctness of vowel dispersion hypotheses can be addressed by a study of the frequency with which different types of apparent exceptions to it appear. The next section discusses how a study of this topic was carried out.

[3a] As mentioned above, considerable energy was expended at UCLA and Stanford to seek out whatever phonetic detail was available in the sources, regardless of the authors' transcriptions. Cross-references were checked, footnotes examined, and a good deal of reading between the lines was done in hopes of adding some detail to the reported phonetic quality.

Skewed vowel systems provide us with particularly valuable information on the phonetic quality of the vowels in question. A linguist who is little concerned with phonetic detail, or who does all his transcriptions on the typewriter, may well report pairs of vowels such as /e o/, /ɛ ɔ/, or even /e ɔ/ simply as /e o/. Even presumably balanced pairs reported as /ɛ ɔ/ or /e^o^/ may be concealing additional phonetic detail, or else emphasizing the vowel height more than is warranted from a purely phonetic standpoint. However, it is extremely unlikely that a pair of vowels reported as skewed (e.g. /e ɔ/) is actually balanced (/e o/). Here there is little doubt that the front vowel is higher than the back -- or, stated slightly differently, that the front vowel is closer to the /i/ region than the back vowel is to the /u/ region. We can therefore draw firmer conclusions about the degree of vowel dispersion in the phonetic space from these skewed systems. We should, on the other hand, refrain from drawing too-firm conclusions from balanced pairs of vowels.

III Method

A. Data

317 languages of the world were tested for vowel dispersion. The sample consists of the 192 languages in the Stanford Phonology Archive and an additional 125 selected at UCLA to augment the areal and genetic balance of the original Stanford sample. The data on each language is based on published accounts of the phonetic and phonemic structure of each language. The sources were carefully examined for any details which might possibly shed light on the true phonetic quality of the vowels under study, and the maximum available detail which could be represented in the coding scheme was retained.

The vowel phonemes from the various sources were all represented on a height scale which includes seven values (high, low-high, high mid, mid, low mid, high-low, low), on a backness scale with 3 values (front, central, back), and on a rounding scale with 2 values (rounded, unrounded). Additional dimensions pertaining to length, nasalization, phonation characteristics (laryngealization, breathiness),and other features (r-coloration, lip compression) were recorded in the archive, but for purposes of clarity and simplicity these were not utilized in the present study.

In discussing these vowels, reference is made to a grouping criterion which may be termed peripheral/interior. The "peripheral" vowels are the front unrounded, back rounded, and low vowels -- all of which lie along the margins of the available phonetic space. It should be noted, however, that the high central vowels, although they occupy one of the margins of the phonetic space, do not fall within the peripheral category; this more restrictive definition of peripherality is justified on phonological grounds, as the high central vowels tend not to pattern with the true peripheral vowels in natural languages, and they are also less common than other peripheral vowels. Thus, high central /ɨ/ and /ʉ/, along with the remaining phonetically centralized vowels, constitute the set of "interior" vowels.

B. <u>Identifying 'defective' vowel system</u>

A basic characteristic of all maximally dispersed vowel systems is that there are no unbalanced gaps in the primary (peripheral) vowel system. A language with a gap is defined as one which fails to utilize a particular region of the vowel space, while fitting one or more vowels into each of the remaining regions. We tested the 317 languages to identify those which contained gaps in the peripheral vowel system.

The test examined whether the five major regions along the periphery of the vowel space -- high front, high back, mid front, mid back, and low central -- were filled with at least one vowel. A high or mid region may, however, be left empty without being considered a gap so long as no other peripheral vowel in the system has a similar value on the height scale. This qualification ensures that balanced 3-vowel systems /i a u/ or /e a o/ will not be classed as defective. Put more formally, this test requires that any $\begin{bmatrix} \text{V} \\ \alpha\ \text{high} \\ \beta\ \text{back} \end{bmatrix}$ be matched by at least one $\begin{bmatrix} \text{V} \\ \alpha\ \text{high} \\ -\beta\ \text{back} \end{bmatrix}$ in both the high and mid regions of the vowel space; there must also be at least one $\begin{bmatrix} \text{V} \\ +\ \text{low} \end{bmatrix}$ in the system.

It should be emphasized that this is a very weak test of dispersion, designed simply to find out whether the *framework* of the vowel system -- that is, the major peripheral subdivisions of the vowel space -- fulfills the requirements of wide and even distribution in the available space. There are many other possible violations of the dispersion theory -- unevenly distributed interior vowels, multiple vowels in a single major subdivision of the vowel space, and the like -- which are not detected by this particular test procedure; future investigations will have to address these more subtle violations. For the present, however, our test will show whether or not the basic requirements of vowel dispersion are met in the languages of our sample.

The various formulations of the dispersion model differ in predicting a more or less wide spacing of the vowels in the available phonetic space. Although we will not be able to resolve the question of whether *maximal* or merely *adequate* dispersion is the correct formulation of this principle, our results may be suggestive. Vowel systems which lack one or more of the "point vowels" /i u a (or ɑ)/-- that is, those vowels with the most extreme values for height and backness--are not exploiting the vowel space to the maximum. Therefore, such systems are perhaps better explained by a theory of adequate, rather than maximal, dispersion. We should not, in any event, allow our test procedure to impose expectations of maximal dispersion by classifying such systems as defective. This calls for an exception:

> Vowel systems which lack all high vowels or all low vowels, but are otherwise balanced, should not be classed as defective systems.

This exception only affects a small number of languages. Only two of the 317 languages were found to lack low vowels. The Cheremis system is centered rather high in the vowel space, although it does count among its eight non-low vowels the low mid central vowel /ʌ/. Tagalog, on the other hand, has a three-vowel system /ɪʊə/ that is somewhat compressed: it descends no lower than the mid vowel /ə/, and its remaining two points fall somewhat short of maximally high /i/ and /u/.[3b] This compression suggests that the Tagalog vowel system is, indeed, only adequately dispersed in the available space. Three other languages, Squamish, Amuesha, and Alabama, lack

---

3b. The Stanford archive notes that /e/ and /o/ occur in a great many loanwords in the dialect of educated Manila speakers. These have been excluded from the UPSID inventory, as such a dialect is not representative of the language as a whole.

high vowels. All of these have a basic three-vowel system /eˆ oˆ a/[4] which is centered rather low in the vowel space, with two high mid vowels, rather than high or low-high vowelsFor these languages and for Cheremis, which are compressed along one edge of the vowel space only, acoustic measurements are needed to determine whether near-maximal, or only adequate, dispersion is in effect.

IV. Analysis of Defective Systems

Forty-five languages (over 14% of the sample) were found to have vowel systems with at least one major gap. These will be discussed below under various headings.

A. Four-vowel systems

The test procedure does not classify vowel systems in the same way as the dispersion models. One particular configuration of vowels which is classified as 'defective' by our criteria is in fact fully in accordance with both the Liljencrants and Lindblom model and Lindblom's later refinement of it. This is the four-vowel system

i u

ɛ (8)

a

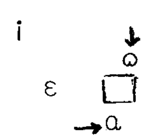
In this case the basic /i a u/ system has been expanded by a single vowel, and the one first chosen is the front vowel /ɛ/. (The corresponding back vowel /ɔ/ is not predicted to appear until the inventory has reached at least five vowels.)

By our procedure all 4-vowel systems with one mid vowel are regarded as defective. This classification seems justified since such systems are in any case rare -- only seven cases, or 2% -- and there is no single predominant pattern. Shasta, Paez, and Moxo have the inventory predicted by Liljencrants and Lindblom (i,ɛ,a,u). However, there are also three languages with four-vowel systems in which the single mid vowel is back and unrounded, contrary to predictions. The remaining language, Cayapa, differs in several respects from the first six.[5]

---

4 Squamish has an additional /ə/ near the center of the space.

5 Cayapa is a Paezan language which is closely related to Paez. They both have four-vowel systems, but the phonetic detail provided in the description of Cayapa suggests that it may have a well-balanced system. It has a high front vowel [i], a less high back vowel [ɷ] and low mid front vowel [ɛ] and a low back vowel [ɑ]. Rather than occupying their expected positions in the vowel space, the back vowels of Cayapa verge on the gap in the low mid region:

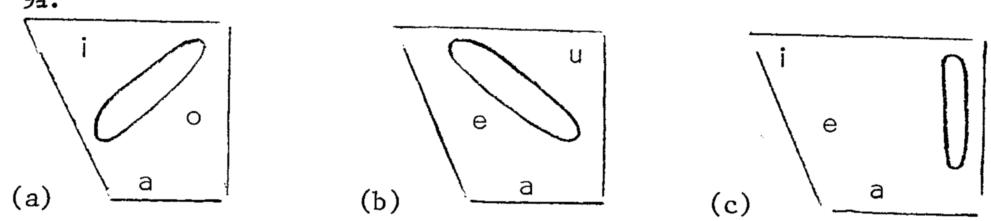


B. Frequency of missing vowels

Crothers, reporting on the Stanford archive languages, notes that the missing vowels are "generally /e/, /u/, or /o/, never /i/" (1978, p. 106). And though he formulates the near universal that "all languages have /i a u/" (p. 115), the counterexamples or "borderline cases" that he reports all have to do with deviations from an expected high back rounded vowel. This he notes, reflects the a > i > u hierarchy observed by both Greenberg (1966) and Jakobson (1941).

Our own frequency count of the missing vowels in UPSID confirms that the high front and the low central vowels are less likely to be missing than the high back rounded vowel (a, i > u). It further shows that the high back vowel is more likely to be absent in natural languages than either the front or the back mid vowels (e, o > u). Of the 37 languages which lack a single vowel, over half (19 languages) lack /u/; nine others lack /e/, seven lack /o/, and two lack /a/. The implied ranking is therefore $\left\{ \begin{matrix} i \\ a \end{matrix} \right\} > \left\{ \begin{matrix} e \\ o \end{matrix} \right\} > u$, rather than the /a>i>u>e>o/ which is generally assumed. This fact does not seem to have been commented on in the literature, and it may have implications that bear on such ideas as markedness and the choice between maximal and adequate dispersion.

In addition to the languages which lack a single vowel, seven lack more than one vowel. The most common pattern (5 languages) involves a missing high back vowel and mid front vowel, creating a gap of positive slope as in figure 9a.



(9)

None of the languages exhibits a gap of negative slope due to the lack of a high front vowel and a mid back vowel. (Figure 9b). Vertical gaps are also rare: only one language lacks front high and mid vowels, and another lacks back high and mid vowels. (Figure 9c). (Languages lacking both high vowels or both mid vowels, it will be recalled, are exempted from the 'defective' category.)

C. Structure of systems with missing vowels

The defective languages demonstrate that vowel systems occasionally *do* avoid certain regions of the space. These systems will be discussed below in terms of the typology discussed in section II above.

(i) Stationary systems

Leaving aside the matter of ambiguity of transcription, our investigation reveals that nine languages fall under the category of 'stationary' systems--that is, systems which do not appear to compensate for the gap in any way. All of these happen to be three- or four-vowel systems, although larger systems can be stationary just as well. In these systems all the vowels are peripheral vowels of the most common types, and the systems are otherwise balanced, with no evidence of skewing from the front to the back. The clearest examples of such stationary systems are Klamath, which has the vowel system /ieao/, and Bardi and Adzera, which have a high back vowel but lack a front mid vowel. Shasta, Moxo, and Paez lack a back mid vowel; there is no evidence for considering these systems as any less defective. Variations on these basic patterns are:

Campa: The Campa system is similar to Klamath, except that the mid vowels are reported as being "mid close", i.e. higher mid. These vowels may nevertheless be very similar, or even identical, to those of Klamath.

Hupa: The basic Hupa system is a not-fully-peripheral /e o a/. However, an /ɩ/ occurs in the language as a surface segment. Depending on the status accorded to this anomalous segment, the language may or may not be viewed as having a gap in the high back region of the vowel space.

Mura: According to the source, Mura has a three-vowel system /i a o/. This could be classed as doubly defective by our test, with /i/ implying a missing /u/ and /o/ implying a missing /e/. However, there is no evidence in the source to indicate whether the transcription reflects phonetic reality or orthographic convention. The system may well be a conventional /i a u/ system, or possibly even a rotated system such as /i˃ o æ/, both of which are maximally dispersed in the vowel system.

As we have discussed, stationary systems, such as Klamath or Bardi constitute a counterexample to the theory of maximal dispersion--if indeed they are stationary and give no indication of compensating for, or balancing out, the empty region of the vowel space. However, we cannot dismiss the possibility that certain important phonetic details of the system have been obscured by a broad phonetic transcription. Therefore, a theory of maximal dispersion cannot be disproven with this set of potential counterexamples. Definitive evidence against such a theory will have to be sought in the form of systems with unevenly spaced vowels, (systems that have been rotated into asymmetrical patterns), rather than in systems which appear not to have compensated for the gap at all.

(ii) Complementary vowels

Some of the languages in our inventory have defective systems that are complemented by a single vowel of unexpected phonetic quality (or of a single anomalous vowel) which shares some of the features of the missing vowel. Vowel systems of this sort can be classed into four major categories: those in which the complementary vowel is a central vowel (11 languages), a front rounded or back unrounded vowel (13 languages), an anomalous segment (3 languages), or a peripheral vowel, very similar to the missing vowel, which lacks a counterpart of equal height and opposite rounding elsewhere in the vowel system (6 languages). Stated more formally, if the missing vowel is $\begin{bmatrix}\alpha\text{high} & \beta\text{back}\\ \gamma\text{low} & \delta\text{round}\end{bmatrix}$, the complementary vowel is either $\begin{bmatrix}\alpha\text{high}\\ +\text{central}\end{bmatrix}$, $\begin{bmatrix}\alpha\text{high}\\ \beta\text{back}\\ -\beta\text{round}\end{bmatrix}$, $\begin{bmatrix}\alpha\text{high}\\ \beta\text{back}\\ +\text{anomalous}\end{bmatrix}$, or $\begin{bmatrix}[-\alpha\text{high}] \text{ or } [-\gamma\text{low}]\\ \beta\text{back}\end{bmatrix}$.

a. Vowels of unexpected backness ($\begin{bmatrix}\alpha\text{high}\\ +\text{central}\end{bmatrix}$)

Two principal patterns of this type are found: /ə/ for an expected missing mid or low vowel (in Acoma, Wu, and Tagalog), or /ɨ/for an expected missing high vowel (in Abipon). Variations of these basic patterns are:

(Acoma type)

Margi: An /ɛ/ does occur, but mostly in loanwords. /ə/ and /o/ are the only native mid vowels.

E. Armenian: There is no mid front monophthong in the language, but /ie/ or /je/ does appear as the only diphthong. The categorization of this language as "defective" is thus somewhat questionable.

Bashkir: The gap in the system is complemented in three alternate ways. In addition to /ə/ there is an /ø/, and there is also evidence of compensation in the peripheral system (see below).

Cheremis: /ə/ offsets a missing low vowel, but there is also evidence of compensation (see below).

(Abipon type)

Cofan: Very similar to Abipon, but with evidence of compensation as well (see below).

Guahibo: At first glance this system resembles both the Acoma and Abipon variants: a high vowel is missing, but the system is complemented by a mid vowel. The auditory properties of high /ɨ/ and mid /ə/ are actually quite similar, however, and it is not unlikely that the more common (less marked) symbol /ə/ might have been used in this case for a vowel that is closer in quality to /ɨ/.[6] The actual phonetic values of the vowels in the system are thus likely to be /i e a o ɨ (or $ɨ^v$)/, as in Abipon.

Chacobo: In addition to the missing high vowel a front mid vowel is also lacking. There is no central vowel to offset the latter. This double gap may perhaps be better explained as a rotation of the entire system.

(iii) Vowels of unexpected lip position $\begin{pmatrix}\begin{bmatrix}\alpha\text{high}\\ \beta\text{back}\\ -\beta\text{round}\end{bmatrix}\end{pmatrix}$

Systems of this sort fit into the "defective" category somewhat more marginally than the previous sort. On a formant chart with F_1 and F_2 as the axes, the front rounded vowels and the back unrounded vowels are more centralized than are their front unrounded and back rounded counterparts, but certainly less so than the true central vowels /ə/ and /ɨ/. Crothers nevertheless classes these centralized vowels together with the true central vowels, and does not regard any such "interior" vowels as fulfilling the requirements for maximal dispersion of the system. Liljencrants and Lindblom take no stand on the status of the front rounded and back unrounded vowels; their model is not designed to generate this particular set of vowels.[7]

Thirteen defective systems in our sample have a vowel with the same height and backness as the missing peripheral vowel, but with the opposite rounding. In eleven of these thirteen examples the vowel complementing the peripheral system is the *only* front rounded or back unrounded vowel in the language; only in Bashkir and Khalaj are the complementary vowels (/ø/ in each case) embedded in a series (/yøy̞/ for Bashkir; /yø/ for Khalaj). This fact suggests that the front rounded and back unrounded vowels do not show up casually in the vowel system. Moreover, in examples such as Gilyak, with /ø/ for a missing /e/, and Island Carib, with /ɤ/ for a missing /o/, the complementary role of the mid interior vowel is underscored by the lack of a high interior vowel such as /y/ or /ɯ/ in the system. It is true of most languages with front rounded or back unrounded vowels that the system "builds down" from /y/ to /ø/ to /œ/. or from /ɯ/ to /ɤ/ to /ʌ/, such that the lower vowel implies the presence of the higher. The isolated /ø/ and /ɤ/ of Gilyak and Island Carib are therefore quite unusual. Their isolation suggests that they do not represent an incipient nonperipheral system, but instead are likely to be closely associated with the gap in the mid peripheral region.

In eight languages of our sample all gaps are complemented by a single front rounded vowel or back unrounded vowel.

---

[6] The opposite case is unlikely. The more marked symbol /ɨ/ is usually reserved for high central vowels exclusively.

[7] It is not clear to what extent we can consider systems with such vowels to be defective. Certain configurations of peripheral and nonperipheral vowels may, in fact, be *more* dispersed in the vowel space than the corresponding peripheral systems.

/ɯ/ for missing high back rounded vowel: Japanese, Nunggubuyu, Alawa

Variants:

Jaquaru: The native vowel system is as in Nunggubuyu, but an additional /ɛ/ appears in a limited number of loanwords.

Ocaina: In addition to the complementation by /ɯ/, there is evidence of compensation in the peripheral vowel system. (see below)

/ɤ/ for missing mid back rounded vowel: Island Carib
/ø/ for missing mid front unrounded vowel: Gilyak
/ɯɤ/ for missing high and mid back rounded vowels: Nimboran

In two other languages, as we have noted, the complementary vowels form part of a front rounded system.

/ø/ for missing mid front unrounded vowel: Bashkir, Khalaj

And in yet another three languages only one of the gaps in the peripheral vowel system is complemented. All of these examples involve a double gap of positive slope (missing /eu/) rather than the other possibility, a gap of negative slope (missing /io/).

/ɯ/ for missing /u/; /e/ also missing: Adzera

Variants:

Nez Perce: Gap at /e/ offset by a peripheral vowel of unexpected height. (see below)

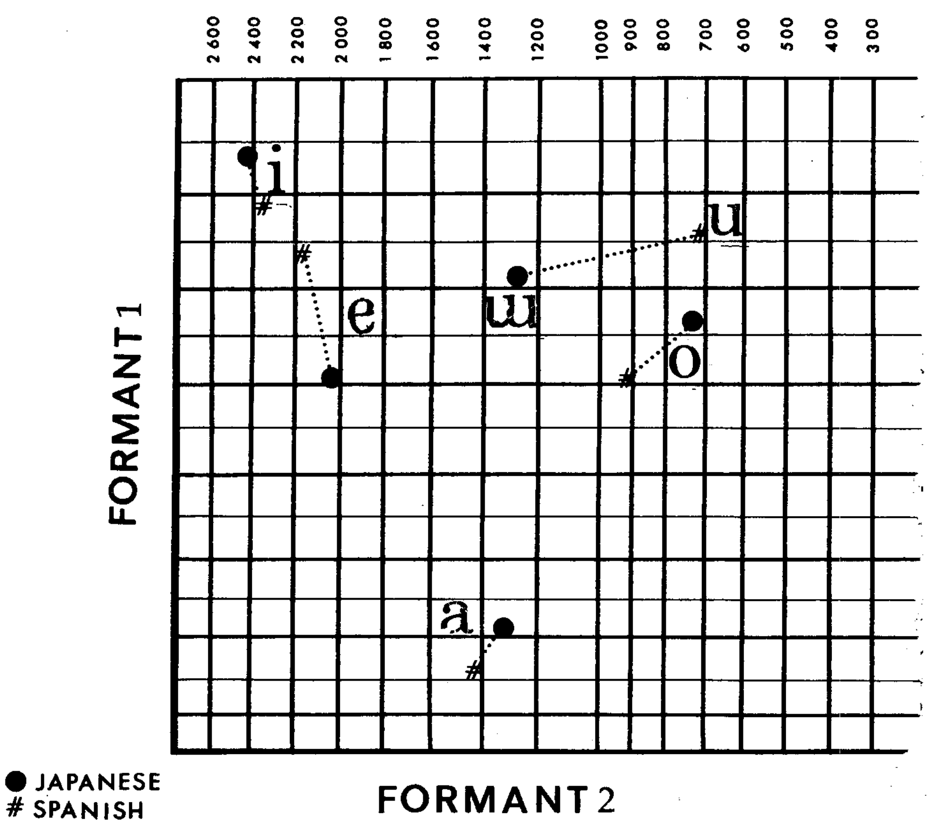
Amahuaca: Gap at /e/ offset by compensation in the peripheral system. (see below)

The complementary vowel in these double-gap systems is always the high back vowel /ɯ/ rather than the front mid vowel /ø/. Such preferential treatment of the high back region of the vowel space is quite unexpected in view of the fact that defective vowel systems are far more likely to be lacking a high back vowel than a front mid vowel. (We might well expect a greater number of /ɯ/s overall, due to the frequent absence of /u/; what is surprising is the complementation of /u/ instead of the complementation of /e/, *ceteris paribus*, in systems which lack both.)

Yet it is even somewhat surprising to note the predominance of /ɯ/ over other complementary interior vowels (/yøœɣʌ/). While this fact obviously relates the frequent absence of /u/, the latter might just as well go uncompensated, or else be compensated by some other vowel or by rotation of the system as a whole. Moreover, while we might expect the presence of /ɯ/

to presuppose the presence of /u/, just as, in general, /y/ presupposes /i/,[8] this is not true in the relatively large number of vowel systems in which /ɯ/ stands alone.

An investigation of the acoustic, perceptual, and auditory quality of the vowel /ɯ/ suggests that this vowel is in fact quite centralized, verging on the quality of central /ɨ/. For example, Ward's (1968) auditory analysis of Pamun /ɯ/ places this vowel well away from cardinal /u̲/, half-way between the central and back regions of the phonetic chart. Hombert (personal communication) has conducted a perceptual test on speakers of three Bamileke languages, Banjoun, Fe?Fe?, and Bangangte; synthetic vowels were presented to the subjects, and the formant frequencies which corresponded to "acceptable" vowels in their languages were noted. For the /ɯ/ in these languages the mean acceptable F_1 was 260 Hz., and F_2 was 1391 Hz.; this corresponds to a high and almost central vowel. Papçun (1976) presents an acoustic analysis of two five-vowel languages, Spanish and Japanese, the former having a rounded /u/ and the latter an unrounded /ɯ/; the differences between these high back vowels can be seen in Figure 10. The vowel /ɯ/ thus appears



(10)

(From Papçun 1976)

---

[8] As we have seen, no defective vowel system has a complementary /y/ in place of /i/.

to be most similar to the vowel /ɨ/. If we combine the nine vowel systems complemented by /ɯ/ with the four complemented by /ɨ/ into a single category, we find that these 13 cases of vowel system complementation constitute the majority of such cases in our sample; in contrast, complementation by the remaining interior vowels /y ø ɤ ə/ occurs in only eleven languages. Clearly, the high central region is a favored area for complementary vowels. The significance of this fact is not readily apparent, and merits further perceptual and acoustic study; it may be the case that the corner of the vowel space in which the high back vowel /u/ is to be found is simply smaller, in perceptual terms, than we might otherwise expect.

(iv.) Anomalous vowels ($\begin{bmatrix} \alpha\text{high} \\ \beta\text{back} \\ +\text{anomalous} \end{bmatrix}$)

These are special cases and should be considered individually.

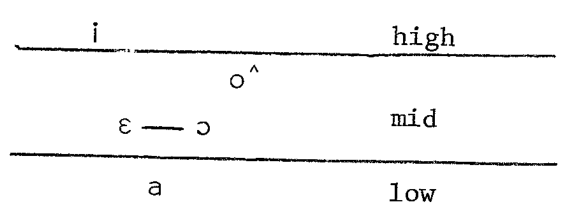
Seneca: Has a high back vowel /u/ with, however, extremely low frequency of occurrence.

Egyptian Arabic: Lacks /i/ and /e/ in its basic vowel system. A high mid /e^/ occurs as a surface segment only, while /i/ is posited as underlying but neutralized in surface forms.

Margi: Lacks a mid front vowel in its basic vowel system. There is, however, an /ɛ/ which is limited to a few loanwords.

(v.) <u>Vowels of unexpected height</u> ($\begin{bmatrix} [-\alpha\text{high}] \text{ or } [-\gamma\text{low}] \\ \beta\text{back} \end{bmatrix}$)

This category is reserved for those vowel systems with two peripheral vowels instead of one in a region vertically adjacent to the gap in the system. Here the complementary vowel is distinguished from other peripheral vowels by its unexpected quality. In the case of non-low vowels we shall define the complementary vowel as the one which lacks a counterpart of equal height and opposite rounding elsewhere in the vowel system. (e.g., the vowel /o^/ in a system such as



(11)

In the case of low vowels, the complementary vowel is always the non-central vowel -- i.e. the low vowel that is *not* /a/.

The precise location of the gap in the system may be inferred from the height of the matching vowel of opposite backness and rounding. Thus, a missing high back vowel matched by /ɩ/ will be considered a missing /ɷ/.

While many patterns of complementation are possible, only three occur in our sample.[9]

High back /u/ for a missing mid back vowel: Kunimaipa

Higher mid back /o^/ for a missing high back vowel: Nootka,[10] Navajo

High-low front /æ/ for a missing mid front vowel: Taishan, Nez Perce, Ket

In the languages of our sample a peripheral vowel of unexpected quality which is several levels of phonetic height away from the gap it complements never occurs. The additional vowels appear in the *immediately* adjacent vertical sector of the vowel space in all but one of these languages. Kunimaipa has an /ɷ/ complementing a gap at /o^/ (cf. the front vowel /e^/); Taishan, Nez Perce, and Ket all have /æ/ complementing a gap at /ɛ/ (cf. the back vowel /ɔ/ in each language); Nootka and Navajo have /o^/ complementing a gap at /ɷ/ or /u/ (cf. Nootka/i˅/ and Navajo /i/). Only in Navajo is the complementary vowel /o^/ on step further removed from the gap, specifically identified as a missing /u/ by our criteria. However, this exception may be simply a transcriptional artifact, since the lower-high vowels /ɩɷ/ are usually reported as high /i u/ in languages with a single pair of high vowels.

(a) Discussion

The vowel systems in which a single vowel of unexpected quality complements a gap in the peripheral system constitute the major portion (31 of 45, or about 69%) of the defective vowel systems in our sample. We have seen for example, systems lacking a high back vowel complemented by a high central vowel /ɨ/, a highmid back vowel /o^/, or a high back unrounded vowel /ɯ/, or else by an anomalous high back vowel; in only a single questionable case (Guahibo) is the complementary vowel much more distant from the gap. Similarly, systems lacking a mid front vowel were found to be complemented by /ə/ or /æ/ or /ø/, rather than by any more remote segment. The

---

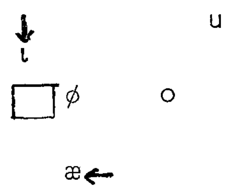
[9] Bashkir and Khalaj are not considered here since in both languages the two low vowels pertain to different vowel harmony sets. Both, however, follow the Taishan example, having a low front vowel for a missing mid front vowel. Hopi and Auca, which also have two low vowels instead of one, are discussed under "evenly spaced vowel systems", below.

[10] The question of whether to consider the Nootka vowel system a defective system at all will be discussed below.

common denominator among almost all thirty-one of the complementary vowels in our sample is that they bear a close phonetic resemblance to the expected vowel. Due to this fact, the systems containing such complementary vowels ought to be considered "not excessively deviant" from the predicted, fully-dispersed vowel system. We may even assume that they obey some sort of weak dispersion hypothesis -- certainly weaker than maximal dispersion, but not so weak as in the stationary systems described above.

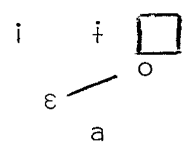
D. Complementary vowels with additional adjustments

We have seen that a few of the systems with complementary vowels seem to be striving for a greater degree of dispersion in the vowel space. Not only is there a complementary vowel near the site of the expected vowel, but the system is skewed such that one or more of the remaining vowels is found closer to this gap than we would otherwise expect, based on the height of the matching vowel(s). For example, the Gilyak system, lacking a mid front vowel, is complemented by the nonperipheral vowel /ø/; moreover, both the high front /ι/ and the low /æ/ are unexpectedly close to the gap, as evidenced by the absence of any other high front vowels or low vowels in the system. (Figure 12).



(12)

In the Cofan system, which lacks a high back vowel, the gap is flanked on two sides by a complementary /ɨ/ and by a higher than expected /o/ (as evidenced by the imbalance in the mid vowels, lowmid /ɛ/ vs. mid /o/).



(13)

Rather than occupying their expected locations, the vowels seem to position themselves around the gap.

Cofan: Lacks a high back rounded vowel, has complementary /ɨ/; also has /o/ in proximity.

Ocaina: Lacks a high back rounded vowel, has complementary /ɯ/ ; also has /o^/ in proximity.

Gilyak: Lacks a mid front unrounded vowel, has complementary /ø/; also has /ι/ and /æ/ in proximity.

Bashkir: Lacks a mid front unrounded vowel, has complementary /ø/ and /ə/; also has a raised /æ^/ in proximity.

(Amahuaca: Lacks a high back and a mid front vowel. An /ɯ/ complements the missing /u/ and the peripheral front vowels position themselves around the gap at /e/. However, as these are not *combined* effects such as those found in the preceding four examples, they will be discussed separately elsewhere.)

Again we must consider the possibility of underspecification of the phonetic vowel quality. Some of the authors undoubtedly have chosen a broader system of phonetic transcription than others, thus possibly obscuring some evidence of rotation of the vowel system. We therefore cannot rule out the possibility that some or even all of the cases of vowel system complementation in our sample are accompanied by some compensation in the peripheral vowel system.

E. Compensation in the peripheral vowel system (skewing and rotation)

Most of the defective vowel systems that we have examined thus far have fallen into two basic categories: ones in which there is either no evidence of any compensation for the gap in the system (stationary systems) or else there is a single vowel of unexpected quality in the system -- either a non-peripheral vowel, an anomalous vowel, or a peripheral vowel of unexpected height -- which may be regarded as complementary (complemented systems). We have, however, noted a few cases of complemented vowel systems whose remaining peripheral vowels are displaced toward the gap as well; these systems thereby achieve an even greater degree of compensation. A number of other languages in our sample lack a complementary vowel but nevertheless evidence of such displacement, even to the point of rotating the entire system.

(i) Compensations involving a single vowel

Vowel systems of this sort are very similar to vowel systems with a complementary peripheral vowel of unexpected height. In both cases there is a vowel in the system which is found closer to the gap than we would otherwise expect, based on the height of the closest corresponding vowel at the opposite end of the backness scale. (The crucial difference, of course, is that in the latter systems this vowel shares its region with another peripheral vowel, while in the former it is alone in its region of the vowel space.) In both Ocaina and Cofan, discussed in the preceding section, there is an /o/ which is higher than the corresponding front vowel /ɛ/, thus compensating for the missing high back vowel to some degree. Bashkir, which lacks a mid front vowel , also shows evidence of compensation. In these examples the vowel system is also complemented by vowels of unexpected lip position. Five other systems lack a complementary vowel , but otherwise follow the Ocaina pattern.

Malagasy: System lacks a high back vowel. /o^/ in proximity, as evidenced by the imbalance of /e/ and /o^/.

Mazatec: System lacks a high back vowel. /o^/ in proximity, as evidenced by the imbalance of /ɛ/ and /o^/.

Tacana: System lacks a high back vowel. /o^/ in proximity, as evidenced by the imbalance of /e/ and /o^/.

Cheremis: System lacks a low central vowel. /ɛ/ in proximity, as evidenced by the imbalance of /ɛ/ and /o/.

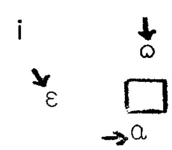
Amahuaca: System lacks a mid front vowel. /ɩ/ in proximity, as evidenced by the imbalance of high vowels /ɩ/ and /ɯ/.[11]

(ii.) Compensations involving multiple vowels

Vowel systems with displacements at more than one point in the peripheral system show an even greater tendency toward dispersion in the vowel space. In the extreme case, most or all of the vowels in the language show a displacement from the expected values (/i u/ for high vowels, /eo/ for mid vowels, /a/ for low vowels) in the direction of the gap; this may be considered a rotation of the entire system. Such a systematic displacement may well result in a maximally dispersed vowel system, which is oriented along a slightly different axis than that of most other vowel systems.

In Gilyak, discussed above, two of the peripheral vowels, /ɩ/ and /æ/, show evidence of displacement toward a gap in the front mid region of the system. Other cases of multiple displacement are rather uncommon.

Cayapa: The vowel system as a whole appears to be compensating for a gap in the mid back region. The high back vowel /ɷ/ is lower than its front counterpart /i/, and the low vowel /ɑ/ is further back than the expected /a/. The quality of the mid front vowel is reported as /ɛ/, and though there is no mid back vowel against which to test its height, we may perhaps infer (from the use of the symbol /ɛ/, rather than the unmarked /e/) that it is lower than expected, backing up the counterclockwise displacement of the low vowel. From what appears to be an anchor point at /i/, the vowels of Cayapa are displaced from their expected values according to the pattern in Figure 14.

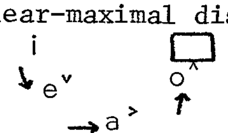


(14)

[11] Further evidence for the lowered high front vowel is afforded by the use of the marked symbol /ɩ/, rather than /i/.

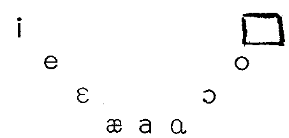
The resulting dispersion is very nearly maximal, in spite of the absence of one phonological category.

Tacana: As mentioned in the previous section, the /o^/ is located in the proximity of the missing high back vowel. There is also a suggestion in the sources that the vowel reported as /a˃/is located between the back and central regions, and that the vowel reported as /e/ is between mid and lowmid. This would have a vowel /ə˃/ and a vowel /eˇ/ backing up the counterclockwise displacement of the mid back vowel, in keeping with a near-maximal dispersion.



(15)

When a gap occurs in the peripheral vowel system, some or all of the remaining peripheral vowels are often found to be displaced in the direction of the gap, regardless of whether the gap is otherwise offset by a vowel of unexpected quality elsewhere in the system. This displacement may be interpreted as a means of establishing a wider and more even dispersion of the peripheral vowels, in spite of the gap. Ten of the languages in our sample -- 22% of the defective systems -- show evidence of some such compensation in the peripheral vowel system. However, only two of these languages, or possibly three, extend the compensation beyond a single vowel. This is largely due to the fact, noted above, that most of the defective vowel systems in our sample lack a high back vowel; in these systems only one peripheral vowel -- the mid back rounded vowel -- verges on the gap. Now, we would naturally expect that the vowel(s) adjacent to the gap would show a greater shift in vowel quality than the vowels further away, which, if anything, need only make subtle adjustments to the adjustments which precede them in the space. For most gaps, then, we can expect to find evidence of a few displacements. However, in the case of systems lacking a high back vowel,



(16)

it is far less likely that the peripheral vowel on the far side of the gap -- the high front vowel -- would be affected from so great a distance across the upper boundary of the vowel space. We should therefore not be too quick to conclude that displacement of a single vowel provides an adequate degree of dispersion for most of these languages. Moreover, a narrower phonetic transcription may well reveal subtle displacement throughout the vowel system which point to a maximal or near-maximal dispersion of the vowels, in spite of the gap in the phonemic system.

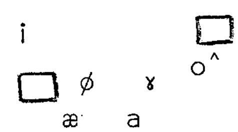
F. Unevenly spaced vowel systems

The languages discussed thus far do not exhaust the list of defective vowel systems in our sample. Vowels in the remaining systems appear to be distributed unevenly in the phonetic space, contrary to the predictions of any dispersion theory. As we have stated, vowel systems of this sort are the only ones which can validly be used to disprove a dispersion theory, for stationary systems--which also run counter to the predictions of such a theory--are indistinguishable from dispersed systems that merely lack adequately reported phonetic detail.

Two areally and genetically diverse languages, Hopi and Auca, constitute the clearest examples of unevenly spaced systems in our sample. They share a number of characteristics, notably a gap in the high back region and an apparent displacement of the remaining vowels *away* from this gap. A similar phoenomenon is found in Nootka, but the phonetic validity of these data is subject to question.

> *Hopi:* "So far as vowel placements are concerned, Hopi is extraordinarily asymmetrical." (Voegelin 1956:24)

The Stanford inventory of Hopi phonemes is based on analyses of three different Hopi dialects; the hybrid system is given as in Figure 17



(17)

(Note the unexpected height of the mid back vowel: even though there is no other peripheral mid vowel to test it against, the fact that this /o^/ is reported as higher than the nonperipheral mid vowels /ɤ/ and /ø/ suggests that this vowel *is* unexpectedly high.)

> While this sole peripheral back vowel does seem to be displaced toward the gap at /u/, the /a/ does not follow along. Similarly, /æ/ -- a front vowel of unexpected height adjacent to the gap in the mid front region -- verges close upon the low central vowel /a/, with no accommodation on the part of the latter. There are, however, indications that the peripheral system is actually more dispersed than the Stanford analysis would lead us to believe. The vowel /a/ of the Mesa dialect of Hopi is, according to Whorf, "as in 'calm'." This suggests a back, or at least a back-central[12] quality, which is as expected if an adequate separation is to be maintained between the two low vowels, and, simultaneously, if the separation between the backmost low vowel and the mid back vowel is not to become too great.

[12] Voegelin nevertheless classifies this vowel as central.

An alternative analysis would view all of these displacements together as consequences of a compensation for the missing high back vowel -- in other words, a near-complete rotation of the peripheral system. According to such an analysis, the peripheral vowels of the defective Hopi system are systematically displaced from the expected /i e a o/ (as in Mazatec, Hupa, Tacana, etc.) to the existing /i æ ɑ o^/, as negative pressure from the gap draws the vowels successively rightward (toward the gap) in a counterclockwise pattern.

```
i                   i                i                 i
                       o^                   o^               o^
  e   o            e   ↑            e                  ↘                (18)
    a                a                   →ɑ              æ     ɑ
```

It would be interesting to see whether there is any diachronic evidence for this systematic vowel displacement in Hopi, and if so, to see which displacements, if any, took precedence over others. We might look to the relative chronology of the /e/ → /æ/ and /a/ → /ɑ/ displacements for an indication of whether the vowel system rotated as a whole (/a/ → /ɑ/ first), or whether the central vowel was displaced in order to maintain an adequate separation in the low region of the vowel space (/e/ → /æ/ first). We might also look for indications that the displacements nearer the gap came before those farther away.

If the displacements described above are indeed accurate, the net result is that the Hopi system achieves a considerable degree of dispersion with a nearly square system and two interior mid vowels. Such an "hourglass" pattern may in a sense be more fully dispersed than the system with six peripheral vowels predicted by the dispersion models.

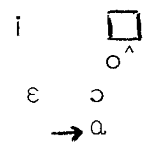
```
i
   ø  ɤ  o^                (19)
æ        ɑ
```

Auca: The Auca system is described as /i e æ a o^/. The /e/ is said to range from [ɩ] to [ɛ], and the /o^/ from [o] to [u] and inward to [ɯ]. No further phonetic detail is provided for /i/ /æ/ or /a/. The unexpectedly high midvowel (or low-high vowel, depending on the analysis) is a common compensation for a missing back vowel, but the juxtaposition of vowels /i e æ a/, well away from the gap, is surprising and unique among the systems in our sample. Certainly, phonetic verification of these vowel qualities is in order.

Nootka (Tseshaht): Whether or not this system is actually defective is subject to some dispute. The transcription provided by Sapir and Swadesh suggests that Nootka has a gap in the high back region of the vowel space[13], and a low vowel /ɑ/ that is displaced *toward*, rather than away from, the unexpected vowel /o^/.

---

[13] This is based on the Sapir and Swadesh transcription and on the statement that /o^/ and /o^:/ "have the tongue position of the vowel of 'coat' and the lip position of the vowel of 'hoot'" (Sapir and Swadesh 1939:13).



(20)

However, a closer examination of the source reveals that the phonetic quality of the vowel /o^/ is as in 'put', and long /o^:/ is as in 'food', comparable to the height of /i/ as in 'pit' and /i:/ as in 'feed'. On the basis of these phonetic descriptions, it is unlikely that the Nootka system is defective at all.

G. Conclusions

The great majority of vowel systems in our sample assume configurations which are predictable from a theory of vowel dispersion, considered in the light of some basic facts about the overall number of vowels, their degree of peripherality, and the like. At first glance there appear to be forty-five languages which are exceptions to the notion of vowel dispersion, in that one or more of the five major regions of the vowel space remain unfilled. However, some measure of dispersion is in force even in many of these "defective" systems. As a result, these systems maintain a degree of balance in spite of the obvious gap.

The defective systems may be classed into three major categories. Nine of these forty-five languages simply tolerate the gap in the system, showing no evidence of any compensatory shifts (although such shifts may in fact be present and merely overlooked in the broad phonological accounts). Two or perhaps three of the languages with major gaps show a displacement of the remaining vowels away from the gap, resulting in an even more uneven distribution of the vowels. The great majority, thirty-three languages, tend toward a balanced distribution of vowels in the available space, either by complementation with a vowel of unexpected quality or by a displacement toward the gap of some or all vowels in the system.

Thus the number of obvious exceptions to a vowel dispersion hypothesis in the UPSID data is extremely small. About 86% of the languages have vowel systems that are built on a basic framework of evenly dispersed peripheral vowels. At least 10% more approach this specification. This strongly indicates that a vowel dispersion theory correctly captures a principle governing the distribution of vowels in natural languages.

References

Crothers, J. (1978) 'Typology and universals of vowel systems' in Greenberg (ed.) *Universals of Human Language*, Vol. 2, pp. 93-152.

Disner, S. (1980) 'Evaluation of vowel normalization procedures.' *J. Acoust. Soc. Am.* 67(1), 253-261.

Greenberg, J. (1966) *Language Universals, with Special Reference to Feature Hierarchies*. The Hague: Mouton.

Jakobson, R. (1941) 'Kindersprache, Aphasie, und allgemeine Lautgesetze. Reprinted in *Selected Writings I,* 328-401. The Hague: Mouton, 1962.

Liljencrants, J. and Lindblom, B. (1972) 'Numerical similation of vowel quality systems: the role of perceptual contrast.' *Language* 48: 839-862.

Lindblom, B. (1975) *Experiments in Sound Structure*. Paper read at the Eighth International Congress of Phonetic Sciences, Leeds.

Maddieson, I. (1977) 'Tone loans: a question concerning tone spacing and a method of answering it.' *UCLA Working Papers in Phonetics*. 36: 49-83.

Papçun, G. (1976) 'How may vowel systems differ?' *UCLA Working Papers in Phonetics*. 31: 38-46.

Sapir, E. and Swadesh, M. (1939) *Nootka Texts*. Wm. Dwight Whitney Linguistic Series. Philadelphia: Linguistic Society of America.

Terbeek, D. (1977) 'Some constraints on the principle of maximum perceptual contrast between vowels.' *Chicago Linguistic Society* vol. 13.

Voegelin, C.F. (1956) 'Phonemicizing for dialect study, with reference to Hopi.' *Language* 32:116-135.

A Survey of Liquids

Ian Maddieson

1.0 Introduction

For reasons to do with both acoustic similarities and common phonological patterning, laterals and r-sounds have been grouped together as "liquids" in phonetic tradition. Although these similarities largely involve voiced non-fricative segments (Goschel 1972, Bhat 1974), the term liquid in this paper will be applied to all lateral segments except lateral clicks and all sounds that are included in the somewhat heterogeneous class of r-sounds. The core membership of this latter class consists of apical and uvular trills, taps and flaps. * Added to this core are a variety of fricative and approximant sounds which seem acoustically or articulatorily similar, or which are related by diachronic processes (Lindau-Webb, 1980).

2.0 Overall Frequency of Liquids

Using the definition above, almost all languages in a sample of 321 in UPSID have at least one liquid--95%. Most languages (72%) have more than one liquid. Details of the distribution are given in Table 1.

No. of liquids	No. of languages	Percent of languages
0	15	5%
1	75	23%
2	133	41%
3	47	15%
4	29	9%
5	12	4%
6	8	2%
7	2	1%

Table 1.

* Provision to distinguish between taps and flaps was made in the variables employed in UPSID, but the sources used do not seem to distinguish them reliably. They have therefore been treated here as a single group. For more discussion of the use of these terms see Ladefoged (1971) and Elugbe (1978).

The patterns found for systems of each size will be analyzed in Section 5 below, following an examination of the occurrence of particular types of liquids. As far as the two major classes of liquids are concerned, some 79% of languages have one or more lateral segments, whereas 76% have one or more r-sounds. The total number of laterals occurring in the surveyed languages is much greater than the difference between these percentages might suggest, since there are more languages with greater numbers of laterals. In fact, about 58% of the liquids reported are laterals.

3.0 Laterals

a) Types of laterals

The laterals occurring may be grouped under four broad headings: lateral approximants, taps/flaps, fricatives and affricates. The occurrence of these types is summarized in Tables 2-5 where frequencies are expressed in terms of the percentage of the total number of laterals counted in the survey (414). Approximant lateral types are shown in Table 2. Plain voiced approximant laterals are by far the most common type of lateral. Other types of approximant laterals are rare and only occur in inventories in which a plain type appears.

Approximant laterals	n	percent
plain voiced	313	76%
plain voiceless	8	2%
laryngealized voiced	8	2%
murmured voiced	1	--
	330	80%

Table 2.

There may be some doubt as to whether linguists have consistently reported on the distinction between voiceless approximants and voiceless fricative laterals, but the distinction is an important one to attempt to maintain. Unlike voiceless approximants, voiceless lateral fricatives are reported in inventories that contain no voiced lateral approximant (Tlingit 701, Nootka 730, Puget Sound Salish 734, Chukchi 908, Kabardian 911) so there may be an important distributional difference between the two types of sounds.

Lateral taps and flaps are reported fairly rarely (see Table 3). There is reason to believe that segments of this type may be more frequent than the reports indicate. Few phonetic manuals mention their occurrence and there is some evidence to suggest that field linguists may have difficulty recognizing them. In such cases lateral taps/flaps are likely to be reported as r-sounds (this is perhaps so in Tiwa 740) or as approximant laterals (as perhaps has happened in Zoque 711).

Lateral taps/flaps	n	percent
plain voiced	4	1%
laryngealized voiced	1	--
	5	1%

Table 3.

The few flaps reported are all voiced. Lateral fricatives, on the other hand, are far more likely to be voiceless than voiced, as Table 4 shows. Two languages in the survey have a voiced lateral fricative without a corresponding voiceless fricative. In Kanakuru (270) /lʒ/ only occurs "in a few words" and there is an approximant /l/, but in Pashto (14) a 'prepalatal' fricative lateral is the only lateral.

Lateral fricatives	n	percent
plain voiceless	36	9%
plain voiced	8	2%
ejective voiceless	3	1%
	47	11%

Table 4.

These examples show that it would be unsound to propose that voiced fricative laterals only occur with voiceless fricative laterals. However, the comparative rarity of voiced fricative laterals mirrors the lower frequency of voiced fricatives generally (Nartey 1979). The ejective lateral fricatives reported (in Tlingit 701, Yuchi 757 and Kabardian 911) are restricted to languages with glottalic consonants of other types and in each case a non-ejective voiceless lateral fricative also occurs. Ejective laterals are far more likely to be affricates than fricatives, but one language in the survey, Tlingit, has both /ɬ'/ and /t͡ɬ'/ in its inventory. This language also lacks a 'normal' voiced lateral approximant. Table 5 reports the types of lateral affricates in the survey.

Lateral affricates	n	percent
ejective voiceless	14	3%
plain voiceless	10	2%
aspirated voiceless	3	1%
plain voiced	2	--
	29	7%

Table 5.

A lateral affricate is almost always voiceless. It has a very high probability of being ejective and, in fact, is the only segment type so far reported with such a high probability of being ejective. The ejective lateral affricates only occur in languages which have a non-ejective lateral fricative or affricative (or both).

b) Places of articulation for laterals

Laterals are almost all articulated with the tip or blade of the tongue but it is not possible to be very much more specific than that. Due to inadequacy of data, no attempt was made in UPSID to distinguish between apical and laminal articulations, and, in a very large number of instances it is not possible to determine if a segment is dental or alveolar in place of articulation. A very large number of laterals is thus classified as "unspecified dental or alveolar." The data on place of articulation if given in Table 6 (percentages add to more than 100 because of rounding). About 85% of all laterals are produced in the dental-alveolar region. Probably alveolar laterals are more frequent than dentals, but this cannot be determined from the available data. Retroflex laterals are the next most frequent, and, of course, these too are tip or blade articulations. Laterals made with the body of the tongue are comparatively unusual. Among them palatals are most frequent. Velar laterals are extremely rare only one example appears in the survey, in Yagana (609). The three doubly-articulated segments reported to have velar and dental/alveolar articulations are all somewhat obscurely described. All three are voiceless and fricative or affricate, being interpreted as /x͡ɬ/, /kɬ/ (Ashuslay 814) and /kɬ'/ (Zulu 126). Apart from this rather dubious instance, there are no significant interactions between lateral manners and places of articulation. All types of laterals are predominantly dental or alveolar.

The preference for tip or blade articulations for laterals is presumably related to the greater opportunity to provide a free air passage behind the front closure if the body of the tongue is not involved in the articulation. Tongue-body laterals probably are more subject to processes resulting in their diachronic loss, as for example in the phenomenon known as 'yeismo' in Spanish which has transformed the palatal lateral /ʎ/ into a palatal approximant /j/ in many dialects (Guitarte 1971).

	Dental	Unspecified dental/alveolar	Alveolar	Retroflex	Palato-aveolar	Palatal	Velar	?Alveolar/velar
Secondary articulation								
none	29	222	88	27	7	14	1	3
palatalized	2	6	1	--	--	--	--	--
velarized	2	7	1	--	--	--	--	--
pharyngealized	--	1	--	--	--	--	--	--
TOTAL	33	236	90	27	7	14	1	3
% of total	8%	57%	22%	7%	2%	3%	--	1%

Table 6.

c) Summary statements on lateral segments

The observations above on lateral segments suggest that the following seven substantive generalizations can be made. The number of conforming cases in the inventory is shown over the number of potentially relevant cases after each statement, together with a calculation of the probability of a case conforming.

(i) A lateral segment is most likely to be articulated with tongue tip or blade (386/414) p.= .93

(ii) A lateral segment is most likely to be voiced (337/414) p. = .81

(iii) A lateral segment is most likely to be an approximant (330/414) p.=.80

(iv) A fricative lateral is most likely to be voiceless (39/47) p.=.83

(v) A voiceless lateral is most likely to be fricative (36/74) p.=.49

(vi) An ejective lateral is most likely to be an affricate (14/17) p. = .82

(vii) A lateral affricate is highly likely to be ejective (14/29) p. = .48

It should be noted that (iv) and (v) are independent observations. The explanation for (iv) is likely to be related to the greater salience of voiceless fricatives in general, compared to their voiced counterparts (Goldstein 1978). The explanation for (v) is likely to be related to the greater salience of voiceless fricatives over voiceless approximants, added to the fact that the only places favored for affricates in general are palatal and palato-aveolar, which are disfavored for laterals. There may however also be a special reason why diachronic processes act to retain few voiced lateral fricatives or voiceless lateral approximants, that is that these sounds are difficult to distinguish from non-lateral counterparts (e.g. [ɮ] → [ʒ]; [l̥] → [h]).

4.0 R-Sounds

a) Types of r-sounds

The sources used to compile UPSID fail to specify the manner of articulation of segments represented by /r/ in thirty-three instances, or 11% of the total of 297 r-sounds. These will be dropped from consideration in the analysis of different types of r-sounds below. It should be remembered that ignorance of how this group of sounds should be distributed into the various classes adds a measure of uncertainty to the conclusions reached in this section.

Of the remaining 268 cases the largest number are reported as trills. * The data is reported in Table 8.

Trills	number	percent
plain voiced	122	46%
plain voiceless	2	1%
laryngealized voiced	1	--
	125	47%

Table 8.

Obviously, trills are overwhelmingly voiced. The same is true of the next most frequent type of r-sound, consisting of taps and flaps. All reported taps/flaps are voiced, with only three being other than plain. The numbers are given in Table 9. Although fewer taps/flaps are reported than trills, the difference is less than the number of r-sounds with unspecified manner and hence no very firm conclusion can be drawn. In any case, trills and taps/flaps are closely related sound types (often both appear as allophones of the same phoneme) and it may be observed that 88% of thosem r-sounds with specified manner are interrupted. **

Taps/flaps	number	percent
plain voiced	109	41%
laryngealized voiced	2	1%
fricative voiced	1	--
	112	42%

Table 9.

*Ladefoged, Cochran and Disner (1977) claim that "very few languages have any trills at all." The data collected for UPSID suggest either that trills are not in fact particularly rare or that very many erroneous reports of trills occur in the literature.

** Including voiceless and voiced "trilled retroflex affricates" reported in Malagasy (410) but not included in either Table 8 or Table 9.

The numbers of continuant r-sounds are shown in Table 10. As with other types of r-sounds, voicing is obviously the norm.

	number	percent
voiced approximant	27	10%
voiced fricative	7	3%
voiceless fricative	1	--
	35	13%

Table 10.

Considerably more approximants are reported than fricatives. However there is considerable room for doubting the validity of the reporting of fricative r's. The criteria for reporting a sound as 'some kind of r' rather than, say, as a voiced retroflex sibilant /ʐ/ are obscure and may not reflect a phonetic difference between /ʐ/ and /ɹ̌/, but instead be based on phonotactic considerations or other non-phonetic characteristics (including orthographic convention).

As with laterals, there is a considerable number of instances where the place of articulation is only known to be in the dental/alveolar region somewhere. The tabulation for places of articulation in Table 11 thus includes an unspecified dental/alveolar column.

secondary articulations	dental	unspecified alveolar/dental	alveolar	retroflex	velar
none	9	133	114	39	4
palatalized	1	4	2	--	--
velarized	0	--	1	--	--
	10	137	117	39	4
	3%	46%	39%	13%	1%

Table 11.

Of the thirty-three r-sounds with an unspecified manner, twenty-nine fall into this underspecified place category. If all unspecified manner r-sounds are excluded then there is a larger number of alveolar r-sounds (115) than any other category. Note also that there are only a few r-sounds reported as dental - only about one-twelfth as many as are specified as alveolar. These two indications suggest that the most common place of articulations for r-sounds may be alveolar, although this cannot be taken as certain due to the very large number of cases with underspecified place.

The only other reasonably frequently occurring place for r-sounds is retroflex. Uvulars are quite rare (and are known to be mainly restricted to prestige dialects of Western European languages). Uvular trills are included in the survey from French (010), German (004) and Batak (413) and a uvular fricative from Eastern Armenian (022).

c) Interaction between place and manner

There is an important interaction between place and manner of articulation for r-sounds which can be seen from Table 12. This table juxtaposes the frequency by manners of alveolar and retroflex types of r's. (The unspecified dental/alveolar category has a distribution by manner similar to the alveolars.) Whereas an r-sound with an alveolar place

		alveolar n	alveolar %	retroflex n	retroflex %
<u>interrupted</u>	trills	54	49%	6	16%
	taps/flaps	47	42	12	32%
<u>continuant</u>	approximants	8	7%	14	38%
	fricatives	2	2%	5	14%
		111		37	

<u>Table 12</u>

is one of the interrupted types (trill/tap/flap) in the great majority of cases, retroflex r-sounds are most commonly found as approminants in the languages surveyed and, relative to alveolars, are very rare as trills but more common as fricatives. Within the approximant class, 14 or 27 instances are reported as retroflex.*

d) Summary statements on r-sounds

The analysis in the preceding section suggests that the following six substantive generalizations about r-sounds can be made:

(i) An r-sound is most likely to be voiced 293/297 p.=.99

(ii) An r-sound is most likely to be dental or alveolar 254/297 p.=.86

(iii) An r-sound is most likely to be interrupted 229/264 p.=.87

(iv) A retroflex r-sound is highly likely to be a continuant 19/37 p.=.51

(v) An approximant r-sound is highly likely to be retroflex 14/27 p.=.52

(vi) A fricative r-sound is highly likely to be retroflex 5/8 p.=.62

The explanation for (i) probably needs to be different for different types of r-sounds. Tap/flap durations are very short and connected

* There may be some reporting bias reflected in this finding. A somewhat retracted articulation of approximant /r/ is labeled 'retroflex' in some analyses of English (e.g. Kenyon 1926). This may have led to a predisposition to label any approximant 'r' retroflex.

speech is predominantly voiced. The voicing of taps/flaps may result from an inability to switch back and forth from voicing to a lack of voicing quickly enough. Approximants in general are predominantly voiced (Maddieson, in this volume) probably because voiceless approximants are poorly distinguishable and tend to fall together as the undifferentiated voiceless vowel /h/. However, there seems to be no equivalent reason for voiceless trills to be so rare. Trills generally have two or three contacts at a rate of vibration of about 28 Hz (Ladefoged, Cochran and Disner, 1977), requiring a substantial duration (on the order of 100 msec). They have some similarity in production to obstruents, which are preferentially voiceless (about 60%). There may, perhaps, be some factor in the aerodynamic conditions required for trilling which lead to favoring voicing because of the associated reduction in the rate of air flow.

5.0 Structure of Liquid Systems

Languages in the survey have up to six laterals and up to three r-sounds, although they most typically only have one of each. The number of languages with the various numbers of liquids are given in Table 13.

No. of Laterals	One	Two	Three	Four	Five	Six
No. of languages	163	60	17	8	3	6
Percent	51%	19%	5%	3%	1%	1%

No. of r-sounds	One	Two	Three
No. of languages	187	51	5
Percent	59%	16%	2%

Table 13

About 28% of the languages have two or more laterals, but only 18% have as many r-sounds. Of the 231 languages with two or more liquids, 75 have more laterals than r-sounds, whereas only 25 have more r-sounds than laterals. The remainder have equal numbers of the two major types of liquids.

a) Languages with one liquid (75, 23%)

In view of the greater overall frequency of laterals, it is rather surprising that languages having one liquid are more likely to have an r-sound (44) rather than a lateral (31), with the most frequent type of r-sound in these languages reported as a voiced flap (30 cases). However, in a number of cases, both lateral and non-lateral allophones occur,* e.g. in Nasioi (624) the flap /ɾ/ occurs as a lateral before /u/ or /o/, in Barasano (832) the alveolar flap /ɾ̠/ has flapped nasal and lateral allophones, with lateral flaps occurring before central and back vowels and [ɾ̠] before front vowels and in all word-final environments. Tucano (834), Apinaye (809), Japanese (071) and perhaps Bribri (801) are also among the languages which have some lateral allophones of their sole liquid and it is likely that other cases are concealed in less detailed descriptions. Such fluctuation between lateral and non-lateral liquids appears, not unexpectedly, to be most frequent in languages with only one liquid.

The most frequent lateral reported as sole liquid is a voiced dental or alveolar lateral approximant (29 of 31 cases). The only exceptions are the retroflex lateral flap of Papago (736) and the velar lateral approximant of Yagaria (609). The only approximants reported as sole liquids are laterals.

b) Languages with two liquids (133, 41%)

The most typical language has two liquids, usually one lateral and r-sound. Table 14 shows how two-liquid systems break down.

---

* The decision to classify these segments as one or the other depends on a judgment of the most widespread or most representative allophone.

	No. of languages	Percent of languages w/ 2 liquids	Percent of languages in sample
2 laterals, 1 r-sound	111	83%	35%
2 laterals	19	14%	6%
2 r-sounds	3	2%	1%

Table 14.

The usual system consists of a trill or tap/flap and a lateral approximant, but these are the most common varieties of liquids. The systems with two laterals or two *r*-sounds do not contrast them by place of articulation but by manner, voicing or secondary articulation or some combination of these. The most frequent system has a plain voiced lateral approximant and a voiceless lateral (13 of 22 cases). S. Nambiquara (816) has plain and laryngealized voiced retroflex lateral flaps, with non-lateral allophones reported for /ḷ/.

c) Languages with three liquids (47, 15%)

The structure of systems with three liquids is outlined in Table 15.

	No. of languages	Percent of languages w/ 3 liquids	Percent of languages in sample
2 laterals, 1 r-sound	22	47%	7%
1 lateral, 2 r-sounds	19	40%	6%
3 laterals	6	13%	2%

Table 15.

A system with two laterals is a little more common than one with two non-laterals. The two-lateral systems are about evenly divided between those with a contrast of place (dental/alveolar vs. palatal or retroflex) and those with a contrast of voicing, manner or secondary articulation between their laterals.. No cases are reported in which laterals contrast in both place and another feature. Languages with two r-sounds are more likely to contrast them in manner (13 out of 19) than on any other dimension e.g. trill vs. tap/flap. In nine cases there is a difference in the primary place of articulation (usually alveolar vs. retroflex) and in five of these this contrast is in addition to a manner difference. The systems with three laterals usually contain a contrast of voicing (5 of 6) with one or two voiceless fricatives or affricates (including ejectives). These systems are only reported from Amerindian languages, predominantly from mm the North-Western coastal region of the North American continent (5 of 6). Note that the 3-liquid systems in the survey do not include any with three r-sounds.

d) Systems with four liquids (29, 9%)

The structure of systems with four liquids is shown in Table 16. A clear majority of these systems (18 of 29) consist of an equal number of lateral and non-lateral liquids.

	No. of languages	Percent of languages w/ 4 liquids	Percent of languages in sample
2 laterals, 2 r-sounds	18	62%	6%
3 laterals, 1 r-sound	7	24%	2%
1 lateral, 3 r-sounds	2	7%	1%
4 laterals	2	7%	1%

Table 16.

Of these, eight languages have laterals that contrast in place of articulation alone; the remainder contrast laterals by manner, voicing or secondary articulations. A similar number, seven, contrast their two r-sounds by place of articulation but in four of these cases there is also a contrast of manner. In total, ten of the eighteen languages contrast their r-sounds by manner.

Of those languages with three laterals, two contrast plain voiced lateral approximants at three different places of articulation. The remaining five contrast their laterals by voicing and/or manner contrasts and include at least two fricative or affricate laterals. The two languages with three r-sounds are both unusual. E. Armenian (022) is reported with an alveolar trill, a retroflex fricative and a uvular fricative beside a dental lateral, i.e. / r̠, ʁ̌, ɹ̣̌, l̪/. Malagasy (410) has /r,ɫ/ plus voiced and voiceless "trilled retroflex affricates." The four-level systems contrast their laterals by manner and voicing, not by place. Note again, no systems occur with all their liquids non-laterals, and, in fact, no language is reported with more than three r-sounds.

e) Systems with five liquids (12, 4%)

The structure of systems with five liquids is shown in Table 17. The most typical of these systems consists

	No. of languages	Percent of languages w/ 5 liquids	Percent of languages in sample
3 laterals, 2 r-sounds	7	58%	2%
2 laterals, 3 r-sounds	1	8%	--
4 laterals, 1 r-sound	2	17%	1%
5 laterals	2	17%	1%

Table 17.

of three laterals contrasted by place and two r-sounds differing in manner. This kind of system is principally represented by Australian languages (5 of 6 cases); Alawa (354) with alveolar, palato-alveolar and retroflex voiced lateral approximants, an alveolar trill and a retroflex approximant, i.e. /l̠ , ʎ , ɭ , r , ɻ/ is a representative example. On the other hand, Ngizim (269) and the languages with four or five laterals in a five-liquid system contrast their laterals by voicing and manner contrasts.

f) Systems with six or seven liquids (10, 3%)

The structure of systems with six or seven liquids is shown in Table 18. Just as with systems with five liquids, the laterals tend either to differ by place (Diyari 367, Aranda 362, Irish 001) or by manner and voicing (Sedang 304, Chipewyan 703, Haida 700, Kwakw'ala 731). There is however one language, Diegueno (743), which includes intersecting contrasts of both place and voicing/manner, having the four laterals /l̪ , l̠ , ɬ̪ , ɬ̠/. The largest number of laterals reported in the

	No. of languages	Percent of languages w/ 6/7 liquids	Percent of languages in sample
4 laterals, 2 r-sounds	4	50%	1%
3 laterals, 3 r-sounds	1	13%	--
5 laterals, 1 r-sound	1	13%	--
6 laterals	2	25%	1%
4 laterals, 3 r-sounds	2	100%	1%

Table 18.

survey is six. The 7-liquid languages are two Australian languages (Kariera-Ngarluma 363, Arabana-Wanganura 366) which contrast laterals at four different places of articulation, trills at two places and also have a retroflex approximant.

g) Generalizations on the structure of liquid systems

The languages in the survey show up to four contrasts between places of articulation for laterals and up to six contrasts of manner and voicing for laterals. They show up to three contrasts of place for r-sounds and up to three contrasts of manner. These are likely to be the maxima for these contrasts. Although laterals are reported at six major places of articulation, no language is known to contrast palato-alveolar and palatal laterals, and velar laterals are so rare that for them to occur with laterals at three other places would simply be improbable.* (Ladefoged, Cochran and Disner, 1977, report dental, alveolar and velar lateral approximants co-occurring in Melpa.) There has been relatively little work done on phonetic differences between laterals (but see Bladon 1979), so it is not clear if the failure to exploit all the places of articulation in one language could be attributed to a lack of phonetic distinctiveness. As there are no languages reported with over six laterals or three r-sounds, these automatically set the maxima for the other contrasts mentioned above.

The most commonly found systems containing one to six liquids are reviewed in Table 19. These patterns suggest that

No. of liquids	Most common structure	n/total	Percent of cases
1	1 r-sound	44/75	59%
2	1 lateral, 1 r-sound	111/133	83%
3	2 laterals, 1 r-sound	22/47	47%
4	2 laterals, 2 r-sounds	18/29	62%
5	3 laterals, 2 r-sounds	7/12	58%
6	4 laterals, 2 r-sounds	4/8	50%

Table 19

*If the one instance of a velar lateral and the four instances of laterals at four different places of articulation in UPSID are taken as indications of the frequency of such occurrences, then the probability that both would occur in the same language might be estimated at less than .00004 (i.e. less than one language in 400,000).

an inventory of liquids is generally expanded by adding more laterals before adding more r-sounds. Several other generalizations concerning the structure of liquid systems also suggest themselves and are presented below.

(i) A language with two or more liquids is most likely to have at least one lateral (228/231) p.=.99.

(ii) A language with two or more liquids is most likely to include a lateral/non-lateral contrast between them (198/231) p.=.86.

(iii) A language with one or more laterals has a voiced lateral approximant (233/243) p.=.96. Although approximant laterals are the most common type of lateral, the probability of (iii) being true is significantly higher than their overall frequency (81%) would suggest.

(iv) A language with two or more laterals contrasts them either in place or in manner and voicing but not both (96/97) p.=.99.

(v) A language with two or more r-sounds is unlikely to restrict their contrast to place of articulation (38/52) p.=.73. These two observations, (iv) and (v), draw attention to a quite marked difference between laterals and r-sounds in the way that the systems are elaborated.

(vi) A liquid with both lateral and r-sound allophones is most likely to be the only liquid in the language (6/7). The data is incomplete on this point so no probability is expressed. However, only one counterexample is known among the languages in the survey. Finally, it may be repeated that:

(vii) A language most often has two liquids (one lateral and one r-sound) 111/321 p.=.35.

6.0 Conclusion

The survey of liquids in UPSID has revealed patterns of occurrences of different types of liquids which may be taken as reliable. These patterns concern both the overall frequency of particular sound types and their relation to the inventory in which they occur. Although such observations have an intrinsic interest of their own, their main value is to suggest avenues of investigation in diachronic phonology, articulatory phonetics or speech perception designed to seek the explanation for these patterns.

REFERENCES

Bhat, D.N.S. (1974) "The phonology of liquid consonants" Working Papers in Language Universals (Stanford University) 16: 73-104.

Bladon, R.A.W. (1979) "The production of laterals: some acoustic properties and their physiological implications" Current Issues in the Phonetic Sciences (Amsterdam Studies in the Theory and History of Linguistic Science, Series 4 - Current Issues in Linguistic Theory 9) Ed. Harry and Patricia Hollien; John Benjamins, Amsterdam: 501-508.

Elugbe, B.O. (1978) "On the wider application of the term 'tap'" Journal of Phonetics 6:133-139.

Goldstein, L. (1978) "Perceptual salience of stressed syllables" UCLA Working Papers in Phonetics 39:38-60.

Goschel, J. (1972) "Artikulation und Distribution der sogennanten Liquida r in den europaischen Sprachen" Indogermanische Forschungen 76: 84-126.

Guitarte, G.L. (1971) "Notas para la historia del yeismo" Sprache und Geschichte: Festschrift fur Harri Meier Zum 65 Geburtstag Fink, Munich: 179-198.

Kenyon, J.S. (1926) "Some notes on American r" American Speech 1: 329-339.

Ladefoged, P., A. Cochran & S. Disner (1977) "Laterals and trills" Journal of the I.P.A. 7: 46-54.

Ladefoged, P. (1971) Preliminaries to Linguistic Phonetics University of Chicago Press, Chicago.

Lindau-Webb, M. (1980, to appear) "The story of 'r'" UCLA Working Papers in Phonetics 51.

Nartey, J.N.A. (1979) "A Study in Phonemic Universals; Especially Concerning Fricatives and Stops" UCLA Working Papers in Phonetics 46.

Vocoid approximants in the world's languages

Ian Maddieson

1.0 Introduction

Apart from those approximants which have a lateral escape or belong to the family of 'r-sounds' the only frequently-occurring approximants in the world's languages are those which have vocoid characteristics (Pike 1943). This paper examines the frequency of such sounds as phonemic units in a sample of 317 languages (see the first article in this volume) and discusses certain co-occurrence restrictions which relate to their role in phoneme inventories.

In the UPSID inventories vocoid approximants have been coded as consonants when they a) do not alternate with syllabic vocoid pronunciations and b) share distributional properties with other consonants. Ninety percent of the surveyed languages have one or more such segments.

2.0 Frequency

The vast majority of languages have a voiced palatal approximant /j/ - some 86% of the 317 languages. A substantially smaller number of languages have a voiced labial-velar approximant /w/ - some 75%. Despite the lower frequency of /w/, there is no implicational relationship between these two segments; they occur together in the inventories of 71% of the languages, which is not significantly different from the 68% expected, given the independently-calculated frequencies of the two segments. However, Stephens and Justeson (1979) conclude that there is "an overwhelming tendency for /w/ to occur only if /j/ is also present...furthermore, /w/ and /j/ show a statistically significant tendency to occur together or not at all." Their conclusion is reached on the basis of a sample of 277 languages from a pre-final stage of the Stanford Phonology Archive project. The reported frequencies of /j/ and /w/ in the two samples are compared in Table 1. They agree in the percentage of languages which have /j/ but no /w/, but differ markedly in the overall percentages of /j/ and /w/ reported and in the percentage of languages with /w/ but no /j/. This result is surprising in view of the fact that the two samples contain a large common core of about 190 languages. It is conceivable that the claim that /w/ tends "to occur only if /j/ is also present" could be a) true of this common core of the two samples, b) true of the additional languages included by Stephens and Justeson, but c) massively violated in the additional languages in UPSID.

	/j/	no /j/	
/w/	71%	5%	75%
no /w/	14%	10%	
	86%		

(a)

	/j/	no /j/	
/w/	58%	1%	59%
no /w/	14%	27%	
	72%		

(b)

Table 1. Percentages of languages in (a) UPSID and (b) Stephens and Justeson with and without /j/ and /w/.

This requires assuming that a constraint violated by only 1% of the languages Stephens and Justeson surveyed is violated by 12% of the unshared languages in UPSID. This is unlikely, and it seems more plausible that the analytical criteria employed at Stanford at the time may have biassed the finding. This could have happened if the occurrence of a *class* of approximants was taken into account when choosing between vocalic or consonantal analyses of non-syllabic high vocoids so that the occurrence of /j/ favored interpretation of [u̯] as /w/ rather than /u/.

Other vocoid approximants are comparatively rare. They may be divided into two groups - those which are modifed variants of /j/ or /w/ and those with different places of articulation. Those in this second group include the labial-palatal approximant /ɥ/ (4 instances) and the velar approximant /ɯ̞/ or /ɣ̞/ (5 instances). These occur in less than 2% of the languages survyed.[T] They are not found to occur in modified forms.

Palatal approximants occur voiceless, laryngealized and nasalized. Labial-velar approximants occur voiceless, and laryngealized. The frequency of segments of this type is given in Table 2.

Segment	occurrences	percent of languages
/j/	7	2%
/j̰̊/	13	4%
/j̃/	3	1%
/ʍ/	11	3%
/w̰/	12	4%

Table 2. Frequency of modified /j/ and /w/.

Laryngealized approximants /j̰/ and /w̰/ occur with approximately equal frequency and are restricted to languages with other glottalic or glottalized segments in their inventories and with plain voiced /j/ and /w/. Greenberg (1970) suggested that /j̰/ fills the place of an anticipated palatal implosive in languages with an implosive series and a palatal place of articulation for stops (cf. Fordyce, in this volume). There does not seem to be support for this supposition in the available data. A diachronic source of this kind for /j̰/ would predict that it should be markedly more frequent than /w̰/ for which no parallel source is proposed. It also predicts that /j̰/ should occur in inventories which contain a subset such as that given in (1).

(1)	p	t	c or tʃ	k
	ɓ	ɗ	j̰	

Only 5 of the 13 instances of /j̰/ in the survey occur in the kind of inventory that would appear to support Greenberg's suggestion. There is a much stronger association between the occurrence of /j̰/ and of /w̰/; in 11 cases /w̰/ and /j̰/ occur together. In other words, there is only one exception to the statement that the presence of /w̰/ implies the presence of /j̰/.

The voiceless approximants /j̊/ and /ʍ/ differ fairly markedly in frequency, /ʍ/ being 1.7 times more frequent than /j̊/. This is particularly surprising when considered in comparison to the relative frequency of their voiced counterparts. The diachronic source of these voiceless segments is likely to be similar in both cases - documented instances seems to predominantly arise from a cluster of a voiceless obstruent and the approximant, or from labialized or palatalized voiceless obstruents, which may be equivalent. Thus (one source of) Hupa /ʍ/ is from Proto-Athabascan */ʃʷ/ (Huld 1980), and Middle English /ʍ/ is derived from Old English /xw/, and */ʍ/ in early Northern Tai is derived from Proto-Tai */xw/ (Li 1977). They are likely to exit from inventories in a variety of ways including vocalization (as in the widespread merger of */ʍ/ with /w/ in Modern English), collapse into an undifferentiated voiceless vowel phoneme (as in the special development of */ʍ/ before /u/ and /o/ in English words such as "who", "whoop"[1] "whole"), or fricativization (as in the idiolectal [ç] for the initial segments /hj/ = [j̊j] in English words like "huge" "human" etc. and the change of earlier /w̥/ in German into modern /v/).

The relative frequency of /ʍ/ and /j̊/ suggests that there may be some factor which favors the development of /ʍ/ over /j̊/ or favors the loss of /j̊/ more than /ʍ/. It is probably the case that a true voiceless palatal approximant is poorly distinguishable from /h/, which occurs in most languages and hence is liable to collapse together with it. If on the other hand it is articulated more forcefully to preserve the distinction it would become a palatal fricative. A voiceless labial-velar approximant may survive better

1. This word now frequently receives a spelling pronunciation with /w/ or even /ʍ/ supplanting the historically derived /h/.

because its two strictures produce two cavities with formant-like resonances which are rather close to each other in frequency and hence reinforce each other (cf. Ohala and Lorentz 1978).

3.0 Approximants and related vowels

The approximants /j/ and /w/ are closely related to the high vowels /i/ and /u/ respectively. The vast majority of languages have both these vowels, but there are more cases in which /u/ is missing than /i/ - in fact, /u/ is the most frequently missing of the major peripheral vowels (Disner, in this volume). The greater frequency of /i/ is undoubtedly a predictor of the greater frequency of /j/. However, for both /j/ and /w/ there are a few languages which have the approximant but lack the corresponding vowel. The numbers are given in Table 3.

	Languages	% of sample
/j/ but no /i/	8	3%
/w/ but no /u/	23	7%

Table 3.

There are about 3 times as many cases of /w/ occurring without /u/ as of /j/ occuring without /i/. Disner suggests that the systems without /u/ may be regarded as falling into two principal classes: those with a 'compensating' vowel which is high or back or rounded but not all three (such as /ɨ/, /ɯ/, /ʉ/ etc) and those which simply have a gap (and whose highest back vowel is usually /o/). This is suggestive of a variety of possible sources for /w/ and may predict that the class of /w/ segments in languages may vary phonetically through a greater range than /j/.

The less frequently occurring approximants /ɥ/ and /ɣ/ were also investigated in relation to the corresponding vowels, in this case /y/ and /ɯ/ respectively. The numbers are given in Table 4.

	Languages
/ɥ/ and /y/	3
/ɥ/ but no /y/	1
/ɣ/ and /ɯ/	1
/ɣ/ but no /ɯ/	4

Table 4.

These suggest that /ɥ/ is most likely to occur if /y/ also occurs in the inventory but that there is no such dependence of /ɣ/ on the occurrence of /ɰ/. However since these numbers are so small no great reliance should be placed on these indications.

4.0 Approximants and related consonants

a) /j/ and palatalized consonants

True palatalized consonants, that is, ones with a palatal secondary articulation usually perceptible because of a /j/-like offglide, occur in about 10% of the languages in the survey. Since disyllabification of high vowels is a major process creating both /j/ and palatalized consonants, it might be expected that palatalized consonants would usually occur only with /j/ (cf. Bhat 1978). There are, however, three languages in the survey which have palatalized consonants but no /j/ phoneme - exactly the number which would be predicted if there was no association between these two classes of sounds. Of these three, Ocaina seems straightforward, but Muinane has a voiced palatal fricative /ǰ/ with [j] as an allophone, and Ket also has /ǰ/, albeit largely restricted to intervocalic positions. Thus, a generalization stating that palatalized consonants occur in inventories containing /j/ or /ǰ/ would have only one exception in the survey.

b) /w/, labial-velar consonants and labialized velars.

As /w/ has two strictures of equal rank it falls into a class with other labial-velar consonants, especially /kp/ and /gb/ which are the most frequently-occurring labial-velar consonants after /w/. These labial-velar stops may vary a good deal in their initiation (Ladefoged 1964) but belong together by virtue of their shared places of articulation. The co-occurrences between /w/ and /kp,gb/ are shown in Table 5 below.

	Languages	% of sample
w, kp, gb	19	6%
w, kp	1	0%
w, gb	2	1%
no w, but kp, gb	1	0%
	23	7%

Table 5.

An assumption that there is no relation between the occurrence of /w/ and /kp/ or /gb/ predicts that we should expect 16 cases of /w/ with /kp/ (20 actual) and 16 or 17 cases of /w/ with /gb/ (21 actual). The observed numbers suggest that there is a tendency for /kp/, /gb/ to occur in systems with /w/ in preference to those lacking /w/. The one exception, Kpelle, will be discussed below.

By far the most frequent labialized consonant types are labialized velar stops (cf. Ohala and Lorentz 1978). Again there is a similarity between members of this class and /w/, besides a historically similar source in many instances. The co-occurrences of /w/ with /k^w/ are shown in Table 6. (/g^w/ only occurs if /k^w/ occurs, so is not separately listed).

	Languages	% of sample
/w/ and /k^w/	35	11%
no /w/ but /k^w/	5	2%

Table 6.

Random co-occurrence of /w/ and /k^w/ would predict that there would be 30 languages in the sample which contained both these segments. The observed number (35) suggests a weak tendency for /k^w/ to be more likely to occur in languages which also have /w/. The exceptions to the trend in UPSID, are Mixtec, Guarani, Wantoat, Chipewyan and Kpelle. In Chipewyan the labialized velars have a rather marginally contrastive status, being largely restricted to occurrence before back rounded vowels where plain velars do not occur. Kpelle is unusual in being the only language in the survey which has both labial-velar and labialized velar stops; and it also lacks /w/ !

5.0 Other Approximants

In addition to the four approximants /j, w, ɥ, ɣ/ discussed above, it may be noted that 6 languages (2%) have a bilabial approximant /β/ and 6 have a labio-dental approximant /ʋ/. All the remaining approximants which are reported to occur in the languages of UPSID have been grouped into the class of liquids and are discussed elsewhere in this volume.

6.0 Summary

Most languages have /j/ and /w/, with /j/ being more frequent. However an implicational hierarchy cannot be set up between these two segments since /w/ occurs without /j/ in too many languages. The greater frequency of /j/ is parallel to the greater frequency of /i/ than /u/, but these facts are not directly related since /j/ may occur without /i/ and /w/ without /u/. Modified varieties of /j/ and /w/ only occur in languages with the plain counterparts. There is some tendency for an association between the occurrence of palatalized consonants and /j/ and between labial-velar stops and labialized velars (and other labialized consonants) and /w/.

REFERENCES

Bhat, D.N.S. (1978) "A general study of palatalization" *Universals of Human Language 2: Phonology* (ed. J.H. Greenberg et al.) Stanford University Press: 47-92.

Greenberg, J.H. (1970) "Some generalizations concerning glottalic consonants, especially implosives" *IJAL* 36: 123-145.

Huld, M. (1980) "Tone in Proto-Athabascan" Unpublished paper. Department of Classics, UCLA.

Ladefoged, P. (1964) *A Phonetic Study of West African Languages* Cambridge University Press, London.

Li, F-K. (1977) *A Handbook of Comparative Tai* University Press of Hawaii, Honolulu.

Ohala, J.J. and J. Lorentz (1978) "The story of [w]: an exercise in the phonetic explanation for sound patterns" *Report of the Phonology Laboratory, U.C. Berkeley* 2: 133-155.

Pike, K.L. (1943) *Phonetics* University of Michigan Press, Ann Arbor.

Stephens, L.D. and J.S.Justeson (1979) "Some generalizations concerning glides" *Proceedings of the 8th Annual Meeting of the Western Conference on Linguistics* (ed. D.I.Malsch et al.) Linguistic Research Inc., Carbondale and Edmonton: 151-164.

On the Nature of Glottalic and Laryngealized Consonant and Vowel Systems

James F. Fordyce

1.1 Introduction.

This paper presents the results of a survey of the occurrence of glottalic consonants and other glottalized segments in a sample of the world's languages, and relates their occurrence to the rest of the segments in the phonological inventories of the languages. Greenberg (1970) discussed the distribution of glottalic consonants cross-linguistically and language-internally, but relatively little subsequent work has been presented in substantiation. In fact, much of the work addressing Greenberg's claims is of a non-corroborative nature (e.g., Campbell, 1973; Pinkerton, 1979). Other works are of an explanatory nature, but suffer from a lack of sufficient data upon which to base generalizations (e.g., Hamp, 1970; Javkin, 1977). Our goal is to determine whether the place of articulation preference hierarchies for implosves and ejectives posited by Greenberg (1970) can be substantiated and to discover other distributional patterns relating to glottalic and glottalized segments. This will be done through a survey of languages selected for the purpose of such sampling primarily on the basis of considerations of genetic balance. Greenberg's generalizations were based on a survey of languages for which data was available rather than on a carefully structured sample and it is possible that the hierarchies proposed reflect only the tendencies of a particular areal or genetic grouping and not more general cases.

This paper will also seek to determine from the distributional findings the underlying phonetic motivation for the patterns exhibited. We believe that these two processes are synergistic and can achieve together an explanation for glottalic systems which would otherwise be elusive.

1.2 The UCLA Phonological Inventory Database.

The generalizations within this paper are drawn from a significantly larger database than Greenberg's. Almost double the number of phonological inventories were available. The data are drawn from the UCLA Phonological Segment Inventory Database (UPSID) established in 1979.[1] The UPSID consists of segment inventories of (at present) 317 languages, 191 of which are included (sometimes in a revised form) from the Stanford Phonology Archive (SPA) made available through the cooperation of the Language Universals Project at the Department of Linguistics, Stanford University. The remaining 126 languages were archived at the University of California, Los Angeles, from original sources. The criteria for inclusion of a language in the UPSID are basically similar to those for the Stanford Phonology Archive except that number of speakers was not considered and only one language from within a closely related group is included. The criteria include considerations of genetic balance, i.e., that one language from each moderately distant genetic grouping be included, and the availability of primary data, generally in the form of published articles and/or grammars of the language. However, the amount and type of data included for each language in the UPSID differs somewhat from that found in the SPA. The information coded for each language in the UPSID consists only of a segment inventory of the phonemic units represented by a phonetic specification of a "most typical" or "most representative" allophone. The 61 phonetic features to which values are assigned for each segment in the inventory are subdivided as features for primary places of articulation (i.e., for particular combinations of aperture and airstream), other consonant features (lateral, sibilant, etc.), secondary articulations (labialization, nasalization, palatalization, etc.), vowels, diphthongs, and voice quality (voiced, voiceless, aspirated, long, etc.). Some of the implications for the description of possible segments entailed in the use of these particular features will be discussed in section 1.4 of the paper. Unlike the SPA, the UPSID presently contains no information regarding tone, stress, syllable structure, phonotactic constraints, and phonological alternations. Length is included where relevant to only a partial set of vowels or consonants (i.e., in those cases in which it cannot be treated suprasegmentally).

1.3 Relation to Greenberg (1970).

Greenberg´s 1970 article is the classic exposition of the patterns of occurrence of glottalic consonants throughout the world´s languages. But, whereas Greenberg "intended to illustrate a methodology in which synchronic systems, and diachronic generalizations regarding change, are mutually supportive, and in which historical comparison and reconstruction enter in an integral way," (1970:37) we have for the most part limited our inquiry and discussion to synchrony. However, facts which appear to elaborate on Greenberg´s diachronic accounts will be noted. The main aim of this paper is to substantiate and further elaborate on Greenberg´s point of articulation hierarchies of ejectives and implosives, the hierarchical relations to the plain series, and to consider hierarchies of manner of articulation of ejectives and implosives not explicitly discussed in Greenberg (1970). In addition the nature of the explanations required for our findings will be discussed.

1.4 Glottalic and Laryngealized Sounds.

The types of sounds under investigation in this paper consist of those articulated using the glottalic airstream mechanism (i.e., ejectives and implosives) and "glottalized" segments where the glottal constriction does not serve as the airstream initiator (i.e., preglottalized consonants and laryngealized consonants and vowels). Only those sounds articulated using the glottalic airstream mechanism will be referred to as glottalic. Pulmonic or velaric "glottalized" sounds will generally be referred to as laryngealized. In general ejectives are those sounds produced through the raising of the larynx with the closed glottis compressing the air within the oral cavity. The occlusion or constriction is subsequently released with considerable outward airflow. Implosives, on the other hand, are articulated by rapid lowering of the larynx although in this case the glottis is usually not closed, but rather the vocal folds are allowed to vibrate through leakage of pulmonic air into the oral cavity (see Catford, 1939). Thus ejectives tend to be unvoiced while implosives tend to be voiced. As noted by Ladefoged (1968, 1976) implosives do not always entail inward oral air flow upon release. The distinction between truly imploded consonants and those which are preglottalized or laryngealized with incidental implosion (Greenberg, 1970; Ladefoged, 1968, 1976) has been almost as difficult to maintain in the UPSID as in Greenberg's survey. That is, in general these sounds are not distinguished in the literature, but where the airstream initiator of the most typical allophone is known to be or may be other than glottalic, the sound is classified in the UPSID as laryngealized. As both Ladefoged (1968) and Greenberg (1970) conclude, the potential phonological contrast of these differing types is not realized in any of the languages known to them directly or through the literature, nor we might add in any included in the UPSID.

It should further be noted with regard to the types of sounds under discussion here that the possible feature specifications available in coding segments in the UPSID limit the description of segments in certain phonetically plausible ways. Thus, ejectives must be specified as either ejective stop (glottalic egressive stops), ejective affricate (glottalic egressive affricates), or ejective fricative (glottalic egressive fricatives). The presumedly phonetically impossible ejective approximant, for instance, is therefore excluded a priori by the coding mechanism. Likewise, implosives are coded only as implosive stops. The phonetically implausible implosive affricate or fricative is therefore excluded. Hoard (1978), however, does report as secondary allophones of the affricates /ts´/ and /tl´/ (both ejectives) the voiced implosive counterparts [ɗz] and [ɗl] for the Tsimshian language Gitksan (not included in the UPSID). From the description provided it is not clear if these voiced alternates of the voiceless ejective affriates are actually imploded or not. However, since they contrast phonetically with corresponding plain voiced affricates in certain environments, it is probable that their articulation is, if not implosive, at least accompanied by laryngeal constriction. Still, on the phonological level the particular allophone characterizable as "most typical" (i.e., that which would be coded by the UPSID) is in both cases the voiceless ejective affricate and hence, does not present a secial problem for the coding mechanism as devised. A similar analysis is presumably maintainable in regard to the voiced palatal implosives of Fula and Serer (also not in the UPSID) which are slightly affricated (Ladefoged, personal communication). The feature "laryngealized" is not limited in this way, reflecting the possible and actual occurrences of preglottalized or laryngealized stops, affricates, fricatives, approximants, and vowels. Thus as Greenberg (1970:2) states, "(t)he phonological opposition in individual languages between ejectives and injectives applies effectively only to obstruents, and is neutralized for sonorants and semivowels." One final redundancy regarding "glottalized" sounds, not built into the UPSID coding system, is the fact that the relevant lateral affricates and fricatives included in the survey are all reported as ejectives. In other words, the phonetically plausible laryngealized fricative or affricate is simply not reported in the literature and hence does not occur in the survey.

2.1 Ejective Systems.

Of the 317 languages in the UPSID, 57 contain ejectives, making ejective systems the most preferred of systems with glottalic or "glottalized" (laryngealized) segments. Twelve of these languages also exhibit implosive stops. Fifteen of these languages also exhibit laryngealized stops, fricatives, sonorants, and/or vowels. Ejective systems are fairly widespread genetically although the vast majority, 37, are to be found in the Amerind family - seven in the Southern Amerind geographical region and thirty in the Northern Amerind region. Nine of the languages with ejectives are of the Afro-Asiatic family, three Nilo-Saharan, three Caucasian, two Niger-Kordofanian, and one each of Indo-European and Khoisan. Ejective systems are not known to occur outside of these major language families although it is at least conceivable that ejective systems may occur in the Austro-Asiatic and/or Austro-Tai families since "glottalized" stops of the preglottalized or laryngeaized variety are found there.

2.11 Ejective Stops.

As Haudricourt (1950) noted, ejectives exhibit a strong preference for back articulations. Greenberg's (1970) survey also demonstrated that intrasystemically this tendency manifests itself in hierarchical order. The implied hierarchy for ejective stops is as follows: a language having a bilabial ejective stop also has the alveolar and velar ejective stops; a language having an alveolar ejective stop also has a velar ejective stop; and a language having only one ejectve stop will have the velar. Javkin (1977) provides the count of the occurrences of ejectives in the Stanford Phonology Archive shown in Table 1.

Labial	Alveolar	Palatal	Velar	Uvular
26	29	7	31	15

Table 1
Ejective Stops in the Stanford Phonology Archive.

He notes that these "implicational relations . . . hold in the languages of the archive . . . although the numerical preferences for ejectives at the three major places of articulation do not seem overwhelming." Javkin further notes that palatals and uvulars do not maintain the tendency to prefer a further back articulation over a further front one "since these places of articulation tend to disfavor stops." We will return to this obsevation in section 5.3 of the paper, but it is important to stress that the disfavoring of palatals and uvulars is not specific to glottalic consonants (see Gamkrelidze, 1978, and Nartey, 1979, for details concerning non-glottalic stop distributions).

In general, the implicational hierarchy posited by Greenberg (1970) for place of articulation of ejective stops is confirmed in a count of occurrences in the UPSID as follows:

	Labial	Alveolar[2]	Palatal	Velar	Uvular
Ejective stops:	38	51	8	53	18
labialized :	-	-	-	16	9
palatalized :	-	-	*	2	-
long :	-	2	-	2	-
Total :	38	53	8	73	27

Table 2
Ejective Stops in the UPSID.

Here as in the case of the SPA, the numerical preferences for simple ejective stops at the three places of articulation (labial, alveolar, and velar) do not appear particularly strong, especially between alveolar and velar articulations. But when additional stops with secondary articulations are taken into account, especially those with labialization, a clear preference for the velar place of articulation over the alveolar emerges. Significantly, athough certain secondary articulations are implausible for certain places of articulation (e.g., palatalized palatals), such plausibility restrictions do not apply to the alveolar position and cannot be invoked in explanation of the lack of secondary articulations there. The above count does not, however, reveal the fact that there are various intrasystematic violations of the implied hierarchy in the UPSID, in contrast to the facts reported fo the SPA. These violations as well as the more frequently found patterns will be discussed in the following sections.

Table 3 shows that the most typical system of ejective stops contains three places of articulation: labial, alveolar, and velar:

# OF PLACES	Labial	Dent/Alveol	Palatal	Velar	Uvular	# OF LANGS	TOTAL
1	y					2	7
				X		5	
2	y			y		1	10
		X		X		9	
3		y		y	y	1	19
	X	X		X		18	
4	y	y ¦ y		y		2	15
	y	y	y	y		2	
		y	y	y	y	2	
	X	X		X	X	9	
5	y	y ¦ y		y	y	2	5
	X	X	X	X	X	3	
X = most common y = other							56

Table 3

Preferences in Number of Places of Articulations for Ejective Stop Systems.

Representative of such distribution is the stop system of Eastern Armenian (IE)[3]:

p t k

p^h t^h k^h

p´ t´ k´

or Tzeltal (NA) whose stop system is as follows:

p t k^h

p´ t´ k´

b d g

The next most preferred ejective stop systems are those containing either two ejective stops - the alveolar and the velar - or four ejective stops - the labial, alveolar, velar, and uvular. Typical of the ejective stop systems with only two places of articulation is Itonama (SA) whose stop inventory is as follows:

p	t	k
	t^j	
	t´	k´
b	d	

or, as part of a more complex stop system, Tigre (AA):

	t^h	k^h		$t{:}^h$	$k{:}^h$
	t´	k´		t´:	k´:
b	d	g	b:	d:	g:

Representative of stop systems with four ejective stops are Quileute (NA):

p	t	k	q	ʔ
p´	t´	k´	q´	
b	d	(g)		
		k^w	q^w	
		k'^w	q'^w	

and Quechua (SA):

p	t	k	q
p^h	t^h	k^h	q^h
p´	t´	k´	q´

The fourth most common system contains only the velar ejective stop, e.g., Wichita (NA):

t	k	?
t^h	k^h	
	k´	
	k^w	
	k^{hw}	

The fifth most common system contains five places of articulation: labial, alveolar, palatal, velar, and uvular. An example of such a system would be Jaqaru (SA):

p	t	c	k	q
p^h	t^h	c^h	k^h	q^h
p´	t´	c´	k´	q´

There are a few exceptional languages which have ejective stops in different places than the languages which have the typical ejective stop system for a given number of ejective stops. These do not affect the general tendencies noted above. They usually, though not always, constitute exceptions to the implied hierarchy of preferences for point of articulation of ejective stops as well. For example, we find two languages, Lelemi (NK) and Jivaro (SA), with only one ejective stop, but it is the labial rather than the expected velar:

Lelemi stop inventory:

kp		t	k
	p´		
gb	b	ḍ	g

Jivaro stop inventory:

p	t	k	?
(p´)			

Note that in Jivaro the labial ejective is characterized in the UPSID as being marginal in that it has "extremely low frequency of occurrence." We also find cases such as K´ekchi (SA) which has three ejective stops, but the third is at an unexpected place of articulation - uvular. K´ekchi does not violate Greenberg´s place of articulation preference hierarchy:

K´ekchi stop inventory:

p	t	k	q	ʔ
	t´	k´	q´	
´b				

In other words, it is not an explicit requirement of Greenberg´s hierarchy that any third ejective stop be the bilabial one, but rather a tendency for this to be the case (as it clearly is). In this regard we might note that K´ekchi actually does exhibit a "glottalized" stop in the expected labial position described as a (bilabial) laryngealized voiced plosive.[4]

Other cases which violate the expected occurrences for place of articulation for a given number of ejective stops are not easily explicable. As we have seen, the Lelemi system contains a voiced/voiceless distinction for stops at four places of articulation: labial, alveolar, velar, and labiovelar. But the voiceless labial is ejective, hence the only ejective is labial rather than the expected velar. Why such markedness obtains only in the labial stops is unclear.

Berta (NS), a language with two ejective stops, one labial and the other velar, violates the tendency towards alveolar and velar articulations for such two stop systems, thus:

p´		k´
	ɗ	
b	d	g

An argument similar to that for K´ekchi above may be made in this case since the alveolar slot is at least filled by a glottalic consonant (here a voiced implosive, actually dental). This implosive stands alone in Berta. Thus a sort of trade off is achieved between the "deficient" ejective system and the isolated alveolar implosive to create a symmetrical series of glottalic stops. This sort of systematic conflation of implosive and ejective stop series within a language is exactly as Pinkerton (178) reports for five Quichean (Mayan) languages in her study. Pinkerton shows that implosive and ejective (as well as laryngealized, and for the bilabials only, plain) sound types correspond in these five languages (1978:11):

Language				
1. K´ekchi	b/´b	t´	k´	q´/$q^{<}$
2. Pocomchi	$p^{<}$	$t^{<}$	k´	$q^{<}$
3. Cakchiquel	b/´b	t´	k´	$q^{<}$
4. Quiche	ɓ	t´	k´	$q^{<}$
5. Tzutujil	ɓ	ɗ	k´	$q^{<}$

These correspondences confirm that close diachronic relationships may exist between different glottalic stops. However, the Berta and Lelemi systems are still unusual and different from the Quichean systems above in having the labial glottalic be a voiceless rather than the expected voiced implosive.

Hupa (NA) seems particularly deviant in having three ejective stops to the exclusion of both the velar (except as an extremely limited extrasystematic phoneme) and the labial, thus:

t	c		q
t^h	c^h		q^h
t´	c´	(k´)	q´

In this language the ejective stop system shares the same unusual places of articulation as the voiceless plain and the voiceless aspirated stop series. Our explanation for this curious ejective series rests on Greenberg´s (1970: 23) observation that "unvoiced ejectives never exist without plain stops." This hierarchy which we hereafter refer to as the ejective to plain hierarchy is clearly not only relevant here, but reveals an important point concerning these particular hierarchies as well as phonological hierarchical relations in general and that is simply this: relationships among phonological hierarchies may themselves be hierarchical. Indeed we must posit that the relationship between the ejective to plain hierarcy and the hierarchy of preferences for place of articulation of ejective stops is such a case, with the ejective to plain hierarchy preceding and commanding the place of articulation hierarchy for ejective stops. By postulating such a relationship between the two hierarchies we maintain the viability of both (since there are consequently fewer exceptions to either) and we are able to explain unusual and otherwise exceptional systems such as that of Hupa.

The four ejective stops of Haida do not violate the place of articulation preference hierarchy for ejective stops, but they do evidence an unexpected pattern - dental, palatal, velar, and uvular:

p	t	c	k	q	k^{w}	q^{w}
p^{h}	t^{h}	c^{h}	k^{h}	q^{h}	k^{hw}	q^{hw}
	$\underset{<}{t}'$	c'	k'	q'	k'^{w}	q'^{w}

The pattern appears explicable only on the grounds that back articulations are generally preferred in the Haida stop system as seen above. And, of course, the occurrence of the particular ejective series evidenced, though unusual for a four ejective system, is predicated upon the occurrence of corresponding plain stops for each ejective as we would expect given the ejective to plain stop hierarchy noted above. It is certainly in accordance with Greenberg's general view that the ejective that is missing is the bilabial one.

The Wappo system of four ejective stops likewise constitutes not an exception to Greenberg's place of articulation hierarchy, But rather exhibits the unexpected contrast between dental and alveolar ejective stops:

Wappo stop inventory:

p	$\underset{<}{t}$	$\underset{>}{t}$	k	ʔ
p'	$\underset{<}{t}'$	$\underset{>}{t}'$	k'	
b	$\underset{<}{d}$		g	

Here again the occurrence of an unusual ejective series can be attributed to the precedence established by the plain stop series in the language.

The two "exceptional" languages with five ejective stops, Nez Perce (NA) and Pomo (NA), also exhibit a dental/alveolar distinction, but do not technically violate the place of articulation preference hierarchy for ejective stops:

Nez Perce stop inventory:

p	$\underset{<}{t}$	$\underset{>}{t}$	k		ʔ
p'	$\underset{<}{t}'$	$\underset{>}{t}'$	k'	q'	

Pomo stop inventory:

p	$\underset{<}{t}$	$\underset{>}{t}$	k	q	ʔ
p´	$\underset{<}{t}$´	$\underset{>}{t}$´	k´	q´	
b		d			

As expected from our observations so far, the ejective system is largely a function of the corresponding plain stop system in both of these languages with the exception of Nez Perce /q´/ which appears in violation of the ejective to plain stop hierarchy. The uvular place of articulation in Nez Perce is, however, represented by the fricative /X/ as well and it is therefore at least possible that the fricative here can be regarded as a precedent for the ejective. Clearly given the rest of the Nez Perce stop system, a fifth ejective at the uvular place of articulation would be significantly more viable than one at the palatal region. We might also speculate that given the ejective to plain stop hierarhchy, /X/ is most likely related diachronically and comparatively to a voiceless uvular stop /q/. Thus the ejective stops fill the positions they are able to, given the point of articulation restrictions of these languages.

Finally, the four ejective stop systems of Kefa and Maidu contain palatal ejectives rather than the more common uvular ejectives. But these languages also lack a corresponding plain stop in the uvular position:

Kefa stop inventory:

p	t	c	k	ʔ
p´	t´	c´	k´	
b	d	ɟ	g	

Maidu stop inventory:

p^h	t^h	c^h	k^h
p´	t´	c´	k´
ɓ	ɗ		

The Maidu case is particularly interesting since it demonstrates that (1) there does not exist an aspirated to plain stop hierarchy requiring corresponding plain stops for any aspirated stop in a system, and (2) that the ejective to plain stop hierarchy entails an ejective to aspirated stop hierarchy as well (although this is certainly the more limited case). This is not surprising since the relationship between aspirates and "plain" voiceless stops is clearly less drastic than that between ejectives and plain voiceless stops (i.e., VOT differences as opposed to differences in airstream mechanism).

2.12 Ejective Affricates and Fricatives.

Ejective affricates and fricatives have rather limited occurrence. Forty-three languages of the UPSID (14%) contain ejective affricates. In all but one of these languages, ejective stops occur as well. The exception is Iraqw which does, however, contain glottalic (implosive) stops. Clearly, then, the implication is that ejective affricates occur only in those systems containing glottalic stops (almost exclusively ejective). The most commonly occurring ejective affricates are the stop plus sibilant /ts'/ and /tʃ'/, and the stop plus lateral /tl'/. The figures are as follows:

	tθ'	tl'	ts'[5]	tʃ'	ts'	kxl'
simple	1	13	35	36	3	1
labialized	-	-	2	2	-	-
prevoiced	-	-	1	1	-	-
prevoiced/ aspirated	-	-	1	1	-	-
long	-	-	1	1	-	-
Total	1	13	40	41	3	1

Table 4
Ejective Affricates in the UPSID.

It is interesting to note that one of the most common ejective affricates is that articulated at the alveopalatal-palatal region, precisely that area least favored and usually null in ejective stop systems. Thus Greenberg (1970:17) notes that, "for the palatal region in particular, it appears that the optimal ejective is the alveopalatal č' rather than a stop. This is true also for the plain obstruent, but in the case of the ejective this is even more the case." This latter statement was made since Greenberg found "no example of an ejective palatal stop" in his sample. However, they not only occur, but actually exist with a contrasting affricate. Thus of the eight languages in the UPSID with the palatal ejective stop, five also have the alveopalatal or palatal ejective affricate, /tʃ'/.

Further, since alveopalatal or palatal ejective affricates clearly are equally frequent as dental or alveolar ones, no preferred point of articulation can be established. Thus, twenty-seven of the languages contain both, two contain only /ts'/, five contain only /tʃ'/, and the remainder contain some combination of /ts'/ or /tʃ'/ and another ejective affricate. Only /ts'/ or /tʃ'/ may therefore constitute the sole ejective affricate in a language system. It should be noted that in only two cases does the lateral ejective affricate /tl'/ occur without the correspondng sibilant ejective affricate /ts'/ in the system. That is, generally the presence of /tl'/ in a language implies the presence of /ts'/ as well. One of the exceptions to this observation is again Haida with the unusual three place (alveolar, palatal, uvular) stop system discussed above. Here, too, there is no corresponding plain affricate (i.e., /ts′/) at this point of articulation. The other exception, Wintu, lacks plain /ts/ as well. However, if we posit an ejective to plain affricate hierarchy corresponding to the ejective to plain stop hierarchy for languages, our two exceptions are accounted for. Thus, given /tl′/ in a language, we would expect to find /ts′/ unless the plain alveo-palatal affricate (/ts/ or /dz/) does not occur. It must also be noted that in every case but one (Ik) /tl′/ occurs with /tʃ′/. In Ik, though, the corresponding plain affricate /tʃ/ does occur. Generally, nevertheless, we posit that the occurrence of /tl′/ implies the occurrence of both /ts′/ and /tʃ′/.

With regard to ejective fricatives Greenberg (1970:17) noted that they are "relatively infrequent and always imply the presence of some ejectives with abrupt onset." In the languages of the UPSID, ejective fricatives imply ejective stops without exception. Most commonly they imply ejective affricates as well, but exceptions do occur. Ten languages of the UPSID contain ejective fricatives (3%), only three of these do not contain ejective affricates. Ejective fricatives reported in the survey are as follows:

	Φ′	f′	s′	ṣ′	ʃ′	x′	X′	ɬ′
simple	1	1	10	1	6	2	1	2
labialized	-	-	-	-	-	1	1	-
palatalized	-	-	-	-	-	-	-	1
Total	1	1	10	1	6	3	2	3

Table 5
Ejective Fricatives in the UPSID.

Only /s′/ and /ʃ′/ occur in systems not containing ejective affricates, they are also the only ejective fricatives which occur alone (i.e., without other ejective fricatives).

2.13 Secondary Articulation Features.

As noted above, approximately 18% (57 of 317) of the languages in the UPSID contain ejectives of one kind or another. All but one of these (Iraqw) contain ejective stops. Not only are plain ejectives found in these languages, but often either labialized or palatalized ejective stops occur as well. Thus, eighten languages of the survey contain labialized stops within their ejective stop systems. Two of these systems contain a palatalized velar stop as well (Hausa and Kabardian). Labialization only occurs as a secondary articulation feature in the velar and uvular ejective stops. Labialization and palatalization also occur as secondary articulation features on certain ejective affricates and fricatives although this is not common. The only occurrences of labialized ejective affricates are in Lak with /ts′ʷ/ and /tʃ′ʷ/. The only labialized fricatives are those found in Tlingit with both /x′ʷ/ and /X′ʷ/. The only palatalized ejective fricative, /$ɬ'^{j}$/ occurs in Kabardian (CA). Although there are few cases from which to draw conclusions, an expected hierarchy is borne out by this data; namely that labialized ejective affricates and fricatives occur only when there are (1) corresponding non-labialized ejective affricates or fricatives and (2) labialized ejective stops. The occurrence of labialized ejective stops as well implies the occurrence of corresponding non-labialized ejective stops in the system. More generally we can say that labialized ejectives imply the presence of simple ejectives which in turn imply the presence of plain stops, affricates, or fricatives at the same point of articulation. A similar implicational universal might be posited for palatalized ejective stops as well. Kabardian, however, presents a counter-example to the latter hierarchies in having both the labialized and palatalized ejective velar stops with no corresponding simple ejective velar stop. This may be explained, however, by the absence of the corresponding non-glottalic velar stop:

Kabardian stop inventory:

p^{h}	t^{h}	k^{hj}	q	ʔ
p'	t'	k'^{j}	q'	
b	d	g^{j}		
		k'^{hw}		$ʔ^{w}$
		k'^{w}	q'^{w}	
		g^{w}		

2.14 Phonation Types and Length.

As Greenberg (1970:2) notes, "the typical ejective obstruent is unvoiced." We find no exceptions to this in the languages of the UPSID due presumably to the phonetic implausibility of voicing while raising the larynx in compression of the oral atmosphere. There do, however, occur pre-voiced ejectives in one language of the survey, !Xũ (KH). The consonant system of !Xũ is in general rather complex containing 102 consonant phonemes, 55 of which are clicks. Snyman (1969) discounts the possibility of the coincidence of voice and ejection noting that "what actually happens is that the vocal cords are activated by pulmonary air and they produce a voiced unemitted sound which we represent [ɗ]. (Both the nasal and oral passages are closed.) The unemitted sound [ɗ] is swiftly followed by the articulation of the ejected sound . . . In close sequence [ɗ] and . . . [the ejected sound] is perceived as a vocalized sound." By convention, then, the coding of a sound in the UPSID as ejective and voiced implies the non-simultaneity of voicing and ejection with voicing preceding ejection.

Aspiration is not a common feature of ejectives though an aspirated alveopalatal ejective affricate is reported for Tigre (AA). Snyman (1969) reports for !Xũ aspiration in the pre-voiced ejective stops and affricates, although non-aspirated pre-voiced affricates also occur.

Length contrasts in ejectives occur in two languages of the survey, Tigre and Amharic (AA), and of course imply non-long counterparts for each long ejective (both stops and affricates occur).

2.2 Implosive Systems.

Thirty-three (9%) of the languages of the UPSID contain implosives. As noted in section 1.4, all implosives are stops; implosive affricates and fricatives do not occur in the languages sampled.

2.21 Implosive Stops.

Greenberg (1870:10) following Haudricourt (1950) and Wang (1968) noted that "injectives tend to have front articulation." He goes on to suggest that "if a language has [one] injective, it is ɓ; if two, they are ɓ and ɗ (the most common pattern); if three, they are ɓ, ɗ, and ʄ (the latter a palatal stop, often replaced, however, by ?y); and, if four, they are ɓ, ɗ, ʄ, and ɠ." The general preference for front articulations was borne out in a count of the Stanford Phonology Archive (Javkin, 1977) shown in Table 6:

Labial	Alveolar	Palatal	Velar	Uvular
17	15	2	2	0

Table 6
Implosive Stops in the Stanford Phonology Archive.

A count of implosives in the languages of the UPSID shows a similar frequency of occurrence of the various stops:

Labial	Alveolar	Palatal	Velar	Uvular
30	31	7	5	0

Table 7
Implosive Stops in the UPSID.

The data in Tables 6 and 7 suggest that a correction needs to be made to the implicational hierarchy posited by Greenberg. Indeed, as Greenberg (1970:11) notes, "there are a few languages whose sole injective is ɗ," thus violating the hierarchy altogether. Table 8 clearly suggests that the hierarchy is blind to the distinction between labial and alveolar implosives. They are essentially equally frequent and either may occur as the sole implosive in any inventory.

# OF PLACES	PLACES OF ARTICULATION				# OF LANGS	TOTAL
	Labial	Dent/Alveol	Palatal	Velar		
1	y				2	5
		X			3	
2	X	X			20	20
3	y	y		y	1	4
	X	X	X		3	
4	X	X	X	X	4	4
X = most common y = other						33

Table 8

Preferences in Number of Places of Articulation for Implosive Stop Systems

There are two languages in which /ɓ/ alone is reported in the UPSID - Kpelle (NK) and Zulu (NK) - and three languages in which /ɗ/ occurs as the sole implosive: Somali (AA), Berta (NS), and Kullo (NS). Those languages with /ɓ/ as the sole implosive come from the Niger-Kordofanian family, those with /ɗ/ as the sole implosive represent two families - Afro-Asiatic and Nilo-Shaharan. It is fairly clear that Greenberg (1970) based his generalizations upon a sample which, from the point of view of genetic balance, overemphasized the Niger-Kordofanian languages and underrepresented some less well documented language groups. The implicational hierarchy needs to be modified to incorporate our observations of equal frequency of labial and alveolar implosives. An actual preference for the alveolar position over the labial one should not be posited since the difference is so small (31 alveolar to 30 labial). Thus we posit that /ɓ/ implies the presence of /ɗ/ or nothing, /ɗ/ implies the presence of /ɓ/ or nothing, while /ʄ/ implies the presence of both /ɓ/ and /ɗ/; and /ɠ/ implies the presence of /ɓ/, /ɗ/, and /ʄ/.

As Greenberg observed, the two implosive system, one labial and one alveolar, is by far the most common or preferred system. An example of such a system is that of Doayo (NK):

kp	p	t	k
gb	b	d	g
	ɓ	ɗ	

or as part of a more elaborate stop inventory in Iraqw (AA):

kp	p	t	k	q	ʔ
gb	b	d	g		
	ɓ	ɗ			
				q^w	

The next most common system appears to be that consisting of only one implosive stop, the labial <u>or</u> the alveolar (these are treated together a single type owing to their function in the place of articulation hierarchy). Kpelle (NK) is a language with the bilabial implosive as the sole occurrence:

kp	p	t	k		k^w
gb	b	d	g		g^w
	ɓ				

Kullo (AA) is a language with the alveolar implosive as the sole occurrence:

	t	k	ʔ
	t'	k'	
b	d	g	
	ɗ		

The third most common system is that containing all four possibilities (labial, alveolar, palatal, and velar), e.g., Nyangi (NS):

p	t	c	k
ɓ	ɗ	ʄ	ɠ

followed by the three implosive system (labial, alveolar, and palatal) as exhibited by Yulu (NS):

kp	p	t	c	k
gb	b	d		g
	ɓ	ɗ	ʄ	

There remains, however, one exception to the place of articulation preference hierarchy for implosives among the languages of the UPSID. Thus, Hamer (AA) is reported to have only the labial, alveolar, and velar implosives, with a gap in the palatal position:

p	t	c	k
			k'
b	d	ɟ	g
ɓ	ɗ		ɠ

And Greenberg's (1970:15) note that such languages often "have ?y [glottalized palatal semivowel] in place of the expected palatal stop" simply does not apply here (?y is also lacking in the language). A rather implausible explanation could be made that /?i/ (i.e., the laryngealized high front vowel) is a sufficient representative of the palatal position. This would be less far fetched were /?i/ to behave as a transitional glide in certain phonological environments, the actuality of which is as yet to be confirmed.

2.22 Phonation Types, Retroflection, and Length.

Both Greenberg (1970) and Haudricourt (1950) noted that an implosive corresponding to a non-implosive dental is often retroflexed or, at least, articulated further back than the non-implosive. In the UPSID a retroflexed implosive occurs as the most common allophone in only one language, Somali. Generally, however, within the coding system of the UPSID entailing that segments are specifed as dental or alveolar only if clearly the case (as determined from the source material), it is difficult to verify the above claim (i.e., if either the implosive or the non-implosive is coded as "unspecified dental/alveolar", the issue cannot assuredly be decided). Of those cases which are distinguishable, six do not show an implosive with a further back articulation than the non-implosive counterpart (i.e., both are either dental or alveolar) while only two report either alveolar or retroflexed implosives with a further front non-retroflexed non-implosive counterpart. Clearly more phonetic detail than is availale in the UPSID and the sources it relies on would be necessary before the general validity of Greenberg's and Haudricourt's observation can be verified in this way.

Overwhelmingly, we may observe, implosives are voiced. Much work has been recently published, however, showing that beyond the shadow of a doubt unvoiced implosives do indeed occur (see particularly Campbell, 1973; Ladefoged, 1976; and Pinkerton, 1979). Greenberg (1970:7-9) notes voiceless implosives in several languages of the Munda group. Pinkerton's (1979) data show voiceless implosives as the unmarked allophone of a uvular glottalic stop in several Quichean (Mayan) languages, apart from K'ekchi. K'ekchi has a voiceless uvular implosive as an allophone of what is more commonly a uvular ejective. The other langages discussed, however, all have voiceless uvular implosives as the unmarked case. Interestingly, there are no reports in the literature of voiced uvular implosives to our knowledge; only voiceless ones are known to occur. Ladefoged (1976) reports for Owerri Igbo not only the presence of voiceless implosives, but contrasting voicing for the bilabial implosive. Igbo thus exhibits voiced and voiceless labial implosives and a voiceless alveolar implosive as well:

Igbo stop inventory:

	p	t	k	k^w
	p^h	t^h	k^h	k^{hw}
	$p^<$	$t^<$		
b^j	b	d	g	
b^hj	b^h	d^h	g^h	g^{hw}
	ɓ			

Length is contrastive only for the retroflexed implosive of Somali (AA):

	t^h		k^h	
b	d	ᶑ	g	G
b:	d:	ᶑ:	g:	G:

2.3 Ejective and Implosive Systems.

Twelve of the languages of the UPSID contain both ejectives and implosives. One might expect, given the point of articulation preference hierarchies for ejectives and implosives, that they rarely occur at the same point of articulation. In many cases, however, they do. Zulu, for instance, has both a labial ejective and a labial implosive stop. Ik (NS) has four implosive stops (labial, alveolar, palatal, and velar) and the velar ejective stop as well. Koma (NS) and Southern Nambiquara (SA) have labial and alveolar ejective and implosive stops and also have the velar ejective. Maidu (NA) has the greatest number of glottalic stops in the UPSID with the labial and alveolar implosive plus the labial, alveolar, palatal, and velar ejectives (see section 2.11). And we might note that Maidu also has a fairly limited set of non-glottalic stops consisting only of a plain voiceless series. Kullo (NS) and Mazahua (NA) both contain velar and alveolar ejective stops with the alveolar implosive and for Mazahua only, the bilabial implosive.

Mazahua stop inventory:

p	t	k	ʔ	k^w
p^h	t^h	k^h		k^{hw}
	t´	k´		$k^{\prime w}$
		g		g^w
ɓ	ɗ			

Hamer (NS) has the labial, alveolar, and velar implosives plus the velar ejective while Otomi (NA) has the reverse - a labial implosive plus labial, alveolar, and velar ejectives:

Otomi stop inventory:

p	t	k	ʔ	k^w
p^h	t^h	k^h		
p´	t´	k´		$k^{\prime w}$
b	d	g		g^w
ɓ				

Iraqw (NS) and Hausa (NS), on the other hand, show no overlap of implosives and ejectives - both have the labial and alveolar implosives, and while Hausa has the velar ejective stop, Iraqw has the ejective affricate /ts´/:

Hausa stop inventory:

	t	k	ʔ	k^w	k^j
		k´		$k^{\prime w}$	$k^{\prime j}$
b	d	g		g^w	g^j
ɓ	ɗ				

Berta (NS) also exhibits no overlap, but its system of glottlic stops appears to at least violate the point of articulation hierarchy for ejectives. It has only the labial and velar ejectives with an alveolar implosive stop (see section 2.11).

3.1 Laryngealized Systems.

As noted in section 1.4, the feature "laryngealized" in the UPSID can imply for obstruents preglottalization, as well as laryngealization (creaky voice). By far the most common occurrence is preglottalization of voiced stops or affricates. Eleven languages in the survey utilize the "laryngealized" feature in their stop system; one for a voiceless series, two for a voiceless click series, and eight for voiced stop series. The voiced series generally mirror the most common distibution for implosives, i.e., two stops, one labial and one alveolar; while the voiceless series mirror the most common distribution for ejectives, i.e., three stops - labial, alveolar, and velar. This similarity of distribution as well as the fact that laryngealized or preglottalized voiced stops do not contrast with implosives in any known language confirm the correctness of Greenberg's decision to collapse these distinctions. He observes "that all the generalizations . . . regarding implosives apply equally to all these types."

3.2 Laryngealized Sonorants.

As Greenberg (1970:2) noted "(t)he phonological opposition in individual languages between ejectives and injectives applies effectively only to obstruents, and is neutralized for sonants and semi-vowels." He goes on to say that "for sonants, . . . the picture is quite different. Here the voiced is clearly the unmarked type in relation either to the laryngealized types which function as representatives of both injectives and ejective obstruents or to the voiceless type." Systems with laryngealized nasals and liquids are fairly common and may be viewed as part of a general tendency for a language which utilizes a glottalic feature to use it to a greater extent than just in its obstruent system. Thus eighteen of the languages in the UPSID have laryngealized nasals and/or liquids, eleven of these have ejective obstruents as well, three have preglottalized obstruents, one has implosives, and three have ejectives and implosives. In other words, laryngealized sonorants imply the occurrence of glottalic or laryngealized obstruents (specifically stops). The frequencies of occurrence of laryngealized nasals and liquids are as follows:

m´	n´	n´	n´	l´	r´
14	14	4	3	8	3

Table 9
Laryngealized Nasals and Liquids in the UPSID.

The distribution of laryngealized nasals can be seen to parallel that of implosive stops rather well in that both the labial and alveolar are much preferred to the back articulations, but yet are not between themselves in hierarchical relation. As with implosives, both /m´/ and /n´/ are found to occur alone (i.e., as the sole representative of their class).

The laryngealized semivowels are relatively common as well and always imply the occurrence of glottalic or laryngealized obstruents. Laryngealized vowels are also reported for three of the languages in the survey and, as expected, occur only in systems with glottalic or laryngealized obstruents. There does not, however, appear to be any hierarchical relationship among the occurrences of laryngealized sonorants, semivowels, and vowels. Any one of those three classes may occur without the occurrence of one of the others.

In terms of phonation types Greenberg (1970:9) noted "there is quite surely no phonological contrast of voicing" for the laryngealized nasals and liquids, nor, we might add, for the laryngealized semivowels and vowels. This observation follows from the fact that the open glottis associated with voicelessness by definition eliminates the possibility of any type of concurrent voicing (of which laryngealization would be one).

4.1 Hierarchical Relations to the Plain Series: Ejectives.

We have seen that ejectives never occur without plain stops in a language. Greenberg (1970:23) calls this "a classic instance of a marked/unmarked relationship" in that plain stops are "found . . . in all languages." In the case of the ejective stops, at least, we might wish to make the stronger claim that ejective stops never occur without corresponding plain stops at the same point of articulation. This claim explains certain exceptions to the point of articulation preference hierarchy for ejective stops if the ejective to plain hierarchy is accorded precedence.

The ejective to plain hierarchy can be extended to apply to affricates and fricatives as well, with some few exceptions for affricates (none for fricatives). Thus the lateral ejective affricate appears without a non-ejective voiceless counterpart in three languages - Ik (AA), Iraqw (NS), and Puget Sound (NA). Puget Sound does, however, have a plain lateral fricative which may correspond systematically to the ejective lateral affricate. In Iraqw there is also no non-ejective /ts/ corresponding to /tsʹ/. Tigre (NS) lacks non-ejective affricates to correspond to its /tsʹ/ and /tʃʹ/, and has only one non-ejective palato-alveolar affricate corresponding to two ejective ones (/tʃʹ/, /tʃʹ:/). Tolowa (NA) has both retroflexed and non-retroflexed /tṣʹ/ and /tsʹ/, but no /ts/. Finally, Otomi also has /tsʹ/ without plain /ts/ occurring. Thus, six of the forty-three languages having ejective affricates have one or more ejective affricates lackinng corresponding non-ejective voiceless affricates.

4.2 Hierarchical Relations to the Plain Series: Implosives.

As Greenberg (1970:23-24) noted, implosives do not seem to participate in a hierarchy requiring plain voiced stops corresponding to implosive ones. Generally, however, there are such correspondences and, when not, there is never an implosive series to the exclusion of a non-glottalic series of stops. Thus, out of the thirty-three implosive systems only ten have no non-implosive voiced stop corresponding to one or more implosive stops of the system. Three of these languages lack a corresponding voiced stop only for the palatal implosive: Kadugli (NK), Yulu (NS), and Ik (NS). Two of the others are maximal implosive systems; i.e., labial, alveolar, palatal, and velar, and constitute the only voiced stops series for the respective languages (Swahili, NK, and Maasai, AA) (see Greenberg 1970:24).

4.3 Hierarchical Relations to the Plain Series: Laryngealized Obstruents.

With so few laryngealized stop/affricate/fricative systems from which to generalize, only cautious statements can be made. It does appear, as expected, that voiced laryngealized obstruents tend to have corresponding non-laryngealized obstruents.

4.4 Hierarchical Relations to the Plain Series: Laryngealized Sonorants.

The ejective to plain hierarchy appears valid for laryngealized nasals, liquids, semivowels, and vowels. That is, laryngealized sonorants always imply corresponding plain sonorants. Exceptions to this generalization are quite few: Wichita (NA) is reported to have a laryngealized /r´/ without a corresponding plain /r/, while Southern Nambiquara (SA) has laryngealized /n´/, but no /n/.

4.5 Hierarchical Relations to the Plain Series: General.

We have seen that for any given language the presence of ejective stops, affricates, and fricatives implies the presence of non-ejective counterparts. Implosives, on the other hand, do not seem to imply corresponding plain voiced stops, but they do at least imply corresponding non-glottalic stops. Finally, laryngealized sonorants generally imply the presence of corresponding non-laryngealized sonorants for any given languages. The general relationship of ejective, implosive, and laryngealized patterns of occurrence with plain segments across languages presents verification of the language-dependent hierarchies we have established. Thus, in a comparison with recent work by Nartey (1979) also utilizing the UCLA Phonological Segment Inventory Database (UPSID) we find the following parallels. While Nartey (1979:4) found /s/ to be the most frequent plain fricative, our survey shows the corresponding /s´/ to be the most common ejective fricative. All glottalic fricatives are , of course, voiceless since they are all ejectives. The most common affricate in the UPSID was found to be /tʃ/ while our study shows that the most common simple ejective affricate is the corresponding /tʃ´/.

The relationships among voiceless ejective stops, implosives, and non-glottalic stops in the UPSID demonstrate the markedness of ejective systems and implosive systems in various ways. Thus, athough the most frequent stop system found by Nartey (1979:17) is that corresonding to the most frequent voiceless glottalic stop system in our survey, i.e., /p, t, k/ and /p´, t´, k´/, the most common voiceless stop overall is the alveolar /t/ rather than the velar obtaining in the ejective cases, /k´/. In terms of voicing, the implosive stop systems clearly reflect the general tendency noted by Nartey (1979:21) for voiceless primary oral stops to outnumber voiced ones.

Finally, we might note that parallelling the non-glottalic generalizations presented in Nartey (1979), in a given language the number of glottalic affricates is less than the number of glottalic stops. The ejective to plain hierarchy seems to be a more specific instance of the generalization that "a language is highly unlikely to have secondary oral stops unless it also has primary oral stops," (Nartey, 1979:24) and corresponding generalizations concerning the relationships between secondary fricatives and primary fricatives as well as between primary nasals and secondary nasals. We might also note that the predominant frequency of occurrence of /m´/ and /n´/ among the laryngealized nasals appears to be entailed by their overall frequency as plain nasals.

5.1 Genetic Distribution of Glottalic and Laryngealized Systems.

We have in section 2.1 that ejective systems are genetically fairly widespread, represented in eight of the thirteen major language families of the world: Amerind, Nilo-Saharan, Afro-Asiatic, Caucasian, Niger-Kordofanian, Indo-European, Khoisan, and Sino-Tibetan. The geographical distribution is consequently great as well. Of course, by far the greater number of ejective systems is to be found in the Amerind family (38) with the majority of those from the Northern Amerind region (30). Thus, of the fifty-one Northern Amerind languages of the UPSID, thirty contain ejective systems. Only three of these languages also exhibit implosives: Maidu, Otomi, and Mazahua. K´ekchi also reports a bilabial, voiced, laryngealized stop.

Implosive systems are found in fewer of the major language families and are more limited geographically than ejective systems. Most implosive systems (27) are to be found on the African continent. The small number of non-African languages involved (6) are mostly Amerind languages which also have ejectives (all have more ejective stops than implosive ones). Thus, implosive systems are represented by the Niger-Kordofanian (10), Nilo-Saharan (9), Afro-Asiatic (8), Amerind (4), Austro-Asiatic (1), and Sino-Tibetan (1) families.

Systems with "laryngealized" (preglottalized or laryngealized) stops and, in some cases, affricates and fricatives are found in the Ural-Altaic (1), Nilo-Saharan (1), Afro-Asiatic (1), Austro-Tai (3), Amerind (3), and Khoisan (1) language families. Thus, although laryngealized stop systems are not common, they are not merely areal phenomena. None of these systems contains ejectives nor implosives.

Laryngealized nasals, liquids, semivowels, and/or vowels are found in most of the language families mentioned above and, of course, not elsewhere. Of the major language families or groupings for which either ejective, implosive, or laryngealized obstruents are attested, only Niger-Kordofanian, Ural-Altaic, Nilo-Saharan, and Caucasian evidence no additional laryngealized sonants, semivowels, or vowels.

5.2 Diachronic Implications.

Greenberg (1970:23) suggests that "it is possible to derive the general diachronic hypothesis that at least one source of injectives might be a sound shift from voiced plain to voiced implosive stops." This is based on observation that languages with implosive stops tend to lack corresponding non-implosive voiced stops. Greenberg also suggests that loss or addition of implosives should follow the point of articulation preference hierarchy for implosives discussed in section 2.21. This would seem to suggest that implosives at the same point of articulation as non-implosive voiced stops should be rather rare. This, however, is not the finding of our survey and as noted in section 4.2, twenty-one of the thirty-three implosive systems show voiced stops corresponding to each implosive stop in the given languages. This does not refute the diachronic hypothesis, but merely the implication of non-correspondence between a plain voiced stop and an implosive stop system. Clearly in those cases for which a corresponding plain voiced stop series is attested in respect to an implosive series, we must posit a concomitant shift in another series (most likely the plain voiceless stop series) replacing the former plain voiced series which has become implosive or possibly split due to different phonetic environments later lost.

5.3 On the Phonetic Nature of Glottalic Systems.

A phonetic explanation for Greenberg's (1970) point of articulation hierarchies of ejectives and implosives has been sought by several linguists (see particularly Javkin, 1977). But while various theories have been offered in explanation, no single set of articulatory, acoustic, or perceptual factors alone seems adequate. We therefore posit that the explanation must lie in a synthesis of several interrelated factors.

We must begin by pointing out that the set of facts for which this study seeks phonetic explanation is somewhat different than that which Greenberg (1970) offered. The hierarchy posited by Greenberg for place of articulation preferences of ejective stops holds for the UPSID, but the hierarchy for place of articulation preferences of implosives does not. Whereas Greenberg (1970) reported the precedence of the labial implosive over the alveolar and Javkin (1977) offered explanation for that observation, our explanation must acknowledge the fact that the point of articulation hierarchy for implosives does not entail that the labial is the single most preferred implosive, but rather that the labial and the alveolar are equally preferred.

Presently there exists no single factor which can be cited in explanation for the tendencies we have noted in implosive and ejective distribution and production nor do we expect to discover one. On the contrary our view is that explanation can best be accomplished through a synthesis of the various factors which have recently been examined in the literature concerning unusual glottalic systems (e.g., Ladefoged, 1976; Pinkerton, 1979), the articulatory and acoustic parameters of stop production especially regarding voicing and aerodynamic variability (e.g., Brown et al., 1970; Lisker, 1970; Lubker and Parris, 1970; Malecot, 1966; Warren and Hall, 1973; among others), the neuromuscular parameters of stop production (e.g., Clumeck,1976; Perkell, 1969; Smith, 1971; among others), and the perceptual parameters of stop consonants (e.g., Miller and Eimas, 1976; Stevens and Blumstein, 1978, 1979; among others).

Javkin (1977) clarified the role played by Boyle´s Law (Boyle, 1662) in seeking an explanation of implosive and ejective distributions. He notes that Boyle´s Law entails that "in a closed chamber, pressure is inversely related to the volume of gas in that chamber." Since ejectives rely on the compression of intra-oral air, their preference for back articulations clearly is predicated at least partially upon the advantage that the smaller chambers created by back closures create. However, Javkin also notes that "it takes the same effort to produce a rarefaction or a compressionn in a chamber of the same size." Thus, since implosives rely upon the rarefaction of intra-oral air through the lowering of the larynx, such rarefaction could most efficiently be achieved by also using the smaller chambers in the oral cavity created by back articulations. But this fails to take into account the possibility that, in fact, it is easier to produce a partial vacuum (rarefaction) by increasing the volume of the chamber as much as the articulatory mechanism will allow. Thus a large chamber capable of proportionally greater expansion than a small one is preferred. Clearly the ability for expansion of the oral cavity is greatest with closure at the lips or at the dental/alveolar region. However, it has not yet been determined whether the potential for volume expansion is actually proportionally greater at the front articulations than at the velum. Thus is is possible that even a slight capability toward chamber expansion with a velar closure would be greater in relation to the small size of the chamber than the expansion capabilities for front articulations are in relation to the obtaining chamber volumes. This is simply an empirical question which has not been determined in the literature to date although Javkin´ (1977) model clearly suggests that indeed the proportion of the ability to change volume to the overall volume of the chamber is greater for the velar closure than for either the labial or alveolar closures. His explanation therefore must call upon voicing as part of the explanation rather than as a factor itself in need of explanation. Thus it is posited that the inflow of pulmonic air due to the lowering of the larynx in implosive production offsets the ability of the velar closure to effect a change in overall chamber volume consequently rendering the otherwise preferable back articulations less preferable than those in the front on terms of their ability to aid rarefaction.

As an alternative to this explanation for ejective and implosive distributions we propose that indeed the preference for back articulations of ejectives is primarily a function of the ability of the rising larynx to effect greater supraglottal pressure given the smaller chambers afforded by back closures. Likewise we posit that the ability of a smaller chamber to produce rarefaction more easily than a larger chamber is outweighed by the greater ability to actively expand chamber size at front articulations thus resulting in the preference for front articulations for implosives. It is interesting to note that such active expansion of chamber size has commonly been cited as a possible mechanism in the achievement of voicing in plain stops, particularly through the expansion of the pharyngeal walls (see Kent and Moll, 1969, and Smith, 1971, aong others). This brings us to the conclusion that the most salient cue in implosive perception is most likely that of voicing rather than the actual implosion, for as Greenberg (1970) and Ladefoged (1968) noted, implosives do not always entail inward air flow upon release. We posit that implosion, or the lowering of the larynx associated with it, is pedominantly a mechanism employed in the more general aim of achieving salient voicing in stop production. It is the active expansion of the oral cavity through lowering of the larynx and through jaw lowering (in the case of labials) or retroflection and apical production (in the case of alveolars; see particularly Hardcastle and Brasington, 1978, on this latter point) that allows and, in fact, forces the pulmonic subglottal air through the vocal folds to effect significant, maintainable voicing. Other work currently in progress (Fordyce and Lindau-Webb, forthcoming) demonstrates that within various Ijo dialects containing implosive as well as plain voiced stops at corresponding places of articulation, the implosives generally are able to maintain voicing throughout stop closure while the plain voiced stops tend to evidence irregularities and decay in voicing characteristics during the closure phase. The back articulations, however, cannot produce a significantly low intra-oral pressure, even with the lowering of the larynx, to allow sufficient air flow through the glottis to achieve voicing. Thus in those systems for which implosives are utilized to render more salient a voicing distinction between or among stop series, back implosives would not aid in the endeavour and would basically serve no purpose. However, it is possible that in some language systems, particularly those without overlapping implosive and ejective series, that it is the glottalic feature that serves as the most salient cue in the perception of the given series. In any case, we would not expect to find them outside of systems in which the glottalic feature seemed more salient. These predictions are borne out by the data presented in Pinkerton (1979) showing the occurrence of a voiceless uvular implosive as part of a conflated glottalic series in five Quichean (Mayan) languages. We would also posit that other instances of voiceless implosive such as that reported by Ladefoged (1976) for Owerri Igbo are also obviously not a function of increasing the saliency of voicing distinctions in the language, but entail the modification of the oral chamber in such a way as to actively prevent voicing. This observation seems to be borne out by Ladefoged's note that the voiceless implosives in Igbo tend to be velarized, thus decreasing rather than increasing the oral cavity volume. The Igbo example forces us to conclude that in the general case which presents itself as a system with either an ejective series, an implosive series, or both (overlapping) then their presence is predicated upon a maximalization of voicing distinctions. But in the case which presents itself as non-overlapping ejective and implosive series, the system then serves the function of maximalizing the glottalic feature. And, finally, in cases which present themselves as overlapping voiced and voiceless implosive series (as in Igbo) their presence, while originally motivated by a maximalization of salient voicing principle, have become fossilized in an opaque manner. It seems possible that the Igbo case is actually one of regularization with regard to the rest of the Igbo stop system of what may originally have been a non-overlapping implosive/ejective series such that the velarized /p</ is a reflex of an earlier ejective /k´/, /t</ simply changed from

ejective to implosive, and /b/ remained implosive, thus creating an implosive system with contrastive voicing parallel to the already present plain and aspirated (breathy voice) systems each with contrastive voice. In other words, pattern pressure forced the glottalic series to split in terms of voicing and to reflect the other two stop subsystems in place of articulation realizations insofar as was possible with only three glottalics. The voiceless ejectives in this scenario became implosive in order to mimic as closely as possible the phonation type of the voiced implosives. The Igbo case notwithstanding, our explanation for the distribution of ejectives and implosives in terms of place of articualtion appears to be capable of explaining their voicing characteristics as well and the way that such factors may interact within a given system.

Underlying both place of articulation preference hierarchies for ejectives and implosives respectively is the observation that the palatal and uvular places of articulation are rather limited in frequency of occurrences of either glottalic type. This is part of a more widespread tendency toward disfavoring palatal and uvular articulations for stop consonants in general. For instance, the number of primary palatal or uvular stop consonants in the languages of the UPSID is less than a sixth of the number of labial, alveolar, or velar stops in both the voiced and voiceless series (see Nartey, 1979). This seems to be due to several factors. The most probable reason is that closure at the palatal region entails significantly more neromuscular effort (in relation to labial or alveolar closure) since it requires higher lift of the tongue body and greater force of occlusion since the area covered is by far the largest obtaining in any stop production (see Clumeck, 1976; Lubker, 1970; nihilani, 1975; Smith, 1971, aong others). This in itself might account for the disfavoring of palatals. It is also possible that these distributional limitations are related to observations made by Stevens and Blumstein (1978, 1979) that stable acoustic patterns are produced by only a limited set of articulatory configurations. Their data indicate that information signalling place of articulation for stop consonants appears to reside in the initial 10 - 20 msec following the release of the stop. But coarticulation effects which are undoubtedly present at least for palatal stop productions could conceivably interfere with place of articulation identification particularly if based upon release configurations of the stops. Since the Stevens and Blumstein data do not specifically demonstrate that place information is signalled in the first 10 - 20 msec of the release for palatals (nor for uvulars) our evidence for such an analysis remains indirect.

Another factor in the disfavoring of palatal and uvular articulations of stops has to do with the salience of front vs. back cavity distinctions in the realization of acoustic cues such as compactness, gravity, and flattening (see Fant, 1960:218). Thus palatals present an instance in which the front/back cavities as defined by the closure are not well differentiated. Lubker´s (1970) study seems to suggest that the extreme gesture required in the articulation of palatal stops would produce unacceptable coarticulation effects on the adjacent vowels in terms of their perceptual viability. For uvular stop articulations a number of other factors conspire to lead to the overall disfavoring of uvular stops.

Thus, following upon Fant's (1960) observations it seems probable that uvular closure provides too large a front cavity to produce the acoustic cues salient to the features of compactness, gravity, and flattening. Or, in relation to tense/lax, that uvular articulations (along with palatals) present the most deformed positions of occlusion of the vocal tract from the neutral position with regard to the tongue. Furthermore, data from Smith (1971:17) seem to indicate that in uvular stop production there are antagonistic muscular activities involved such that as genioglossus activity is brought to bear in the lifting of the tongue back toward closure with the uvula, glossopharyngeal activity is simultaneaously moving the tongue body back in antagonistic fashion toward the upward movement. This seems to follow from the general observation that tongue movement is neuromuscularly constrained in such a way as to allow backing articulations involving the styloglossus more toward the velum or toward pharyngeal constrictions without the potential for direct movement toward the uvula (see Ladefoged, 1979). Thus, palatal and uvular stops entail relatively less acceptable neuromuscular complexity and perceptual complication following from it as compared to stop production at the labial, alveolar, or velar positions, those most common in glottalic inventories.

Appendix: Universal Tendencies Restated.

1. A language having the bilabial ejective stop also has the alveolar ejective stop and the velar ejective stop; a language having the alveolar ejective stop will also have the velar ejective stop; and a language having only one ejective stop has the velar. (Subject to #15 below.)

   Exceptions: Lelemi, Berta, Hupa, Jivaro.

2. A language having the uvular ejective stop also has the bilabial, alveolar, and velar; a language having the palatal ejective stop also has the uvular, bilabial, alveolar, and velar. (Subject to #15 below.)

   Exceptions: Kefa, Haida, Hupa, Maidu.

3. A language having the velar implosive also has the palatal, alveolar, and labial; a language having the palatal implosive also has the alveolar and 1.

   Exception: Hamer.

4. A language having ejective affricates also has ejective stops.

   Exception: Iraqw.

5. If a language has /tl'/ it also has /ts'/ and /tʃ'/. (Subject to #16 below.)

   Exceptions: Ik, Haida, Wintu.

6. A language having ejective fricatives also has ejective stops.

7. A language having ejective fricatives also has ejective affricates.

   Exceptions: Berta, Koma, Socotri.

8. If a language has laryngealized sonorants, it also has laryngealized or glottalic obstruents.

   Exception: Tiddim.

9. If a language has only one ejective fricative it will be /s´/ or / ´/.

10. Labialization as a secondary articulation feature occurs only on the velar and/or uvular ejective stops.

11. If a language has a labialzed ejective affricate, it also has a corresponding non-labialized ejective affricate.

12. If a language has a labialized ejective fricative, it also has a corresponding non-labialized ejective fricative.

13. If a language has labialized ejective affricates or fricatives, it also has labialized ejective stops.

14. If a language has a labialized ejective stop, it also has a corresponding non-labialized ejective stop.

    Exception: Kabardian.

15. If a language has an ejective stop, it also has a corresponding non-ejective stop.

    Exceptions: Berta, Nez Perce.

16. If a language has a non-lateral ejective affricate, it also has a corresponding plain affricate.

    Exceptions: Tigre, Iraqw, Tolowa, Otomi.

17. If a language has an ejective fricative, it also has a corresponding plain fricative.

18. If a language has a laryngealized sonorant, it also has a corresponding plain sonorant.

    Exceptions: Wichita, Southern Nambiquara. (Both involve laryngealized nasals or liquids.)

19. (a) Ejectives are voiceless.
    (b) Implosvies are voiced. Exception: Igbo.

20. Laryngealized obstruents do not occur in languages with glottalic consonants.

21. Laryngealized sonorants (nasals, liquids, semivowels, vowels) are voiced.

22. The most typical implosive system consists of two implosive stops: labial and alveolar.

23. The most typical ejective system consists of three ejective stops: labial, alveolar, velar.

24. If a language has only one ejective affricate, it will be /ts´/ or /t ´/.

25. If a language has only one implosive stop, it will be /ɓ/ or /ɗ/.

26. If a language has only one ejective stop, it will be /k´/.

    Exceptions: Lelemi, Jivaro.

27. If a language has only one laryngealized nasal, it will be /m´/ or /n´/.

    Exception: Southern Nambiquára.

28. The most frequent ejective affricate is the alveopalatal /tʃ´/. The next most frequent is /ts´/.

29. The most frequent ejective fricative is the alveolar /s´/. The next most frequent is /ʃ´/.

30. The most frequent ejective stop is /k´/. The next most frequent ejective is /t´/.

31. The most frequent implosive stop is /ɗ/. The next most frequent is /ɓ/.

32. The most frequent laryngealized sonorant is /m´/. The next most frequent laryngealized sonorant is /n´/.

33. If a language has no ejective affricates, then only /s´/ or /ʃ´/ (or both) may occur as ejective fricatives.

Notes.

1. The UPSID was established primarily through the efforts of Professor Ian Maddieson. Assistance with coding and research was provided by myself and my collegues Sandra F. Disner, Vivian Flores, Jonas N. A. Nartey, G. Diane Ridley, and Vincent van Heuven. Assitance with research was aided by Stephen Franks, Bonnie Glover, Peter Ladefoged, Mona Lindau-Webb, Robin Thurman, Anne H. Wingate, W. Andreas Wittenstein, and Eric Zee.

2. The label "Alveolar" in this table includes both the dental and alveolar places of articulation. It is the policy of the UPSID to specify these places of articulation in cases where it is possible to do so. All other cases are coded as "unspecified dental/alveolar" (UDA). Most commonly for ejective stops, the actual place of articulation is not discernable in the literature source and has been coded in the UPSID as UDA. The implication that this conflation is not contrastive must not be made, however, since ejective stops contrasting at the dental and alveolar places of articulation are claimed for two languages in the survey - Pomo (NA) and Wappo (NA).

3. The two letter abbreviations signify the family or major subgrouping to which the language belongs according to the following key:
   IE Indo-European
   UA Ural-Altaic
   NK Niger-Kordofanian
   NS Nilo-Saharan
   AA Afro-Asiatic
   AS Austro-Tai
   ST Sino-Tibetan
   NA Northern Amerind
   SA Southern Amerind
   CA Caucasian
   KH Khoisan

4. Pinkerton's (1978) more recent work, however, shows this to be an implosive rather than a plosive with a "voiced, non-glottalized variant."

5. This includes "unspecified dental/alveolar" as per note 2.

References

Boyle, R. (1662) "A Defence of the Doctrine Tounching the Spring and Weight of the Air." In The Works of the Honorable Robert Boyle, ed., Thomas Birch. London, (1744), Volume 1.

Brown, W. S., R. E. McGlone, A. Tarlow, and T. Shipp. (1970) "Intraoral Air Pressure Associated with Specific Phonetic Positions." Phonetica 22:202-212.

Campbell, Lyle. (1973) "On Glottalic Consonants." International Journal of American Linguistics 39:44-46.

Catford, J. (1939) "On the Classification of Stop Consonants." Le Maitre Phonetique, 3rd series, 65:2-5.

Clumeck, Harold. (1976) "Patterns of Soft Palate Movements in Six Languages." Journal of Phonetics 4:337-351.

Fant, Gunnar. (1960) Acoustic Theory of Speech Production. The Hague: Mouton.

Fordyce, James F. and Mona Lindau-Webb. (forthcoming) "Phonetic Differences Among Implosives."

Gamkrelidze, Thomas V. (1978) "On the Correlation of Stops and Fricatives in a Phonological System." In Joseph H. Greenberg, ed. Universals of Human Language, Volume 2, Phonology. Stanford: Stanford University Press.

Hardcastle, W. J. and R. W. P. Brasington. (1978) "Experimental Study of Implosive and Voiced Egressive Stops in Shona: an Interim Report." Work in Progress 2, Phonetics Laboratory, University of Reading, 66-97.

Haudricourt, A. G. (1950) "Les Consonnes Preglottalisees en Indochine." Bulletin de la Societe Linguistique de Paris 46:172-182.

Hoard, James E. (1978) "Obstruent Voicing in Gitksan: Some Implications for Distinctive Feature Theory." In E.-D. Cook and J. Kaye, eds., Linguistics Studies of Native Canada. Vancouver: University of British Columbia Press.

Javkin, Hector. (1977) "Towards a Phonetic Explanation for Universal Preferences in Implosives and Ejectives." Berkeley Linguistic Society III:559-565.

Kent, R. D. and K. L. Moll. (1969) "Vocal Tract Characteristics of the Stop Cognates." Journal of the Acoustical Society of America 46:1549-1555.

Ladefoged, Peter. (1968) A Phonetic Study of West African Languages. Cambridge: Cambridge University Press.

Ladefoged, Peter. (1976) "The Stops of Owerri Igbo." Studies in African Linguistics, Supplement 6, 147-164.

Ladefoged, Peter. (1979) "Articulatory Parameters." UCLA Working Papers in Phonetics 45:25-31.

Lisker, L. (1970) "Supraglottal Air Pressure in the Production of English Stops." Language and Speech 13:215-231.

Lubker, J. F. (1970) "Aerodynamic and Ultrasonic Assessment Techniques in Speech Dento-facial Research." ASHA Reports 5:207-223.

Lubker, J. F. and P. J. Parris. (1970) "Simultaneous Measurements of Intraoral Pressure, Force of Labial Contact, and Labial Electromyographic Activity During the Production of the Stop Cognates p and b." Journal of the Acoustic Society of America 47:625-633.

Malecot, A. (1966) "The Effectiveness of Intraoral Air-pressure-pulse Parameters in Distinguishing Between Stop Cognates." Phonetica 14:65-81.

Miller, J. M. and P. D. Eimas. (1976) "Studies on the Selective Tuning of Feature Detectors for Speech." Journal of Phonetics 4:119-127.

Nartey, Jonas N. A. (1979) A Phonetic Study in Phonemic Universals. UCLA Working Papers in Phonetics 46. Los Angeles: University of California.

Nihilani, P. (1974) "An Aerodynamic Study of Stops in Sindhi." Phonetica 29:193-224.

Perkell, Joseph S. (1969) Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study. Cambridge: Mass.: M.I.T. Press.

Pinkerton, Sandra. (1979) "Quichean (Mayan) Glottalized and Non-glottalized Stops: A Phonetic Study with Implications for Phonological Universals." Manuscript. Phonology Laboratory, Department of Linguistics, University of California, Berkeley, California.

Smith, Timothy S. (1971) A Phonetic Study of the Function of the Extrinsic Tongue Muscles. UCLA Working Papers in Phonetics 18. Los Angeles: University of California.

Snyman, J. W. (1969) An Introduction to the !Xũ (!Kung) Language. Cape Town: A. A. Balkema.

Stevens, K. N. and S. E. Blumstein. (1978) "Invariant Cues for Place of Articulation in Stop Consonants." Journal of the Acoustical Society of America 64:1358-1368.

Warren, D. W. and D. J. Hall. (1973) "Glottal Activity and Intraoral Pressure and Rate of Flow During Speech." Folia Phoniatrica 25:121-129.