

## Patterns in allophone distribution for voiced and voiceless stops

Patricia Keating, Wendy Linker and Marie Huffman

*Phonetics Laboratory, Department of Linguistics, University of California, Los Angeles, California 90024, USA*

*Received 9th April 1983*

### Abstract:

Do languages prefer certain types of stop consonants in certain environments? A survey of 51 languages reveals two general trends. First, many languages limit the occurrence of final stops; final devoicing is one instance of this phenomenon. Second, other variation is largely limited to languages whose contrasts in initial position involve short lag vs. long lag VOT; that is, aspiration contrasts. Among these languages, many sorts of variation are found, which therefore must be described by language-specific rules.

It has often been observed that languages prefer voiceless over voiced consonants. At the same time, however, it is generally thought that intervocalic consonants will most naturally be voiced. If this is indeed the case, then we should see either consistent patterns of contrast neutralization, or of allophonic variation, with voiceless consonants being found everywhere except intervocally. Work with an aerodynamic model (Westbury & Keating, 1980, and further work) leads one to expect just that: voiceless unaspirated initial and final stops but voiced medial stops. Nonetheless, data from natural languages does not appear to lend overwhelming support to these predictions, particularly in terms of medial voicing (Houlihan 1982). It became clear to us that better data on positional variation in consonant voicing was required to compare with results from physiological modeling studies. Therefore we began a survey of the allophones of voiced and voiceless stop phonemes in various languages.

### Allophone survey

The languages surveyed include those in Lisker & Abramson's (1964) study of word-initial VOT values, and the various languages whose initial VOT values have been similarly determined since then. These languages form less than half the database. Additional languages were added to increase the overall variety, both genetic and phonetic, of the database, and to provide sufficient examples of phenomena of particular interest. All stop consonants, including glottalic segments, were included in the survey. The data consist primarily of state-ments about allophone and phoneme distributions, largely derived from impressionistic transcriptions used for classical phonemicizations. In many cases the sources consulted are those used in compiling the UCLA Phonological Segment Inventory Database (UPSID, Maddieson 1981), a listing of the phonetic feature description of one principle allophone for each of the surface phonemes in 317 languages. UPSID was designed primarily to abstract away from the kind of positional variation that we are interested in, but its sources were often well suited to answering our questions. However, just as often such sources do not

consider a sufficiently wide range of environments, and so where possible we supplemented information from the literature with our own or others' acoustic measurements, and with information from UCLA student phonetic projects and field notes. The major sources consulted and the allophones that were found are described for each language in the Appendix. Here we will describe the relevant trends observed across these data.

A major, though expected, trend across languages is the overwhelming preference for voiceless unaspirated stops. All of the languages surveyed except Bobo, Breton, and possibly Yidjɔ̃ use at least this category.\* Languages with a [voice] contrast generally use the voiceless unaspirated category in opposition to one or more other categories; exceptions include various languages with a [b] but no [p]. Beyond this preference, variation across positions in a word is also found. In the discussion that follows, we will focus on two kinds of variation: first, restrictions on final consonants, which are found in languages with all kinds of voicing contrasts, and restrictions on medial consonants, which hold largely of languages with initial aspiration contrasts.

Languages limit the consonants that can occur in final position to a greater or lesser extent. Some languages simply disallow final consonants, at least phonetically: Alyawarra (with a consistent final [a]), Bobo, Kaititj, Kikuyu, Tiwi, marginally Hawaiian, with a final [ʔ], and Swahili, with final consonants only in loans. Other languages restrict final consonants to sonorants of one sort or another: Akan, Japanese, Mandarin, Tamil, Tswana, and Yidjɔ̃. Hausa allows final sonorants, and rarely [s], as well as having phonetic [ʔ] after short vowels. Spanish limits its final consonants to continuants. More common, however, is that stops are allowed but restricted in their manner. So, for example, Tagalog typically has unreleased final stops; English and Choctaw have optionally unreleased final stops. Burmese allows final nasals and a phonemic glottal stop /ʔ/ which in phrases may have various assimilatory phonetic realizations. Probably the best known type of final stop restriction is that final stops must be voiceless: Basque, Bulgarian, Cantonese, Choctaw, Dutch, Efik, Ewondo, Finnish, Gaelic, German, Polish, Russian, and Zoque. The voiceless stop restriction may, but need not, involve a synchronic phonological rule of final devoicing. It should be noted, however, that the phonological domain varies: in German it is syllable-final stops that are voiceless, while in Polish it is phrase-final stops. (Word-final but phrase-medial Polish voiceless stops can actually voice in some dialects (Mikoś, 1977).) A few languages have only unreleased *and* voiceless final stops: Korean, Nama, Thai, Tikar, Vietnamese. Thus it can be seen that "final devoicing" is just part of a larger trend to fewer segments and contrasts in final position. Other languages surveyed either have the same segments finally as initially and medially, or variation that involves, e.g., spirantization as part of the phonological contrast. Note that languages with four contrasting categories, e.g. Bengali, Hindi, Marathi, generally allow them all in final position, although not always in colloquial speech; Armenian and Tzeltal with three categories also allow them all in final position. Thus, while there is clearly a trend towards fewer contrasts in final position (involving other features such as place of articulation as well), and some variation in the allophones of the contrasting phonemes, it is certainly possible to sustain as many contrasts finally as elsewhere.

The situation for medial position is quite different, with variation being largely limited

\*UPSID does contain a few examples of languages with voiced but no voiceless stops, but they were not included in our sample. Such descriptions are open to reinterpretation in light of the typical transcription of short lag VOT stops (voiceless unaspirated) as voiced, following English. The description of Yidjɔ̃ makes such a reinterpretation likely but not certain.

to one particular kind. Recall that one motivation for this study was an interest in the occurrence of medial voiced stops. If phonetically voiced stops are found in initial position, it is almost always in contrast with some other category. In this case, a "preference" for voiced stops in medial position would have to involve neutralization of the contrast to a voiced variant. As noted before by Houlihan & Iverson (1979), languages typically do not have neutralization in medial position. On the other hand, if voiced stops do not occur in initial position, they could be found in medial position through allophonic variation in some category that does occur initially. First, a language with no voicing contrast could have voiced stops medially but voiceless stops elsewhere. We would in fact expect this to be the general pattern for languages without a contrast. Second, a language with an initial contrast between phonetically voiced and voiceless unaspirated stops could have a rule that changed the medial [+voice] stops into, e.g., fricatives, and [-voice] stops into voiced stops. Third, a language with an initial contrast between voiceless unaspirated and aspirated stops (e.g. most Germanic and Chinese languages) could have a rule that voiced one of those categories in medial position. Let us examine the data for each of these possible sources of medial voiced stops.

A reasonable hypothesis would be that languages without a [voice] contrast should show more allophonic variation than languages with a contrast, since a contrast might be thought to constrain possible variation within the phonetic space. Of the six languages without a contrast in our sample, only one is described as having a phonetic difference between initial and medial position: Yidjip appears to have systematic medial voicing, as predicted. However, none of the other five are described in this way. Spectrograms from one speaker of each language indicate that in Hawaiian medial stops are less aspirated than initial stops though they may have noise during the closure, while in Alyawarra (an Arandic language of Australia) medial stops are if anything slightly more aspirated than initial stops. Kaititj (related to Alyawarra) shows no systematic differences, and Tiwi (also in Australia) is explicitly described as having no intervocalic voicing. Nama (a Khoisan language of Africa) voices /t/ and /k/ after rising tones, but these segments only occur word-initially; labials, which do occur medially and finally as well, do not undergo this change. Overall, then, these languages show surprisingly little allophonic variation, and what there is, is not consistent across the six languages.

To this number can be added four languages with rather marginal voicing contrasts. Finnish's only voiced stop, /d/, occurs only medially, and even then only as a product of gradation rules. In addition, Suomi (1980) describes Finnish /p t k/ as often "voiced" intervocalically, but his own measurements show only a small amount of closure voicing. Next, the contrast in Kikuyu involves prenasalization as well as voicing, and there is no difference described between initial and medial positions. Third, in Cuna the initial voiceless unaspirated [p t k] are not part of a [voice] contrast, but in medial position they become voiced and contrast with voiceless geminates [p : t : k:]. Similarly, Tamil has no initial contrast but a medial contrast of short voiced stops (or fricatives) and geminate voiceless stops. Lastly, Choctaw's [voice] contrast is found only for labials; for all of its places of articulation, medial stops before unstressed vowels can be voiced. Again, then, there is some evidence for a preference for medial voiced stops, but no universal trend.

Languages with three or more contrasting categories in initial position often have the same categories in medial position (Bengali, Hausa, Ibibio, Thai). However, Ibibio is a classical example of allophonic medial voicing, with initial [p t k] but medial [b d g]. E. Armenian is described as having a slight phonetic difference, with /b d g/ being more voiced medially, and possibly finally, than initially. Tatar has allophonic prenasalization,

and Tzeltal has allophonic spirantization, of its medial [+voice] stops, with the other categories unchanged. Interestingly, Swahili is described as being in the process of eliminating its aspiration contrast by associating aspiration with initial position and with stress, a correlation that will re-appear below. Finally, word-initial voiceless unaspirated and aspirated stops in Burmese can become voiced in phrases, especially after a "weakened" (toneless and reduced) syllable. The initial consonant in the weakened syllable is then itself eligible for voicing. That is, Burmese voicing is a phrasal rule, somewhat like Polish voicing of word-final stops before a vowel-initiated word.

Languages with two contrasting categories in initial position, one voiced and one voiceless, and neither of which is aspirated, are languages with traditional "voicing" contrasts (Arabic, Basque, Bulgarian, Dutch, Efik, Ewondo, French, Japanese, Polish, Russian, Spanish, Tagalog, Zoque). Their VOT values in initial position, where ascertained, are seen to be lead vs short lag (Yeni-Kornshian, Caramazza & Preston, 1977; Lisker & Abramson, 1964; Caramazza & Yeni-Kornshian, 1974; Shimizu, 1979; Keating, Mikos & Ganong 1981; unpublished UCLA data). Of these languages, only Efik, Ewondo and Spanish show any real differences between initial and medial positions, and these differences do not just involve voicing. Efik medial alveolars become flaps, and Ewondo [+voice] stops may become continuants, as they do in Spanish. However, even in these cases where the [+voice] stops become something else, the [-voice] stops remain voiceless and unaspirated (cf. Dent, 1976, on Spanish). Overall, there is little variation across these positions in these languages, and no tendency towards medial voicing, devoicing, or aspiration.

In contrast, languages whose initial stop contrast involves aspiration have been found to contrast short lag with long lag VOT values (Lisker & Abramson, 1964; Fischer-Jørgensen, 1954; Keating, 1983). These languages generally do show differences between initial and medial position. The most common pattern across languages involves medial deaspiration of initial voiceless aspirated stops, and/or medial voicing of initial voiceless unaspirated stops. Degree of stress is often implicated in descriptions of this pattern: English, Gaelic, Hindi, Kirghiz, Mandarin, Swedish, and sometimes German, are described as having unaspirated [-voice] stops before unstressed vowels. Mandarin, and sometimes German, are described as having voiced [+voice] stops before unstressed vowels. Choctaw, English, Kirghiz, Norwegian, Persian, Swedish, and Yidjin are all described as having voiced [+voice] medial stops, but not specifically before unstressed vowels. Danish treats medial stops differently, with spirantization of the [-voice] stops and voicing of the [-voice] stops. Akan and Cantonese are exceptions in that they appear to have no variation across positions.

In many of these cases, a sufficiently wide range of environments was not covered in the sources for us to be sure of the importance of stress as a conditioning factor. To clarify just two such cases, we collected acoustic data on English and Swedish, both languages with initial aspiration contrasts. While these are both Germanic languages, we emphasize that the phenomena of interest are not limited to such languages, as can be seen from the examples just given.

#### Acoustic data

Because a correlation between stress and phonetic detail was strongly indicated, we systematically varied both position in the word, that is, initial vs medial, and stress, that is, main stress vs other degrees of stress on a following vowel, for stops in each language. From half a dozen real words were used for each position + stress condition. All stops occurred before low vowels, but the vowel before each stop was not controlled. Words were read in isolation, so all initial stops were post-pausal. Six speakers of each language read their list once in a

Figur

sou  
des  
to

wa

no

of

cc

fr

ol

ir

p

T

v

s

t

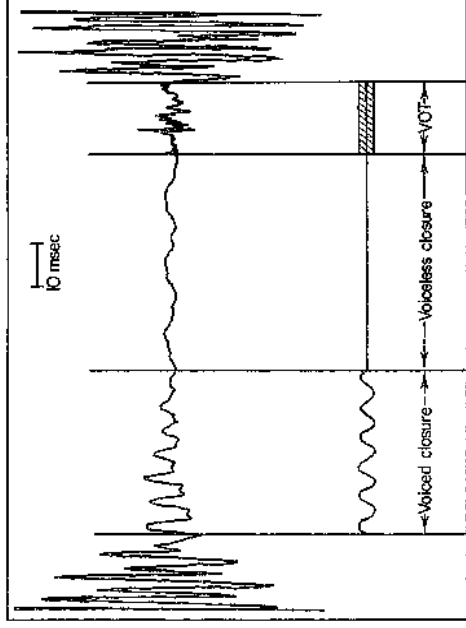


Figure 1 Acoustic measurements made for English and Swedish stops.

sound-treated booth. The items were not randomized, so as to make implicitly clear the desired stress contour. However, neither the experimental goals nor the lists were described to or discussed with the speakers.

Three measurements were made for each stop, shown in Figure 1. The top picture is a real waveform of the sort measured in a computer display; it includes some typical low-frequency noise that must be ignored. The point of closure in the waveform displays was determined on the basis of both a sudden decrease in amplitude, and a loss of frequency components corresponding to F2 and higher. Closure voicing containing energy at the fundamental frequency and sometimes F1 typically continued past the point of closure for at least one or two pitch periods. The moments of stop release and of voicing onset were determined in the usual ways. Below the waveform is a schematic version of the type we will use in presenting our results, and on the bottom is an indication of the measurements we made. The first was the duration of any voicing during closure, the second was the duration of voicelessness during closure, where that could be measured from an acoustic display. The sum of these two measures is the total closure duration. The third measure was the duration of voicelessness after the release, that is, lag VOT.

Results for English are shown in Figure 2. This experiment is similar to others in the literature (e.g. Lisker & Abramson, 1964, 1967; Flège & Brown, 1982), but it has the advantage of using the same speakers in a variety of contexts. For our speakers, while initial /b d g/ are generally voiceless and unaspirated, some speakers sometimes prevoice. Second, medial /p t k/ before reduced vowels generally have much lower VOT's, that is, are unaspirated. At the same time, medial /b d g/ are typically voiced, no matter what stress the following vowel has. However, for some speakers there is proportionately less voicing for /b/ (a surprising result) before a syllable with main stress.

Results for Swedish are shown in Figure 3. The categories we tested were divided according to word accents, which require some explanation. A Swedish word has one of two pitch accents, which in the Stockholm dialect are markedly different from each other. Accent 1 words have one pitch peak, while Accent 2 words have two pitch peaks, one on each syllable. Only Accent 1 words can vary in where the stress falls; traditionally Accent 2 words are described as having primary stress on the first syllable and secondary stress on the second syllable. Therefore we separated Accent 1 from Accent 2 words in our list. Since our results

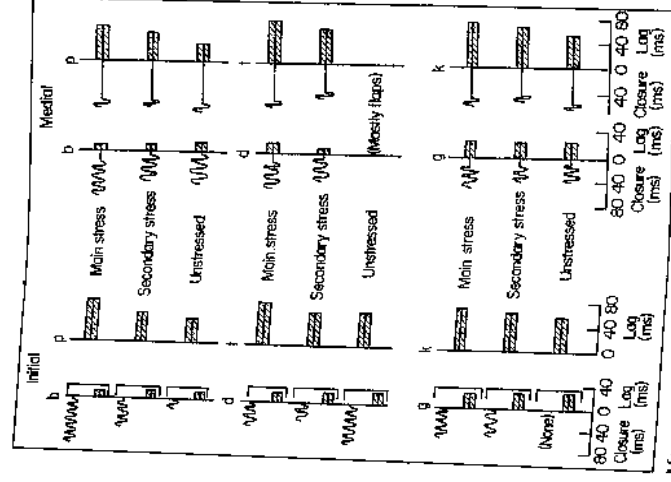


Figure 2

Mean acoustic measurements for English stops of six speakers. Initial b, d, g are divided into prevoiced and short lag groups.

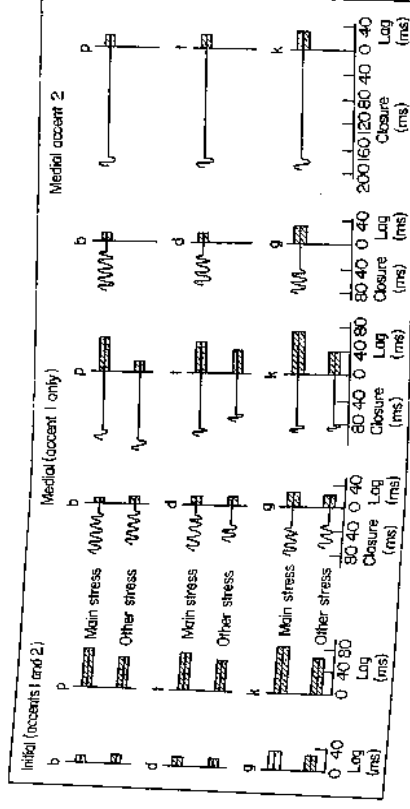


Figure 3

Mean acoustic measurements for Swedish stops of six speakers.

show that as far as VOT is concerned, stops in Accent 2 words look like the corresponding unstressed stops in Accent 1 words, we have collapsed these categories for initial stops. The closure voicing of these categories is also the same. However, since stops in Accent 2 words have much longer voiceless intervals during their closures, we have not collapsed these categories in our presentation of the medial data. Medial Accent 2 data are shown separately in Figure 3, although in their VOT values they behave like unstressed Accent 1 stops.

To summarize the overall Swedish results, initial /b d g/ do not have closure voicing while medial /p t k/ exhibit main stress. However, other medial /p t k/ are separated, as are these data are like those of Löfqvist (1976) in which four speakers read nonsense words containing the stop consonants combined with /a/ in various stress and position contexts.

Our interest here has been grosser, categorical changes of the sort that might be noted in an impressionistic transcription and codified in a phonetic alphabet. However, various other effects of these variables on these measures have been the objects of others' investigations (e.g. Zue, 1976; Flege & Brown 1982). To some extent the patterns noted in those studies are seen here as well, but not always. However, we cannot expect to see effects on closure measures clearly, since the vowel preceding medial stops in our lists was not controlled. Nonetheless, we note briefly the effects evident in our data. For English, we see that stress increases the VOT for /p t k/; that place of articulation affects VOT, particularly for /b d g/. Place of articulation also affects the duration of closure voicing and the total closure duration; to a lesser extent the voicelessness measures, closure voicelessness and total voicelessness. Stress affects the duration of closure voicelessness for /b d g/, and of total voicelessness for all stops. For Swedish, the VOT measures are affected as for English. Also as for English, place of articulation affects the duration of closure voicing and for /b d g/ of total voicelessness. Stress affects the duration of closure voicing, for /b d g/ total closure, and for /p t k/ total voicelessness.

#### Discussion

To summarize, there is a clear overall trend in these languages to maintain aspirated /p t k/ before more stressed vowels, and to voice medial /b d g/. At the same time, there are differences between these two quite similar languages. While prevoicing is seen in English, it is at best rare in Swedish. There is also more variation across speakers in English regarding amount of medial voicing. In addition, there are differences with regard to closure durations, especially as conditioned by the pitch accents in Swedish.

Our language survey suggested a generalization that languages with initial "aspiration" contrasts account for most of the allophonic differences seen between initial and medial stops. This variation appeared to be correlated with stress. In the two languages considered here in some detail, there is remarkable agreement in how stops vary across contexts, with position determining /b d g/ allophones, and position plus stress determining /p t k/ allophones. To the extent that this pattern holds across a variety of languages, and is not due to the genetic relation between English and Swedish, then it can be attributed to some general phonetic principle. The small-scale effects of place of articulation and stress that were observed may also reflect some universal pattern. Variation due to universal principles need not be accounted for in the grammar of any one language; in further work we can attempt to elucidate these principles, e.g. by aerodynamic modeling.

However, there are also subtle differences between the two languages, in terms of category use, exact VOT values, amount of variation across speakers, and closure durations. Such quantitative differences apparently can not be attributed to universal principles, and, as part of speakers' knowledge of their language, must be represented somewhere in their grammars. That is, the very lowest level of the grammar must contain language-specific quantitative rules. Our premise here is that we can use cross-linguistic comparisons to determine which aspects of variation within a language are under language-specific control. This paper is intended as a first step in the direction of teasing apart the contributions of language-specific phonological contrasts, their phonetic implementations, and universal phonetic patterns to the observed acoustic structure of voiced and voiceless stops.

This work was supported by UCLA and by an NSF grant to Peter Ladefoged. Earlier versions of this paper were presented at the 1982 Annual meeting of the LSA in San Diego and at the Conference on Causality and Language Change, UCLA, May 1982. The latter appears in



inding  
The  
words  
these  
ately  
Alle  
are  
ted.  
ords  
x(s).

WPP 54. We thank Don Brusasco and Ellen Jackson for help in data gathering and analysis, Jane Simpson for Australian language data sources, and Kathy Houlihan, Peter Ladefoged, and Ian Maddieson for comments.

### References

- Abdel-Massih, E. T. (1975). *An Introduction to Egyptian Arabic*. Ann Arbor, Michigan: Center for Near Eastern and North African Studies, University of Michigan.
- Allen, W. S. (1950). Notes on the phonetics of an Eastern Armenian speaker. *Transactions of the Philological Society*, 180-206.
- Armstrong, L. E. (1932). *The Phonetics of French*. Twelfth printing, 1978, London: Bell.
- Armstrong, L. E. (1940). *The Phonetic and Tonal Structure of Kikuyu*. London: Oxford University Press.
- Aronson, H. I. (1968). *Bulgarian Inflectional Morphophonology*. Slavistic Printings and Reprintings 73. The Hague: Mouton.
- Arun, V. B. (1961). *A Comparative Phonology of Hindi and Punjabi*. Hoshiarpur, Punjab: VVRI Press.
- Azuke, R. M. (1949). *El Vasconce*. Bilbao: Junta de Cultura de Vizcaya.
- Beach, D. M. (1938). *The Phonetics of the Hottentot Language*. Cambridge: Heffers.
- Borgström, C. Hj. (1937). The Dialect of Barra in the Outer Hebrides. *Norsk Tidsskrift for Sprogvidenskap VIII*. Oslo.
- Caramazza, A. & Yeni-Komshian, G. (1974). Voice onset time in two French dialects. *Journal of Phonetics*, 2, 239-245.
- Chao, Y. R. (1947). *Cantonese Primer*. Publications for the Harvard Yenching Institute. Cambridge, Massachusetts: Harvard University Press.
- Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley, California: University of California Press.
- Cole, D. T. (1955). *An Introduction to Twana Grammar*. London: Longmans, Green and Co.
- Daniels, F. J. (1958). *The Sound System of Standard Japanese*. Tokyo: Kenkyusha Ltd.
- Dasgupta, S. (1971). *Grundzüge einer Laut- und Schriftlehre des Bengälischen*. Heidelberg: Julius Groos Verlag.
- Dent, L. (1976). Voice onset time of spontaneously spoken spanish voiceless stops. Presented at the 91st meeting of the Acoustical Society of America, April, 1976.
- Dixon, R. M. W. (1972). *A Grammar of Yidi*. Cambridge: Cambridge University Press.
- Dow, F. D. M. (1972). *An Outline of Mandarin Phonetics*. Canberra: Faculty of Asian Studies in Association with the Australian National University Press.
- Fischer-Jørgensen, E. (1954). Acoustic Analysis of Stop Consonants. *Miscellanea Phonetica*, 2, 42-59. Also in *Readings in Acoustic Phonetics*, Lehiste, I. ed. Cambridge, Massachusetts: MIT Press.
- Flege, J. E. & Brown, W. (1982). Effects of utterance position on English speech timing. *Phonetica*, 39, 337-357.
- Haas, M. R. (1956). *The Thai System of Writing*. Washington, D.C.: American Council of Learned Societies.
- Han, M. S. (1962). *Japanese Phonology: An Analysis Based Upon Sound Spectrograms*. Tokyo: Kenkyusha Ltd.
- Hardie, D. W. F. (1948). *A Handbook of Modern Breton*. Cardiff: University of Wales Press.
- Harris, J. W. (1969). *Spanish Phonology*. Cambridge, Massachusetts: MIT Press.
- Hashimoto, O. K. Y. (1972). *Phonology of Cantonese*. (*Studies in Yue Dialects I*). Princeton, New Jersey: Princeton University Press.
- Hebert, R. J. & Poppe, N. (1963). *Kirghiz Manual*. Uralic and Altaic Series 33. Indiana: Indiana University Publications.
- Holmer, N. (1947). *Critical and Comparative Grammar of the Cura Language*. Göteborg: Elanders boktrycker aktiebolag.
- Houlihan, K. (1982). Is Intervocalic Voicing a Natural Rule? Paper presented at Annual meeting of the Linguistic Society of America, San Diego.
- Houlihan, K. & Iverson, G. (1979). Functionally-constrained phonology. In *Current Approaches to Phonological Theory*. (Dimmson, D. A. ed.) Indiana: Indiana University Press.
- Jha, A. (1977). *An Outline of Marathi Phonetics*. Poona: Deccan College.
- Jones, D. & Ward, D. (1969). *The Phonetics of Russian*. Cambridge: Cambridge University Press.
- Kahn, D. (1976). *Syllable-Based Generalizations in English Phonology*. MIT dissertation, distributed by Indiana University Linguistics Club.
- Kadwala, L. (1972). *Comparative Historical Phonology*. Translated by Irene Kadwala. In *The Language of Languages*. (Uemachi, I. & Inoue, S. eds.) Japan Linguistics Series, Part 134. The Hague: Mouton.
- Kao, D. L. (1971). *Structure of the Syllable in Cantonese*. The Hague: Mouton.
- Kaufman, E. (1971). *Tribal Phonology and Morphology*. Berkeley, California: University of California Press.
- Keating, P. A. (1983). Phonetic and phonological representation of stop constant voicing. *UCLA Working Papers in Phonetics*. Los Angeles: University of California.
- Keating, P. A., Mikos, M. J. & Ganong, W. F. (1981). A cross-language study of range of voice onset time













