I. INTRODUCTION

An examination of the phonologies of the languages of the world permits the following generalizations:

1. Sequences of /w/ followed by a back rounded vowel (e.g., /u, o/) are less frequent than /w/ followed by nonrounded vowels (e.g., /i,

1 The phonetic transcription used throughout is that approved by the International Phonetic Association, as of 1979. Forms in square brackets [ . . . ], represent detailed or
Table I. Development of (Af)fricated Release in Stops in Bantu Dependent on Quality of Following Vowel

<table>
<thead>
<tr>
<th>Proto-Bantu</th>
<th>Maganja</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>*-biad-</td>
<td>-bžal-</td>
<td>plant</td>
</tr>
<tr>
<td>*-pi-</td>
<td>-pc-</td>
<td>become cooked</td>
</tr>
<tr>
<td>*-kimba</td>
<td>-cimba</td>
<td>excreta</td>
</tr>
<tr>
<td>*-pet-</td>
<td>-pet-</td>
<td>bend</td>
</tr>
<tr>
<td>*-kaan-</td>
<td>-kan</td>
<td>deny</td>
</tr>
</tbody>
</table>

But: *-pet-


e, a/). For example, in English, sequences of #stop w + u do not exist; in Yao, sequences of Cw + [ɔ, u] or [uɔ] do not occur (Purnell, 1965).

2. Stops are more likely to develop an (af)fricated release before a following high vowel or glide (e.g., /i, y, u, u/) than before a nonhigh vowel (e.g., /e, e, æ, a/: see Tables I and II).

3. In the segment inventories of languages, stops are more likely to be differentiated by voice onset time (VOT), voice quality, and airstream mechanisms, than are other segment types (see Table III).

These generalizations are good candidates for phonological universals, i.e., cross-language regularities in the behavior or patterning of speech sounds. As this term has come to be used (Greenberg, Ferguson, & Moravcsik, 1978), a sound pattern need not be unexceptionless or manifest in every human language to count as a universal. It is assumed, however, that the causes of these patterns are universal in an absolute sense and are potentially manifest in every human language, indeed, in every human speaker. However, their actual realization may vary from language to language due to language-specific (i.e., nonuniversal) factors, for example, historical, psychological, and cultural.

narrow phonetic transcriptions, those bounded by slashes /.../ represent broad or phonemic transcriptions, and those in italic are purposely ambiguous as to the level of phonetic detail that they represent (in some cases they represent standard orthographic representation). Forms marked with an asterisk (*) are unattested in speech or texts but represent a reconstructed form. The symbols (>) and (<) stand for ‘became’ and ‘derived from,’ respectively. A tilde (~) between cited forms means ‘freely alternates with.’ V, N, and C are cover symbols for, respectively, any vowel, any nasal consonant, and any consonant (or, when used with N, any obstruent consonant). The information cited for English and Japanese is based on personal observations of the author.

Table II. Development of Affricates in English

<table>
<thead>
<tr>
<th>Orthographic representation</th>
<th>Original pronunciation</th>
<th>Modern pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>soldier</td>
<td>['səldər]</td>
<td>['səldər]</td>
</tr>
<tr>
<td>actual</td>
<td>['ækʃuəl]</td>
<td>['ækʃuəl]</td>
</tr>
<tr>
<td>nature</td>
<td>['netʃuər]</td>
<td>['netʃuər]</td>
</tr>
</tbody>
</table>

Thus, the affrication of dental stops is a fixed form of allophonic variation in Japanese and occurs before high vowels /i/ and /u/ (e.g., /tutii/ is [tsutʃi] ‘ground’), but it is a sporadic feature in English and occurs primarily before /j/ and /w/ (e.g., "Tuesday" [tʃuˈziːdi] or [ʃuˈziːdi]; "truck" [tʃuˈtʃuːk] or [ʃuˈtʃuːk]). Nevertheless, all human speakers are subject to the aerodynamic constraints causing the affrication, i.e., that audible air turbulence, frication, is more likely to arise as the channel through which the air escapes after the release of the stop becomes narrower.

If we focus on widespread cross-language sound patterns, we can be assured that their ultimate causes reside in the only thing that all speakers have in common, namely, the anatomical and physiological mechanisms that serve speech production and speech perception. It follows, then, that a careful, informed, even inspired examination of phonological universals can tell us—or, at least, hint to us—how speech is produced and perceived. This is why phonological universals are worthy of the attention of the speech pathologist—as well as others in speech science (Olab, 1975a).

In the sections to follow I will offer examples of types of phonological universals, discuss their origin, and then suggest some specific ways that they may be of use to the speech pathologist.

II. HOW PHONOLOGICAL UNIVERSALS ARE MANIFESTED

Some of the ways phonological universals are realized have been illustrated by the examples already given. A more systematic list is the following:

a. Allophonic Variation. Phonemes can show allophonic variation conditioned by specific phonetic environments. For example, it has been ob-

2 In the interest of brevity, only a few examples are given of the phonological universals discussed in this chapter. However, it should be understood that many more examples could be provided—from languages widely separated from each other in time, in family membership, and in geography.
served that the phonemes /b, d, g/ in English, although voiced intervocally, are most often devoiced in utterance-initial position and after voiceless obstruents (Lisker & Abramson, 1964; Smith, 1978). Of these three stops, /g/ has the greatest likelihood of being devoiced. Similar tendencies have been noted in Rundi (Meeussen, 1959).


c. Morphophonemic Variation. Just as allophones of a given phoneme may alternate as a function of their phonetic context, so may phonemes alternate as a function of their morphological context. For example, in Nubian, word-final stops become geminate with the addition of certain

<table>
<thead>
<tr>
<th>Noun stem</th>
<th>Stem + ‘and’</th>
<th>English Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>/fab/</td>
<td>/fabːn/</td>
<td>father</td>
</tr>
<tr>
<td>/segəd/</td>
<td>/segətːn/</td>
<td>scorpion</td>
</tr>
<tr>
<td>/kadəɭ/</td>
<td>/katʃːn/</td>
<td>donkey</td>
</tr>
<tr>
<td>/muɡ/</td>
<td>/mukːn/</td>
<td>dog</td>
</tr>
</tbody>
</table>

Table V. Dialectal Variation Evident in Cognate Words in Various Bantu Languages

<table>
<thead>
<tr>
<th>Language</th>
<th>Dialectal variants</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sagala</td>
<td>m—bale</td>
<td>variety of palm tree</td>
</tr>
<tr>
<td>Lwena</td>
<td>ma—bale</td>
<td></td>
</tr>
<tr>
<td>Matengo</td>
<td>n—dahi</td>
<td>bamboo</td>
</tr>
<tr>
<td>Pokono</td>
<td>mu—yanzi</td>
<td></td>
</tr>
<tr>
<td>Luyana</td>
<td>in—dowa</td>
<td>mud, clay</td>
</tr>
<tr>
<td>C. Kongo</td>
<td>loba</td>
<td></td>
</tr>
<tr>
<td>Yao</td>
<td>n—gunda</td>
<td>garden</td>
</tr>
<tr>
<td>Kikuyu</td>
<td>mo—yonda</td>
<td></td>
</tr>
</tbody>
</table>

suffixes; in these cases /b/ retains its voicing but stops articulated further back than that become voiceless (Bell, 1971; see Table IV). A similar pattern occurs in Brabantic Dutch (D. L. Goyvaerts, personal communication.)

d. Dialect Variation. One often finds dialectal alternations between morpheme-initial voiced stops and some other segments, the voiced stops being found in the environment of a preceding nasal consonant (see Table V). (This pattern also shows up as allophonic variation in languages such as Spanish, e.g., [sofra] 'surplus,' but [sombra] 'shadow'; [boďa] 'marriage', [bondaď] 'goodness').

e. Segment Inventories. The lists of segments in many languages exhibit common patterns, e.g., among voiced stops, including voiced implosives, it is back articulated stops that are most often missing (Gamkrelidze, 1975; Greenberg, 1970; see Table VI).

f. Phonotactics. In every language there are certain restrictions as to which phonemes can appear next to which other phonemes. Such "phonotactic" or "morpheme structure" constraints show up in substantially identical forms in many unrelated languages. For example, Ebrie allows /d/ only before /i, u, j, w/ (Dumestre, 1970).

<table>
<thead>
<tr>
<th>Table VI. Stop Inventories of Languages Lacking Voiced Stops Having Back Articulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thai</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Quileute</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Mbay</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Kisi</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Zulu</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Other possible manifestations of phonological universals could be added to this list, e.g., frequency of occurrence of specific sounds in the lexicons of languages or in connected discourse, but there are not many cross-language data on such patterns. The list just given represents the kinds of phonological data that are most readily available and from which cross-language similarities emerge.

III. ORIGIN OF PHONOLOGICAL UNIVERSALS

All of the preceding manifestations of universal phonological patterns can be viewed as stemming from only two basic processes: allophonic variation and sound change. First, the physical constraints of the speech apparatus cause some allophonic variation, presumably in all speakers and in all languages at all times in history—given, of course, that the proper circumstances prevail (e.g., normally one cannot expect to find allophonic variation in the degree of voicing of word-final stops in a language that does not permit any word-final obstruents). Second, some instances of allophonic variation will lead to sound change which may be manifested in a variety of ways. Some sound changes will give rise to morphophonemic and dialect variation if the sound change affects a word in one morphological context but not in another, or if it affects one sister language or dialect but not another. Furthermore, some sound changes may result in the appearance or complete disappearance of some speech sounds, thus affecting languages' segment inventories. If this only happens in the environment of specific phonemes, it will then affect the phonotactics of the language.

In essence, then, the preceding list reduces to two fundamental processes: (1) inherent physically caused variation in speech; and (2) sound change. It will be useful to examine these two more closely.

A. Inherent Variation

By inherent variation in speech is meant any unintended ("nonprogrammed") feature of pronunciation of a word or other element of speech, which feature can be traced to physical phonetic causes, that is, universal human constraints of a neurological, physiological, aerodynamic, anatomical, acoustic, auditory, or similar nature. One example of this sort has already been mentioned, viz., the aerodynamic constraint that noise from air turbulence will be generated when the air escaping after release of a stop passes through a sufficiently narrow channel at or above some critical
velocity. This is a physical necessity; it will happen whether or not the speaker wishes it. As a further example, all of the sound patterns cited in Section II may ultimately be traced to some physical factors affecting voicing—traced, that is, to physically caused inherent variation that leads to sound changes. I will consider this example in some detail (based on Ohala and Riordan (1979) and references cited therein).

Voicing, that is, the vibration of the vocal cords, requires that the vocal cords be in the proper physical configuration and that there be sufficient air flow through the glottis. We may assume for all relevant cases considered above that the vocal cords have the configuration required for voicing: the presence or absence of voicing, then, depends entirely on factors affecting air flow. During a stop, the air flowing through the glottis accumulates in the oral cavity, causing oral air pressure to build up. Eventually, oral pressure will approach subglottal pressure and the transglottal pressure drop will fall below the minimum required (~2 cm H₂O) to maintain the rate of airflow necessary for voicing. Expansion of the oral cavity volume is required if voicing is to continue beyond some 10 to 15 msec of stop onset. Passive cavity enlargement, due to tissue compliance, seems capable of prolonging voicing to about 70 msec after stop onset; active cavity enlargement (lowering the larynx and jaw, raising the velum, etc.) will permit voicing to persist uninterrupted for an even longer period. Nevertheless, the longer a stop is maintained, the more difficult it becomes to continue voicing. Moreover, the tendency to devoice is greater, the smaller is the surface area of the oral cavity (i.e., the less is the capacity of the oral cavity to absorb the glottal air flow). These principles will explain the patterns previously noted.

Presumably, the implosives in Sindhi developed, in part at least, in response to the need for active oral cavity expansion in order to maintain the voicing throughout geminate stops. Nubian, however, shows a different strategy: the geminates stops were “allowed” to devoice, except in the case of the labials which have sufficiently large surface area to absorb the glottal flow. The infrequency of voiced velar stops in languages’ stop inventories is another manifestation of this pattern. Similar factors must have determined the pattern cited for Ebrié: stops coarticulated with high vowels have greater oral cavity surface area than stops coarticulated with nonhigh vowels (due to the enlarged pharyngeal cavity associated with high vowels). The pattern of dialectal variation in Bantu is due in part to the fact that leakage of the oral pressure via the nasal cavity during a preceding nasal preserves the auditory impression of voicing before the release of the stop, without the necessity of active cavity enlargement nor the risk of devoicing.

There are, of course, a great many more physical constraints of the speech mechanism, including auditory constraints, which account for inherent variation in the speech signal as the listener perceives it. Basic research in speech continues to uncover more of these constraints, from which arise a myriad of features in utterances which are not intentionally programmed.

B. Sound Change

In most cases, listeners can tell which features of pronunciation are intended. This is a consequence of their long experience with speech in general and with their native language in particular. Indeed, given the high redundancy of speech, not only unintentional distortions of speech sounds but even their complete obliteration may be overlooked (Warren, Obusek, & Ackoff, 1972). Nevertheless, sound change (change in pronunciation) can occur when a listener is unable to differentiate intended from unintended features of speech. When this listener turns speaker, he may purposely incorporate previously unintended features in his pronunciation. The process is somewhat analogous to the copying errors of medieval scribes. Like them, one scribe’s unintentional slip of the pen could be propagated throughout many other manuscripts when other scribes faithfully copied it.

The scenario whereby sound changes can propagate and eventually be characteristic of an entire linguistic community’s pronunciation is unknown at present, but the important point for our purposes is that they can, ultimately, be traced to a misperception or misapprehension of the speech signal on the part of an individual speaker/hearer. This point used to be controversial; it had been previously suggested that an entire population’s pronunciation shifts at the same rate in the same direction in indetectably small increments. I do not think that this notion can seriously be maintained anymore. Two areas of recent research support the view of sound change originating with the individual and involving abrupt and often large changes in pronunciation. First, it has been found that the universal intrinsic variations in speech parallel sound changes both in regard to the specific environments causing the variations or change and in regard to the direction of the change (Hombert, Ohala, & Ewan 1979; Ohala, 1971, 1974a,b, 1975b, 1976, 1978; Ohala & Lorentz 1977). Second, various studies in speech perception have demonstrated that individuals’ perceptual (re)evaluations of speech sounds also parallel widely attested sound changes (Janson, 1979; Javkin 1977; Jonasson 1971; Ohala, Riordan, & Kawasaki, 1978; Wright 1975, 1980). These parallels cannot be coincidental.

What are the overall implications of this view of the origin of phonologi-
cal universals for workers in speech pathology? It allows us to apply the principles derived from a study of the phonological macrocosm to our analysis of the phonological microcosm (i.e., to the speech behavior—normal or aberrant—of the individual speaker/hearer). I will elaborate on this in the sections to follow.

IV. PARALLELS

Experienced speech pathologists have a good “feel” for which misarticulations are due to maturational problems and which are due to organic causes (anatomical, physiological, or neurophysiological deficiencies). The former may solve themselves in time, or at least require less profound therapeutic measures; the latter require all the art and scientific expertise that the field can muster. It may be possible to sharpen or reinforce this sense by a study of phonological universals since the same factors are often responsible for them and for delayed or aberrant phonological acquisition. I will detail in what follows two parallels between children’s phonological mistakes and “adult” phonologies [The following discussion is adapted from Greenlee and Ohala (in press), which should also be consulted for further documentation of the data presented.]

A. Voiceless Nasals Substituted for /s/ + Nasal Clusters

Several investigators have reported that some children learning English substitute voiceless nasals for initial s + nasal clusters (e.g., “Smith” [smtp], “smack” [smæk], “snip” [snɪp]). Voiceless nasals have developed via sound change from exactly the same source (e.g., Burinese /mʰ/ ‘nose’ < (earlier) sma, /mʰ/ ‘ripe’ < (earlier) smiŋ). Likewise, some instances of Awadhi /-n/- (a breathy voiced nasal) derive from original Old Indo-Aryan sn-, e.g., /dʒɔnɔːjaː/ ‘moonlight’ < jyotsna. Similar origins of voiceless nasals from s + nasal clusters in Indo-European are reported for Primitive Greek and Old Irish.

The voiceless nasals in Burmese are actually only partially voiceless; they start voiceless but end voiced (Ladefoged, 1971, p. 11). The same was undoubtedly true in the case of Greek and Irish as well as in the various cases reported in the language-acquisition literature. As such, then, a sequence of a voiceless followed by a voiced nasal is a very plausible substitute for an s + nasal sequence. The voiceless nasal, like the s—also voiceless—conveys the same auditory impression of high-frequency noise. Moreover, the NN sequence is much simpler to articu-

late, requiring only a single supraglottal articulation and a change from the voiceless to voiced state at the glottis.

B. Treatment of Final NC Clusters

It is well known that young children have considerable difficulty mastering consonant clusters and that some sort of simplification of the clusters often occurs. Many investigators, however, have reported a curious asymmetry in children’s simplification of syllable-final nasal + stop clusters, viz., a nasal + voiceless stop usually has the nasal deleted but a nasal + voiced stop is either mastered earlier or has the stop deleted (Hernández-Chávez, Vogel, & Clumeck 1975). For example, English “don’t” [dɒt], “ants” [nts], but “hand” [hænd], “mend” [mɛnd]; Spanish “elefante” [elɛfante], “chancho” [tanko] ~ [tanko], but “grande” [ɡɡaŋ], “mambo” [mambo].

This asymmetry is manifested also in many languages’ phonologies. The phonotactics of the native vocabulary in Hindi, for example, permit sequences of long nasalized vowel + (homorganic) nasal + voiced stop as well as long nasalized vowel + voiceless stop (e.g., [sɪːk] ‘twig’ and [sɪːŋ] ‘horn’), but does not permit sequences of long nasalized vowel + voiced stop or long nasalized vowel + nasal + voiceless stop, i.e., [sɪːŋ] and [sɪːɡ] are not possible words in the native vocabulary (M. Ohala, 1972). Also, Timm (1978) reports that in Breton, the sequence vowel + nasal + obstruent may become nasalized vowel + obstruent, especially when the obstruent is voiceless (e.g., [jowɔŋ] > [jowɔŋ] ‘young’, [dɔm] > [dɔm] ‘to us’). Malécot (1960) and Lovins (1978) have shown that the same pattern can be found in American English.

The origin of this pattern undoubtedly lies partly in the fact that vowels and sonorants before syllable-final voiceless obstruents are shorter than those before final voiced obstruents. They may, in fact, become so short as to be perceptually insignificant and thus liable to omission when imitated by language learners.

V. WHAT TO LOOK FOR

If phonological universals are indeed macroscopic reflections of the processes shaping speech at the “microscopic” level of the individual speaker/hearer, they could, properly interpreted, tell us what to look for in individual speech behavior. This could be a more efficient process of uncovering significant facts about those with aberrant phonologies than
waiting for the speech pathologist to discover them by chance. I provide some speculative examples in the following section.

A. The Mastering of Speech Sounds in Specific Environments

Speech scientists concerned with the acquisition of phonology recognized fairly early that one could not say simply that a child had mastered or had not mastered a given phoneme at a specific time. First of all, the notion of “mastery” is a slippery one since performance may vary between recognition and production, and even in production it may vary between spontaneous, unprompted speech and speech which is a direct, prompted imitation of others. Furthermore, a phoneme has various allophones and all may not be mastered at the same time. Finally, there is the question of whether the concept “phoneme” in an incomplete phonological system has the same meaning—or any meaning—as it does in reference to a fully acquired phonological system (cf. Moskowitz, 1970).

Although I will not propose any solution to these problems, I think it is worth pointing out that phonological universals indicate that generalizations on the “hardness” or survivability of certain speech sounds or classes of speech sounds must often make reference to very specific phonetic environments—far more specific than the “initial, medial, and final position” one finds in many of the diagnostic checklists that speech pathologists are urged to use.

As mentioned earlier, stops, particularly apical stops, have a tendency to be slightly affricated before high close vowels and glides. This tendency might imply that the contrast /t/ versus /ts/ and /d/ versus /dz/ would be more difficult to learn before high vowels or glides.

The aerodynamic constraints that tend to produce affrication also tend to increase the degree of aspiration of voiceless stops (i.e., aspiration: delayed voice-onset time following stop release) is greater before high vowels and glides than in other environments. This occurs because the increased resistance created by the close constriction retards the flow of air following the release and thus delays the achievement of the transglottal pressure drop needed for voicing (Ohala, 1976; in press). So strong is this tendency that in some languages voiceless stops are reported to show marked aspiration only before high vowels, e.g., Fe?e? (Hyman, 1972, p. 23), Ebrié (Vogler, 1968), and Tucano (West & Welch, 1967). This pattern implies that the contrast between aspirated and unaspirated stops will be greatest before high vowels. Therefore, it would not be unlikely that such contrasts would be first mastered in these environments or, at least, less likely to be a source of confusion among language learners.

It is a common observation that languages often have fewer nasal vowels than oral vowels, and, further, that the reduction in the nasal vowel inventory generally occurs along the height dimension rather than the front–back dimension (Ruhlen, 1975; Williamson, 1973). Table VII shows these patterns in selected languages. In addition, there are many reports that nasalization induces shifts in vowel quality, again, generally in the height dimension. For example, in Binumarian /i/ > [I] and /u/ > [ɔ] before NC sequences (Oatridge & Oatridge, 1966) and in Basque, contextually nasalized mid-vowels are raised and tensed (Lochak, 1960). In a related pattern, many dialects of American English neutralize the contrasts between /i/ and /e/ before nasal consonants.

Distinctively nasal vowels and vowels nasalized by proximity to a nasal consonant, whatever their actual tongue shape, are realized acoustically with formants having very broad bandwidth (Ohala, 1975b) and this makes them more difficult to differentiate from each other (cf. Wright, 1980). Moreover, the first formant is the most heavily distorted formant, accounting for vowel height being less clear and more variable than the front–back dimension (the auditory dimension of “vowel height” is primarily cued by a vowel’s first formant).

On the basis of this pattern, we might expect that the mastery of a language’s vowels would be accomplished later or less perfectly in the environment of nasal consonants than elsewhere (in languages which have

| Table VII. Inventories of Oral and Nasal Vowels in Selected Languages Showing that There Are Often Fewer Nasal Vowels than Oral Vowels and that Neutralization of Vowel Quality Generally Occurs in the Height Dimension, Not the Front–Back Dimension |
|---|---|
| French | Albanian |
| i | i | y | u | ï | ŭ | e | ø | ə |
| e | ø | o | ə | ɛ | ɔ | e | e |
| Ebrié | I | u | Akan | i | u | ï | ɨ | 0 | e | ø | o |
| e | ø | e | ə | ɛ | ɔ | a | õ |
| Mazahua | ikwerre | i | u | ï | ɨ | 0 | e | ø | e | ə |
| e | ø | e | ə | ɛ | ɔ | a | ə | ə | ə | ə |
extensive assimilatory nasalization of vowels in the environment of nasal consonants).

B. Perceived Nasality May Not Be Physiological Nasality

The development of nasal vowels normally stems from loss of a contiguous nasal (e.g., French bon /bɔ̃/ from earlier /bon/). There are, however, a fair number of instances where nasalization appears on vowels in the complete absence of an earlier neighboring nasal consonant. Some examples follow.

Maier (1969) reports that Cua has contractive nasalization only on a few words, and these words are usually initiated with an /h/ and terminate with a voiceless stop, /t, k, c, j/ or /h/, e.g., /hɛŋ/ ‘fat’ versus /heŋ/ ‘finished’. In addition, similar such cases, including the development of nondistinctive nasalization near /h/, are abundant among Southeast Asian languages (Matisoff, 1975).

Phillips (1976) notes that in Los Angeles Spanish si ‘yes’ is often pronounced as [si] or even [i]. In Awadhi (and many other modern Indo-Aryan languages) there are pronunciation variants such as /gəːs/ ~ /gāːs/ ‘grass’ and words which today have nasal vowels, whereas their earlier forms did not, e.g., [səːp] < sarpa ‘snake’, [əːkʰ] < aksi ‘eye’. In various dialects of Korean, there is evidence of nasal consonants being inserted where there was no apparent historical motivation for them (Ramsey, 1975, p. 61ff.), e.g., Chinese (loan word) /pha chyo/ > Korean /panchyo/ ‘plantain.’ A similar pattern was found in Margi by Hoffman (1963, p. 42).

It is a curious but significant fact that almost all such cases of ‘sponaneous nasalization’ involve vowels flanked by fricatives, affricates, or aspirated consonants, i.e., consonants characterized by heavy air flow. In addition, there is strong evidence that consonants having greater than normal air flow have a more open glottis (Slis, 1970) and thus that, through assimilation (anticipatory or perseveratory), vowels adjacent to these consonants will be produced with a slightly open glottis, although still sufficiently close to allow vibration [cf. air flow records supporting this (Ohala, 1976)]. I have speculated that this open glottis condition during vowel production will create acoustic correlates that imitate those produced by nasalization: broad bandwidth in formants, especially the first formant (Ohala, 1975b). This hypothesis is supported by acoustic studies such as those by Fant (1973, p. 8) and Fujimura and Lindqvist (1971); perceptual tests of the hypothesis are currently underway. If this hypothesis is correct, it is conceivable that some listeners hearing what they think is vowel nasalization—since that is what it sounds like—will produce these vowels with actual nasalization, which would explain the aforementioned sound patterns.

It would be of interest to look more carefully at pathological speech to see if vowel nasalization “erroneously” appeared on vowels adjacent to fricatives, etc., more often, or to a greater extent than, on vowels near other consonants. There is some evidence that this is the case (Lintz & Sherman, 1961; Moore & Sommers, 1973) although it is not possible to say definitively whether these effects are due to actual nasalization or (as hypothesized here) an open glottis condition that produces acoustic characteristics resembling nasalization.

VI. THE AUDITORY BASIS OF SPEECH: SUGGESTIONS FOR THERAPEUTIC METHODS

There is a classic method in behavioral analysis for discovering the underlying purpose or goals of an organism’s behavior. Simply described, given a well-defined behavior, one puts artificial obstacles in the way of the organism and observes which aspects of its behavior vary and which aspects do not vary (Bernstein, 1967). One may conclude that those aspects of behavior which do not vary represent the goals underlying the organism’s behavior; those aspects which do vary are done for the sake of the goal. For example, when it is very cold, a human will shiver; and when it is very hot, he will perspire. Throughout these changes, body temperature remains constant at about 37°C. This warrants the conclusion that shivering and perspiring are done for the sake of maintaining constant body temperature in the face of extremes of temperature. Of course, there is a hierarchy of goals and one goal may serve and therefore yield to a higher goal. A higher goal than constant body temperature is the overall well-being of the body; thus body temperature is allowed to rise during a fever in order to combat harmful microorganisms.

I believe that it is possible to take a “long” look at sound change and make informed speculations on the underlying goal of speech. Of course, the ultimate goal is communication of a message. But the message must be encoded for transmission. What is the ultimate goal in this encoding process? Is the brain trying to make articulatory gestures which can be heard; i.e., is articulation primary? Is the brain trying to make acoustic–auditory signals by means of articulatory gestures; i.e., is the acoustic–auditory signal primary? I do not think there is any question that the brain has stored and knows how to produce both articulatory and acoustic–auditory patterns (Fowler & Turvey, 1978; MacNeilage, 1973; Ohala, in press...
sound change involves labialized velars (or velar followed by /w/) to labials (Ohala & Lorentz, 1977). For example, Proto-Yuman *imalik’i > Maricopa /mel’epu/ (Haas, 1963); Proto-Indo-European *gʰiwos (cf. English quick in sense of ‘living’) > Greek bios ‘life’ (whence the first morpheme in biology, biosphere, etc.).

Voiceless stops > [ʔ]. Glottal stop is a common substitute of voiceless stops. For example, Standard English ['bɑɹ] or ['bɑt] but Cockney ['baʔ] ‘bottle.’ (The acoustic similarity of the components of this example should be obvious; the acoustic basis of the other two sound changes are provided in the references cited.)

It might be claimed, however, that the above data only argue for the acoustic-auditory pattern being primary in the social domain, that is, in the way language is used and passed on from one speaker to another, and does not necessarily indicate what the primary purpose of the individual speaker is. (In a similar way ethnologists can speak of animals having an instinct for “preservation of the species,” but there is no necessity for assuming that the same purpose motivates each individual engaging in reproductive behavior.)

Logically, the criticism is correct, but it is implausible—to me, at least—that the purposes of the individual speaker and the community of speakers as a whole could be so dissimilar. Moreover, there is additional evidence such as that coming from studies of ventriloquists’ modifications of articulation and from studies of compensatory articulation (see Lindblom & Sundberg, 1971; and Riordan, 1977), all of which support the conclusion that I have derived from the sound-change data.

Although the issue is likely to remain controversial for some time, the implications of this view for speech pathology could be far-reaching. At the very least, it provides additional support for those who are pessimistic about the effectiveness of the exclusively “oral” approach in education of the profoundly deaf and who favor instead the “manual” approach (use of American Sign Language) or the “total communication” approach (oral plus manual). It also reinforces the view that in some cases faulty articulations, especially those due to anatomical deficiencies, need not always be corrected by drastic corrective surgery or intensive training designed to get the patient to approximate normal articulation. Rather, looking for substitute articulations which produce sounds auditorily close to the target sound might be more appropriate. Indeed, it seems that, left on their own, many speakers with anomalies of the speech organs spontaneously discover such substitute articulations (Eskew & Shepard, 1949; Goldstein, 1940; Weinberg, Christensen, Logan, Bosma, & Wornall 1969; Weinberg

Palatalized labials > dentals. As documented in Ohala (1978), there are numerous cases of palatalized labials (or sequences of labials plus palatal glide or vowel) changing to corresponding dentals. For example, Standard Czech /pʃet/ corresponds to East Bohemian /tet/ ‘five’, Roman Italian [bʃjano] corresponds to Genoese [dʒanu] ‘white’, and Pre-Classical Greek *kom-yo- corresponds to Classical Greek koînos ‘common’.

Labialized velars > labials. A well known and extensively documented
In addition, in cases where it is judged worthwhile to train the patient to approximate normal articulation, the prime importance of providing auditory feedback would be indicated.

VII. THE IMPORTANCE OF ACOUSTIC MODULATIONS

It is a common assumption in speech science that what the speaker produces, and what the listener in turn hears, is a string of phonemes. At some level these phonemes are assumed to be quasi-steady-state entities. Of course, it is well known that the rapid formant transitions between consonants and vowels are important cues for the perception of place of articulation and other features of consonants. Recently it has been shown that these formant transitions also help to cue vowel quality as well (Strange, Verbrugge, Shankweiler, & Edman, 1976). Nevertheless, there is still a tendency to think of these transitions as cuing phonemes.

Research in sensory physiology, however, suggests that sensory organs are more sensitive to sharp modulations in the “carrier signal” to which the particular sensor is sensitive than they are to slow modulations (e.g., the cells in the retina respond more markedly to “edges,” spatial modulations in light intensity, than they do to the same intensity differences in areas not immediately adjacent to each other). It is for this reason that flashing lights, frequency-modulated sirens, etc. are used to attract attention. We can safely conclude that, in listening to speech, our auditory system derives more information from the transitions (i.e., any rapid change in spectrum, amplitude, or periodicity of the speech signal) than it does from an extraction of the same parameters of those portions of speech in between the transitions. If this is so, a word such as “deep” with unreleased final stop, would have two, not three peaks in the function expressing rate of information transfer. It is not clear how or where “phonemes” can be fitted into this scheme.

Certain phonological universals reinforce this view. In Section I, it was mentioned that sequences of /w/ plus rounded vowels are often missing from languages’ lists of permitted phoneme sequences or have a severely restricted distribution. Besides English and Yao, the languages cited there, this phonotactic constraint can be found in several other languages, including Japanese, Touna (Ivory Coast) (Bearth, 1971), and New Caledonese (Rivierre, 1973). The reason, of course, is that the transitions between /w/ and rounded vowels are not very large (i.e., the second formant is low for both segments).

Similar constraints hold for sequences of /j/ plus the palatal vowels /i, e/.

In English and Swahili, /j/ is rare (Tucker & Ashton, 1942) and is missing entirely from Japanese. The reason is that the second formant is high for both segments. Moving from the /j/ to /i/ or /e/ does not produce a very large modulation of the second formant. Other examples of the same sort could be mentioned (e.g., the rarity of initial /tl/ and /dl/ clusters, all of which can be accounted for by reference to the magnitude of the acoustic modulation created by such sequences).

There are potentially significant implications in all of this for speech pathology. At the very least, it would indicate that the speech pathologist should put as much emphasis on articulator movements, on timing, and synchronization of one articulator’s movements with respect to other articulators (in order to maximize the acoustic modulation produced) as is normally put on achieving precise articulatory positions.

VIII. CONCLUSION

In this work I have proposed that phonological universals, gleaned from the phonologies of languages of the world, can tell us how speech production and speech perception works. The same universal principles that shape speech in languages of the world, synchronically and diachronically, also shape speech in individuals, including those with abnormal speech and hearing.

Sound change, in particular, can be viewed as the results of a vast experiment run by Nature to discover the principles underlying speech production and perception. The phonological literature, accumulating for almost two centuries, is a valuable record of these experimental results. Unfortunately, it is not the case that most phonologists have organized and interpreted their data in a way that would be immediately useful to those with specific applications in mind. Nevertheless, there is a growing literature on phonological universals that deserves the attention of workers in speech pathology. I have marked such items in the References section of this chapter with an asterisk (*). However, it is still the task of the speech pathologist to give his own interpretation to these data in light of his own experience and needs. I believe that these efforts will be well rewarded.

Acknowledgments

This research was supported in part by the National Science Foundation. The opinions expressed and the accuracy of the data are the responsibility of the author.
References


Application of Phonological Universals


Ohala, J. J. Articulatory constraints on the cognitive representation of speech. In T. Myers, ...
Application of Phonological Universals


References

J. Laver, & J. Anderson (Eds.), The cognitive representation of speech. Amsterdam: North Holland, in press.


Smith, B. L. Effects of place of articulation and vowel environment on "voiced" stop consonant production. Glossa, 1978, 12, 163-175.


