In its most general form, a language universal is a proposition of universal scope, that is, it is asserted to hold over all languages. An assertion's universal scope does not insure that it is a significant generalization about the essence of language. In linguistics, as in other sciences, law-like assertions must be distinguished from those which are not lawlike. It is usual to exclude two sorts of universals as nonsignificant. First are those universals which are definitional. A definitional universal makes an assertion about language which distinguishes it from other communication systems. Within linguistics, definitional universals are tautological, and cannot be refuted. The second sort of nonsignificant universal is the accidental universal. This is a universal which happens to be true but is of no consequence for the nature of language.

One kind of accidental universal is an assertion which holds for the world's existing or recorded languages, but not for one or more possible nonexistent ones. This distinction implies that the world's languages are scattered over a universe of possible languages in an unknown manner.

This is the source of the concern voiced by Hockett (1966) when he wrote that 'a feature can be widespread or even universal without being important,' evoking in illustration the situation in which all the world's languages but English become extinct. Chomsky and Halle (1968), speaking of 'apparent linguistic universals which may be the result merely of historical accident ... and which are of no importance for general linguistics,' echoed this apocalyptic view, nominating Tasmanians as survivors.

This shows that it is necessary to understand 'all languages', in the statement of a universal, as referring to a wider class of languages than just those now existing or those that have been described. Does this mean that generalizations based on samples of languages are not to be taken seriously? No, for empirical evidence from even one language is of some weight. A proposition that holds for a few hundred languages must be taken seriously. But the distinction between existing languages and possible languages does remind us that even if the proposition were to be validated for all the world's languages, we would still not be assured of its linguistic significance.

The search for linguistic universals can be divided into three steps. The first is discovering a generalization, and testing it against empirical evidence. This might be called puzzle-finding. The next stage, which is what concerns me most here, is assessing
the significance of the generalization. In other words, is it a linguistic puzzle? Explaining the nature of the significance is a further step.

How can a lawlike generalization be recognized? It is not a matter of overcoming the difficulties of sampling, as we have seen. Our ignorance about the representativeness of the world's languages with respect to language in general is a problem mainly relevant to the puzzle-solving stage, not to establishing significance. The answer is that a significant universal must have consequence. If a universal is proven false, and our reaction is 'So what?', then the universal could not have been significant.

One approach is to show that a universal has consequence for the processes of linguistic change. This approach has been discussed at length by Greenberg (1966a, 1969). For example, consider this proposition:

'In a given language the number of nasalized vowels is never greater than the number of nonnasal vowel phonemes.' (Ferguson 1966).

This generalization is significant because it has consequence for a hypothesis about processes of historical origin and mutation: namely that nasal vowels arise uniquely through the agency of a nasal consonant, and that the merger of oral vowels implies the merger of corresponding nasal vowels. If the generalization were proven false, these diachronic hypotheses would have to be revised.

Now look at the second proposition:

For all languages, if a language possesses clicks, then it possesses nasal consonants.

This statement is not logically trivial, as it would be if we said that all click languages possessed vowels. All languages possess vowels, but some languages do not have nasals. Furthermore, the statement has no known exceptions. Nevertheless, it must be accidental. Otherwise we would be led to unacceptable conclusions about diachronic process, namely that the processes that lead to loss of nasals are inhibited by the presence of clicks, or that clicks could not develop in the absence of nasal consonants.

These two examples show that the lack of exceptions is no sure guide to significance. It also happens that the existence of exceptions to a generalization may not indicate its lack of significance. Greenberg (1969) discusses such a case. The generalization concerns the place of occurrence of voiceless vowels within the word. In general, voiceless vowels occur word-finally if they occur anywhere. An exception to this generalization is the Japanese described by Bloch (1950), in which voiceless vowels do not occur finally. As always, exceptions may indicate that a generalization rests upon an inappropriate typology. In this case, one might consider whether one should perhaps subsume final silence under a category of voice-
less segments. The new generalization, now covering the exceptions, would state that voiceless vowels occur between voiceless segments if they occur anywhere, and would imply that final position in itself has no special role.

This move is revealed as incorrect when the significance of the generalizations is investigated in terms of diachronic process. Japanese has no final voiceless vowels not because they arose only elsewhere, but because a successive process of loss has eliminated final devoiced vowels. The originally stated preference for final position is indeed significant for the creation of voiceless vowels, but it is obscured on the synchronic plane by the action of an interfering process.

Such 'statistical' universals, as they are sometimes called, are unfortunately much more numerous than exceptionless synchronic universals. In fact, whenever synchronic states are the product of several processes of origin and mutation, as is usual, exceptions to generalizations can be expected. It should be clear that they are not therefore to be automatically discarded, but deserve equal consideration with exceptionless generalizations.

This is true even where the 'exceptions' are so numerous that it is obvious that no implicational hierarchy of occurrence holds. In classifying a sample of languages according to what kinds of syllabic nasals occurred in them, I obtained the distribution shown in the table below.

<table>
<thead>
<tr>
<th>Inventory of syllabic nasals</th>
<th>Number of language-systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>11</td>
</tr>
<tr>
<td>n</td>
<td>5</td>
</tr>
<tr>
<td>η</td>
<td>5</td>
</tr>
<tr>
<td>m n</td>
<td>11</td>
</tr>
<tr>
<td>m η</td>
<td>4</td>
</tr>
<tr>
<td>n η</td>
<td>1</td>
</tr>
<tr>
<td>m n η</td>
<td>7</td>
</tr>
</tbody>
</table>

First of all, observe that all possibilities occur. One might conclude that the classification represents nothing but the chance of the sample. Certainly that would be a likely possibility. Yet the predominance of occurrence of m is suggestive, particularly considering that for nasals in general, it is n that is favored. As it turned out, by careful consideration of the processes of origin and mutation of syllabic nasals, I was able to show that in some circumstances m or η will become syllabic if any nasal does.
This last example takes us beyond the idealized schema in which one first discovers a generalization, and then assesses its significance. One could hardly say that there ever was a generalization about syllabic nasals until the diachronic regularity was discovered. The merging of the two steps is an advance that has much promise, and it is worth conscious pursuit. In this way process and state, diachrony and synchrony, are considered simultaneously. The scope of generalizations is taken to be the universe of historically transmissible languages, but that universe is not regarded as a static inventory of objects. Because languages do not just happen, but are derived from other languages by processes of change, these processes are considered as an integral part of the reference universe, so that no language exists in isolation, but is related to other languages by diachronic processes.

An idealized paradigm for the process-state approach to typological comparison might be the following. Over some aspect of language structure a typology is constructed (according to type of vowel system, or type of syllable structure, for example). Based on this typology, a process-state model is constructed. Evidence for the model is not simply attested states and attested processes. The necessary connection between the two makes extended hypotheses inevitable: if there is an attested state, some process must lead to it, and if we find no acceptable process leading to an unattested state, we must consider the possibility that it does not occur.5

In the case of the syllabic nasals, the concept of a process-state model was a neuristic guide. In studying the problem of the primacy of the CV syllable, I found an explicit use of a process-state model to be fruitful.6 The relevant empirical synchronic generalization is that all languages possess syllables of at least two types, of which one is CV (Greenberg, Osgood, and Jenkins 1966).

Upon examining the various possible syllable types and the processes that relate them, there does not appear to be any absolute stricture that would prevent a language with the unique syllabic form CV. The same holds for a language with just the syllable types CVC and VC, and similarly for other nonoccurring inventories. This is perhaps one reason why the significance of the generalization has been doubted. Yet the conditions for the creation of these syllabic inventories are so special, and the generality of the processes that would reduce them to types falling within the categories specified in the generalization is so great, that such languages must clearly be very rare and of short duration within the type. Thus the process-state approach in this case suggests that the generalization stated in terms of nonoccurrence of types is an accidental one, and that the significant linguistic regularity has to do with the relative likelihood of occurrence of the relevant processes.

The last two examples possess an interesting sort of complementarity. Concerning the CV syllable, I concluded that an exceptionless universal was significant only in a probabilistic sense, whereas the statistical preference for syllabic m and n was relateable to an implicational diachronic regularity.
There are other paths to the study of universals than the one I have been discussing. Prominent among them is the transformationalist approach, represented by Chomsky. The differences between the two approaches sometimes appear to be so great that it is hard to realize that they have the same goal. Much of the difference, I think, can be understood as consequences of different concepts of assessing significance.

The transformationalist view is that significant universals must have consequence for the acquisition of language. The following statements are typical of many that have appeared in the last five years:

The significant language universals are those that must be assumed to be available to the child learning a language as an a priori, innate endowment. (Chomsky and Halle 1968:4).

... nontrivial universals ... are properties that all natural languages have, though not by definition, and indeed any possible natural language must have if it is to be learned and understood by humans in the normal manner. (Moravcsik 1967:224).

Let us call universals of language that are essential to acquisition 'acquisition-significant' universals. An appropriate term for universals of the sort discussed earlier is 'transmission-significant' universals.

An acquisition-significant universal is necessarily transmission significant. A language cannot be transmitted if it cannot be acquired. The converse is not necessarily true, since the historical transmission of language involves cultural processes not associated with acquisition. In fact, there do exist universals that are transmission-significant but not acquisition-significant. Most of the generalizations considered so far are examples. Take Ferguson's proposition concerning nasal vowels. It is unlikely to be acquisition-significant; in the absence of an operational test, I presume that a language with more nasal vowels than oral vowels may be acquired in a natural way.

With the help of the figure below we may look at the difference between the two from another angle.

Diagram of sets of languages.

E = existing languages
T = historically transmissible languages
A = naturally acquirable languages
P = logically possible languages
Transmission-significant universals hold over the set of historically transmissible languages. This is a proper subset of the naturally acquirable languages, over which acquisition-significant universals hold.

With this background, we can return to the comparison of the two approaches. An example of an acquisition-significant universal will be helpful. The best-known is probably the principle of cyclic ordering in phonology: '... the phonological component of a grammar consists of a sequence of rules that apply in a cyclic manner...' (Chomsky 1967:415).

That this generalization, like most acquisition-significant universals, concerns underlying structure, whereas those claimed to be transmission-significant usually concern surface features of language, is no accident. It is a consequence of the different notions of significance. The statement of Chomsky (1965:209) '... that only descriptions concerned with deep structure will have serious import for proposals concerning linguistic universals' is only understandable if we take it to refer to acquisition-significant universals. Languages' surface structures exhibit much regularity, but for the most part it is hard to conceive that such regularities are essential to language acquisition. On the other hand, the lack of evidence for the diachronic development of underlying structure accounts for the virtual absence of attempts to show the significance of regularities in the underlying structure to historical processes.

The constraints imposed upon language by acquisition-significant universals tend to be less evident and less explicit than transmission-significant universals. In the case of the cyclic principle, its significance to acquisition is not so much that it defines languages that a child could not acquire, but that it must be '... one part of the schematism that the child brings to language learning' (Chomsky 1967:416). The claim that the schematism is highly restricted is based on indirect arguments concerning the rate of acquisition, nature of the linguistic data available to the child, etc. In contrast, the usual argument for the significance of a universal to transmission specifies the kinds of languages that cannot occur since their absence can be attributed to conditions on the processes of change that affect them.

The weaknesses and strengths of the approaches lie mainly in a final contrast. The approach of transmission-significance permits investigation of a broader range of linguistic phenomena, but it makes no commitment as to where one should look for the explanation of linguistic universals in terms of nonlinguistic phenomena, although the desirability of such wider explanation is patent.

The claim of significance to acquisition presupposes that the ultimate explanation lies in man's psychophysical capacities. The claim may be mistaken, of course. Even the principle of the ordering of phonological rules may turn out to be only transmission-significant, as Chafe's (1968) discussion suggests. The risk is acceptable,
for the prize is great, not to be won easily. If such ultimate explanations are to be found at all, it is necessary to focus on the limited set of linguistic structures that are significant to acquisition and consider them simultaneously with the nature of cognitive processes.

Linguistics needs both approaches to the search for universals. There will always be those who will seek the mother lode on the mountains of acquisition-significance, but the fertile plains of transmission-significance will also yield riches to those who will till them.

NOTES

1An earlier version of this paper was written while I was a member of the Stanford Project on Language Universals (NSF Grant GS 1880). I am indebted to comments by J.H. Greenberg, Fred Householder, Edith Moravcsik, and Julius Moravcsik.

2A general review and discussion of the problem may be found in Hempel (1966), Chapter 5.

3Quileute and a few other Salishan and Wakashan languages, according to Ferguson (1966).

4The actual situation is more complex than this schematized account may indicate. See Bell (1970) for a full account.

5A similar principle was enunciated by Greenberg, Osgood, and Jenkins (1966:xxiii):

... no synchronic state can exist which is not the outcome of possible diachronic processes (except perhaps de novo for artificial and pidgin languages) and no diachronic process can be posited which could lead to a synchronic state which violates a universally valid synchronic norm.

6Some of this research will be reported in a paper to be delivered at the LSA meeting in Washington, D.C., December 28-30, 1970.

REFERENCES


