'Out of Chaos Comes Order'; Physical, biological, and structural patterns in phonetics

Peter Ladefoged
Los Angeles, USA

Science is always seeking order; and fortunately for us Nature is always producing patterns out of chaos. In speech this happens at three levels, which we must be careful to keep distinct. At the lower level we can observe the physical patterns, such as the repetitively organized variations in air pressure that arises when sounds are generated. Next there are the patterns that are due to our human biological endowment, such as the tendency for vowels to be dispersed in symmetrical ways. Finally, there are characteristics that are generated by language being a self-organizing social institution. As I will show, these include the appearance of units such as phonemes.

There is not much that need to be said about the physical patterns. We can all appreciate the laws of physics that generate regular movements of the vocal cords when the outgoing breath passes through the suitably adjusted glottis. We can also appreciate the way these same laws generate the complex patterns of air pressure that are produced in the vowel tract as a result of the glottal pulses. The elasticity of the vocal cords, the harmonic structure of formants, and a host of other similar patterned phenomena are now reasonably understood.

At the biological level our knowledge is more meager. But we can explain some of the patterns that occur in language by reference to general principles of human behavior. As a general biological principle, organisms achieve their goal with the least possible effort. We can state this principle equally well the other way round: behaviors that are reinforced and survive are those that use the fewest resources. The notion of intent in achieving a goal is not critical to anything that follows.

Languages exist so that humans can communicate (and so that they can categorize their experiences; this is no doubt important in considering human biological pressures in syntax and semantics, but it is of little relevance in considering the sounds of language). The speaker's goal is to communicate as efficiently as possible. This involves producing a sufficiently distinct sequence of sounds for the listener to be able to get the message in a sufficiently short length of time. Within this aim there are a number of trade offs possible between articulatory effort, auditory distinctiveness, and rate of speech. Usually the speaker is able to take the initiative in setting this balance. Only occasionally does the listener have to interrupt and ask for clarification in some way.

The balance between the conflicting forces is clearly different in different circumstances. Two close friends exchanging information may be able to talk fast, and with a considerable reduction in the degree of precision of the articulatory movements. A speaker addressing an unfamiliar audience on a complex topic may talk more slowly with a more careful articulation. On some occasions, when the listener almost certainly knows the words to expect, virtually no distinctive articulations may be needed. Soldiers on the parade ground can interpret the wordless bellowing of the drill sergeant with great rapidity.

In general, as listeners become more familiar with particular words or phrases, speakers will be able to use more articulatory assimilations. This topic has been well enough covered in the vast literature on historical phonology to need little further documentation here. We might, however, note that as well as obvious assimilations that occur in pronouncing items such as 'handkerchief' and 'in between' as ['hæŋkərtʃif] and [ɪmbətwin], many cases of apparent *dis*similation are actually examples of economy of effort. For example, this is the case for Grassman's law, which states that the first of two aspirated stops in a word will become deaspirated (so that, for example Indo-European *thrikhos becomes classical Greek trikhos 'hair'). This can be interpreted as dissimilation, the consonants in a word becoming more unlike one another. But it is also an example of economy of effort. Aspirated consonants are very distinct from all other sounds (Singh and Black, 1966); but they are also costly in that they use considerable respiratory energy. A word with two such sounds is very costly, and an obvious candidate for pruning in any attempt to reduce the overall effort required for an utterance. Ohala (1981) has given a convincing account of dissimilation in terms of the listener as the source of sound change. I would only add to his account that the conditions are ripe for it to occur because it satisfies the biological drive for economy of communication (not that Ohala would put it that way).

The real cases in which the pressure for auditory distinctiveness affects the sounds of a language are those in which an increase of this kind can be achieved with a relatively low cost in articulatory effort. Many of these affect language paradigmatically, altering the items that can occur at a given place in a phonological structure, rather than syntagmatically, altering items in a sequence as a result of assimilations. The best documented case of the drive for auditory distinctiveness is the dispersion of vowels to fill the vowel space in the most efficient way (Lindblom, 1983). Given that there are a certain number of contrasting vowels in a particular language, it takes little added articulatory effort to increase the distinctiveness of these vowels by dispersing them appropriately. This principle is far from sufficient to account for all the differences observed in languages with similar numbers of vowels. But it is undoubtedly a powerful underlying force.

The drive for communicative efficiency is not the only source of phonetic effects that may be ascribed to the behavior of individual speakers and

listeners. Whenever people speak they identify themselves as belonging to a certain speech community. Every accent of every language is a peculiar set of sounds that is distinguished from all others sets of sounds of all other accents. One cannot speak without an accent of some kind. Part of the function of language is to convey information, part to categorize experience, and part to convey this accentual, sociolinguistic information.

For each of us there is a biological drive to identify oneself as belonging to a particular group, as well as to have one's own idiosyncratic characteristics. My accent is not right for me unless the words have a British RP accent tinged with small Americanisms, and my own particular voice quality. Speakers of every language have to use *exactly* the right vowel and consonant qualities, intonations, rhythms, etc. on pain of being wrongly labeled if they do not. There can be very subtle phonetic differences among languages resulting from this drive to be correctly identified as part of a group; but these phonetic phenomena are important to speakers and listeners. They cannot be ascribed to any general universal principles; they are due to the vagaries of local history and personal desire. But their maintenance can be regarded as ascribable to the behavior of individuals.

We can illustrate the explanatory power of these notions by reference to some recent cross-linguistic studies. In one study of this kind Disner (1983) showed that there were both similarities and differences between the vowel systems of Italian and Yoruba, both of which have seven vowels. The similarities can be ascribed to the fact that both Italian and Yoruba speakers are subjected to the same human drive for greater communicative efficiency. The differences are due in part to the physical laws of sound production affecting speakers of Italian and Yoruba differently, and in part to the biological drive for group identification.

Disner's charts of the formants of a group of 25 Italian speakers and a group of 10 Yoruba speakers are shown in Figure 1. Each vowel is represented by an ellipse (solid lines for Italian and dotted for Yoruba), with its center at the mean for that vowel, its axes along the two principal components of the distribution of the points, and the radii corresponding to two standard deviations. Roughly speaking, therefore, we may say that the ellipses enclose 95% of the points corresponding to each vowel.

Despite some obvious differences which we will discuss in a moment, there is a great overall similarity in the pattern of the vowels in the two languages. In each case the vowels are distributed in a V-shape that takes advantage of the space available. The drive for communicative efficiency results in most of the vowels of each language being fairly distinct from one another.

Some of the differences between the two languages are due to the shapes of the lips of Italian as opposed to Yoruba speakers. The mean first formant frequencies are similar in both sets of vowels, indicating that there are no overall differences in headsize between the two groups of subjects. But, with the exception of /i/ and to a lesser extent /e/, the second formant is lower for the Italian vowels than for the Yoruba vowel. These differences are precisely

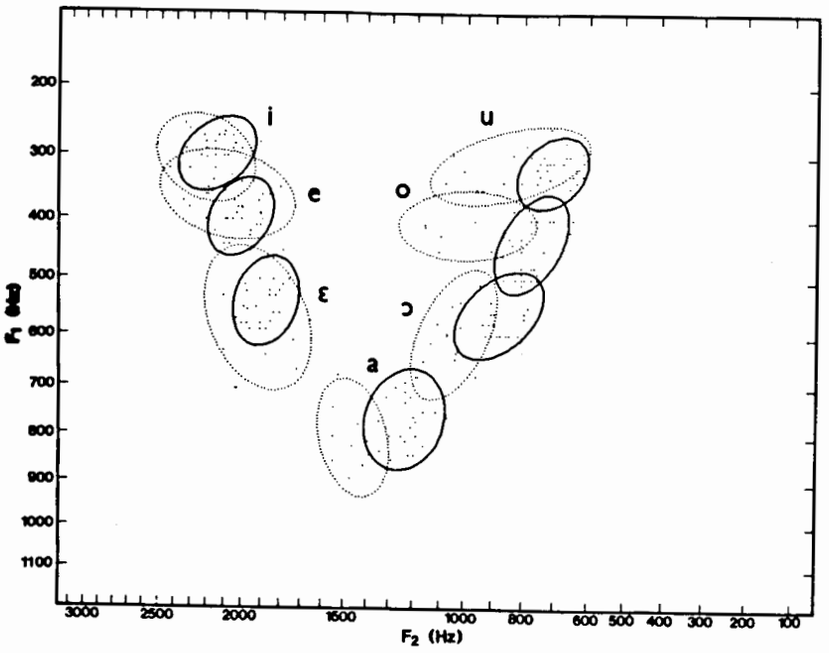

Figure 1. Formant plots of the vowels of 25 speakers of Italian (solid lines) and 10 speakers of Yoruba (dotted lines).

those that one would expect if Yoruba speakers, on the whole, used a larger mouth opening than that used by the Italian. Rounding affects the third formant rather than the second for vowels like [i]; and in other vowels the second (as well as the third) formant is lowered by decreasing the lip aperture. Accordingly, in the case of these two languages, there may be a physical phonetic explanation for at least some of the differences in the formant frequencies between the two groups of speakers. The possibility of overall differences in mouth opening is certainly compatible with the apparent facial differences between speakers of Yoruba and Italian. (This does not of course, imply that a Yoruba could not learn perfect Italian. Any individual speaker could compensate for the overall, statistical, difference in headshape shown in Figure 1, and thus learn perfect Yoruba.)

However as Figure 1 demonstrates, these are far from the only differences between Italian and Yoruba. The most obvious differences are in the F_1 dimension. The vowels of Italian are more evenly distributed than those of Yoruba in which /e/ and /o/ are much closer to /i/ and /u/ than to /ɛ/ and /ɔ/ respectively. The uneven distribution of the Yoruba vowels may be attributed to historical facts concerning the way in which the vowels of the original 9 or 10 vowel system have merged to produce the current 7 vowel Yoruba system. But, nevertheless, this historical explanation does not disguise the fact that present day Yoruba speakers choose to have vowels that are evenly distributed. To some extent they resist the biological pressures for

communicative efficiency that undoubtedly exist, and undoubtedly account for phonetic phenomena that are widespread in the languages of the world, preferring instead the uneven distribution of vowel qualities that characterizes them as Yoruba.

A second illustration demonstrates even more clearly that differences are subject to the whims of fashion. Some articulatory gestures can produce the same sounds - the same acoustic structures - as other quite different movements of the vocal organs; and some languages habitually use one way of producing a given sound, and others another. The evidence on this point is somewhat indirect, but it seems that some languages use one kind of the lip gesture to produce a high back rounded vowel of the [u] type and others produce virtually the same vowel, using a different gesture, with less rounding of the lips being compensated by other vowel gestures such as more lowering of the larynx.

Linker (1982) photographed the lip positions of the vowels in the speech of eight speakers of Cantonese and eight speakers of French. She examined the correlation between formant frequencies and lip position in each of these (and other) languages and showed that there are reliable, statistically significant, differences in the articulatory-acoustic relations. As she says, 'if ...a (mean) speaker of Cantonese wanted to produce an [u] with a given set of formant frequencies, he would have considerably less horizontal opening than a speaker of French producing the same vowel.' She concludes 'These results indicate that languages differ greatly in the lip gestures they use to make the same acoustic distinctions among vowels.'

Presumably children learning to become part of a particular group that speaks a certain language, achieve this goal by watching as well as listening. Children see the lip positions that are typical in their language and learn to make these sounds in that way. There may be a physical explanation for the difference between French and Cantonese associated with the different anatomies of the different racial groups, but it seems unlikely. It is much more likely to be a case of language specific behavior that can be ascribed only to the whims of fashion, and the desire of speakers of each language to maintain their group identity.

There is certainly no anatomical basis for another case in which visual cues during language acquisition are probably the cause of adults having noticeably different articulations when producing acoustically very similar sounds. As has been shown by Ladefoged (1979) a considerably higher proportion of American English speakers use an interdental fricative, as compared with the dental fricative which is more common in British English. In Californian English 75% of the speakers protrude the tongue between the teeth when saying words such as 'thief', whereas virtually no speakers of British English make the fricative in this way. Again, there are no reasons for this other than the desire of British English speakers to do things one way, and American English speakers to behave in another way.

The final illustration concerns coarticulation. We all know that in English,

as in many other languages, the place of articulation of a stop is affected by the articulation of the following vowel. Thus the velar stop in 'keep' has a more forward articulation than in 'cop'. But this difference in the stops is only partially explained by claiming that there is coarticulation with the following vowel. We cannot claim that this is a necessary thing that languages have to do. If this were so, it would be nice, not only from the point of view of making biological explanations, but also from the point of view of simplifying the theory of phonology. Life would be easier if there were a set of universal, language independent, rules that took a string of segments specified in terms of a narrow, allophonic, transcription (or a matrix of distinctive features) and converted it into a sequence of continuously varying parameters of the kind required for a complete description (synthesis) of the utterances represented by the string. But there is no force compelling speakers to have a given degree of coarticulation. As indicated in Ladefoged (1972), it seems likely that French and English differ in the ways in which coarticulations occur. The French velar stops in *pique* [pik] and *Pâques* [pak] coarticulate more with the preceding vowel than the corresponding stops in the English words *peak* [pik] and *pock* [pɑk].

Recently Nartey (1982) has given a more rigorous demonstration of the language dependent nature of coarticulation. He recorded (among other data) eight speakers of Amharic and eight speakers of Navaho. Both languages have, among other fricatives, two sibilants which may be transcribed /s,ʃ/. They also have similar vowels that may be transcribed /i,a/. The fricatives in each language were recorded in the context /i-i/ and /a-a/. The acoustic spectra at the midpoint of each fricative was determined using the UCLA WAVES computer analysis system. These spectra were then convolved with an auditory filter as suggested by Bladon and Lindblom (1981) to produce representations of the auditory information.

The upper part of figure 2 shows the mean auditory spectrum of Amharic [ʃ] in the two vocalic contexts; the lower part shows Navaho [ʃ] in similar contexts. There are differences in the sharpness of the spectrum in the two languages. But, over and above this, it may be seen that the Amharic sounds show greater effects of coarticulation with the vowel, the spectrum having a noticeably higher peak in this Bark scale representation when in the context of [i] than when in the context of [a]. In Navaho there is very little coarticulation between this fricative and the adjacent vowel. Again we see that speakers of different languages choose to distinguish themselves by using different phonetic mechanisms that cannot be predicted from the principle of communicative efficiency, but are simply learned behavior that distinguishes one group from another.

The physical laws of speech production, and, at another level, the drives for communicative efficiency and for group identity account for many rule governed phenomena, But many other phonetic facts about languages cannot be explained in terms of these physical and biological principles. This leads to a difference in emphasis between Lindblom and myself. I do not

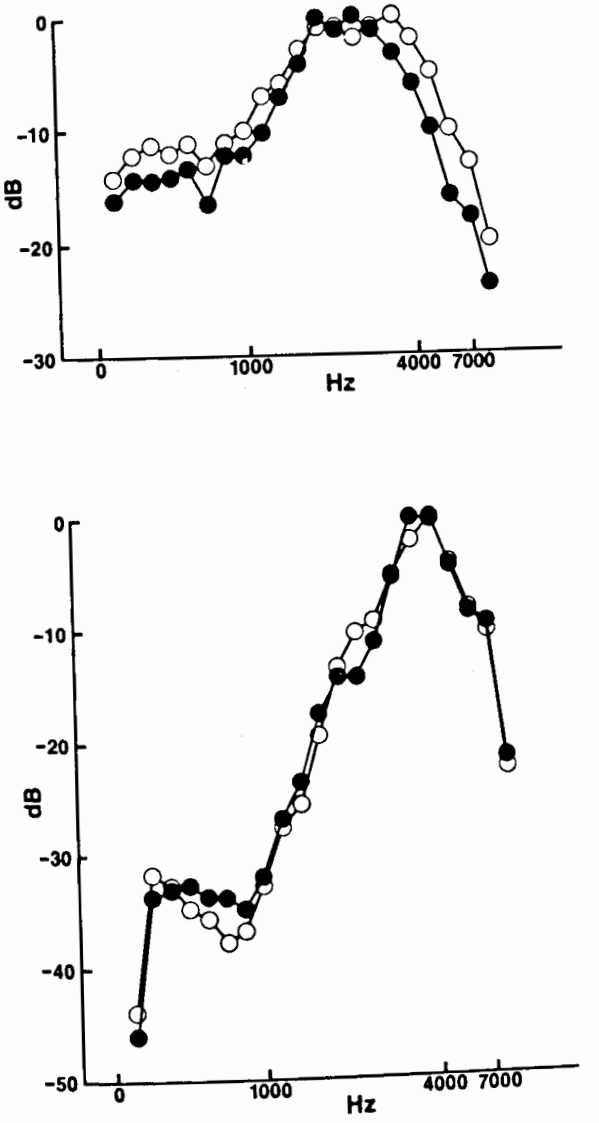

Figure 2. The mean auditory spectra of [ʃ] in the context [i-i] (open circles) and [a-a] (solid points). Data for 8 speakers of Amharic in the upper part of the figure, and for 8 speakers of Navaho in the lower part of the figure.

think that we can directly derive 'the fundamental units and processes of linguistic structure deductively from independent premises anchored in physiological and physical realities' (Lindblom 1983). Instead I agree with Anderson (1981), and other generative phonologists who find that many interesting linguistic observations can be made in formal terms, independent of any other data. I disagree with scholars of the latter group only when they

try to relate these linguistic observations to observable mental capacities. It seems to me that the mental nature of language has been somewhat misleadingly presented by Chomsky (1975). His notion that language is an organ of the mind is not very helpful. It is somewhat like saying that digestion is an organ of the body. Digestion is an ability that involves many components, including some things that are normally called organs such as the liver and the pancreas, as well as a number of other things such as saliva, mastication, and bowel movements. Digestion is like language in that it is a system. But neither of them is an organ in the usual sense.

A better way of describing a language is to consider it as an observable social institution, without having to consider what goes on in people's minds. When we consider any social institution we find that it is governed by different principles from those that govern the behavior of individuals. Principles such as communicative efficiency and identification with a group apply to descriptions of what people do. But a language considered as a system where everything hangs together is to some extent self-organizing.

In order to make this point clear it is worth considering two other examples of self-organizing social institutions. We may begin by comparing a language with a moral code - a system of value judgments applicable in a given community. Any moral code is clearly a product of a society, and is strongly influenced by the surrounding culture. Moral judgements that originally had some utilitarian function rapidly become ritualized. Like pronunciations of words they are as they are because that is the way things are done a a certain society. But morality is also property of an individual, at least to the extent that the individual can choose to perform moral acts. Morality, or at least the capacity for performing moral acts, may even be like language, or the capacity for language, in being innate. Certainly one way to think of morality is as an over-developed herd instinct -- a self-organized, innate drive for the preservation of the herd rather than the individual.

The moral code that we observe (or feel guilty about) is only one example of a social institution. As another, very different, example consider the economic system. There are obvious market forces affecting the price of goods and the cost of labor (the far from inexorable 'laws' of supply and demand). There are also Galbrathian forces such as the conflict between the company management (whose aim is usually growth, which leads to bigger managerial responsibility and salaries) and the company ownership (the shareholders) who want bigger profits, which may well be achieved without growth and with less management). All these forces, and many more (government, international affairs, and perhaps morality) add up to form a social institution, the economic system, which nobody understands and which is certainly not part of anybody's competence. Without people there would be no economic system. It is like language in that it takes at least two to trade. Furthermore, just as people 'know' the rules of their language, in the same sense everyone 'knows' their economic system. We all understand what money can do. But it is obviously ridiculous to take a mentalist approach. Nobody would call economics an organ of the mind.

Many sound patterns are the result of language being a self-organizing social institution, and are not generated at the level of individual behavior. We may start by considering those that result from the filling of a gap in a phonological system. It has often been observed that languages tend to fill holes in the patterns of their segmental inventories. Thus Antilla (1972) shows that Proto-Baltic Finnic had a system 'in which the short vowels had one degree of height more than the long ones, and contained the only front rounded vowel in the whole system ... (Modern) Finnish has filled every single gap and ended up with perfect symmetry.' To take another example, it is not at all surprising for a language such as English, which at one stage had four voiceless fricatives /f, θ, s, ʃ/ and three voiced ones /v, ð, z/, to acquire the missing voiced fricative [ʒ], as we have done recently. But it should be noted that this does not occur because of the biological drive acting on individual speakers and listeners. Filling holes in an abstract phonological system does not increase communicative efficiency for the individual. There is no increase in ease of articulation or auditory distinctiveness for any *existing* possible utterance. Nor, in general, does it help to identify the speaker in any way. (The exception is when the new sound is being brought in by borrowing from some other language or dialect. Using this sound may mark the speaker as wishing to identify with the speakers of the other language or dialect, as, for example, when a speaker of Southern British English starts using a voiceless fricative /χ/ in words of Scottish origin, such as 'loch'.)

The discussion of possible linguistic changes in terms of holes in phonological patterns is sometimes formulated in a slightly different way. There are a number of occasions when linguists talk about the segmental inventories of a language being such that they facilitate or hinder possible sound changes. Thus Maddieson (1984) suggests that if a language does not have /v/ it is more likely to develop a phonemically contrastive /β/ as a result of phonologizing an intervocalic [β] allophone of /b/.

Adding a new sound is like the emergence of a new species in biological evolution. It is possible to claim that it happens because God sees a gap and wants it to be filled because it is easier for the world to be that way. But an equally good claim is that if there is an ecological niche to be filled, events (the random mutation of genes, evolution is not purposive) will conspire to fill it. In the same way a new phoneme is more likely to occur (to be borrowed or to be phonologized from an existing allophone), if it fits nicely into an existing pattern. Note that this cannot be explained in terms of the behavior of individual speakers and listeners, just as the development of a new species is not due to the action of individuals. The communicative efficiency principle does not apply to how languages organise their sounds. When we discuss phonological systems we have moved to considering language not as part of an individual's behavior, but as a self organizing institution.

Many of the patterns currently described by linguists are patterns that occur simply in language considered as a social institution. In addition to the

hole in the pattern phenomena, there are what Kisseberth (1970) calls phonological conspiracies. Dauer (1983) has assembled an excellent case for regarding stress timing in English in this way. Many people have observed an apparent tendency in English for stresses to recur at regular intervals of time. But it seems that this may be due to a fortuitous combination of circumstances. The fact that English words have a somewhat regular stress pattern, the possibility of alternative stress patterns in some words, and of dropping stresses in some sequences of words, the reduction of weak syllables, and the clitic-like nature of many grammatical formations, all these things and more combine to lead to the occurrence of stresses at appropriate intervals.

Perhaps the most startling conspiracy - one that seems to have deceived by far the majority of linguists - is the appearance of phonemes. Accounts of human behavior in terms of phonemes are nearly always examples of what has been called the psychologist's fallacy - the notion that because an act can be described in a given way that it is necessarily structured in that way. As far as I can see, phoneme size units play only a minor role in human behavioral acts such as normal speaking and listening. I have argued this point elsewhere (Ladefoged 1980) and will not consider it in detail here. All that is necessary is to examine some apparent counter arguments.

The first concerns children babbling, playing with newly discovered sounds. Typically they make sounds such as [babababa....dadadada] but sometimes they will produce sequences such as [bibibebe]. Does this mean that they have discovered the notion of a segment, and are trying out different vowels? I doubt it. I think it is much more likely to be random, uncontrolled behavior, and it is only the linguist who hears it as a set of different vowels. But in any case I do not want to deny that there is a strong commonality among all syllables beginning with the same consonant. After all the Aramaic scholars and others who invented syllabaries spotted this and wrote syllables such as [bi, ba, bu] with the same symbol. But the important point is that no child ever babbles sequences such as [badaga badaga]. And for hundreds of years the Aramaic scholars never noticed that there was something in common between syllables such as [ba da ga]. According to Gelb (1952), writing has been invented many times, and there have been many independent developments of syllabaries. At least in some languages the syllable is a clear, intuitively apprehended, concept. But only once in the whole history of writing has it occurred to people to break syllables into separate segments that can be recombined to form other syllables; and that occurred because of fortuitous circumstances. It took the happy chance of the coming together of speakers of Semitic languages, who had symbols for syllables beginning with pharyngeal and other non-IndoEuropean consonants, and Greeks, who did not need these symbols with their original values and chose to use them for vowels. The Greek writing system, the only original alphabet, was produced out of the spare symbols of a syllabary. If phonemes are behavioral concepts, it is difficult to see why the alphabet was invented only once; and, for that matter, why children do not alternate phonemes and babble [badaga badaga].

Speech errors are sometimes cited as another piece of evidence against the claim that phoneme size units play only a minor role in speech production and perception. One of the commonest errors in rapid speech is transposing two elements as in a spoonerism, saying, for example, 'fast pew' instead of 'past few' (Fromkin 1973). In such cases it is often apparently segments that are transposed. But note that it is not just any two segments that get out of order. The previously cited phrase could never have appeared as 'paf stew' with the final element in one syllable becoming initial in another. The segments that move always have to be in the same positions in different syllables. I do not know how speech errors involving transpositions occur. But it is no way obvious that they are simply transpositions of segments.

A final piece of seeming counter-evidence is the existence of a number of language games and secret languages that can be described in terms of phonemic interchanges. For example, Cantonese speakers who are unfamiliar with alphabetic notions can nevertheless produce a form of secret speech in which vowels and consonants are reversed (Wu, personal communication). In this speech words such as [ma] are produced as [am], but this is still not a truly segmental approach; words such as [tʃaŋ] come out as [aŋtʃ]. Speakers do not produce [ŋtʃ] (the reverse of the phonemes), and do not even recognize [ŋaʃt] as the true reversal of the phonetic segmental order. Again what is even more important is that in all the language communities I know of that have secret languages or games of this kind, there are always some speakers who find it very difficult, or even virtually impossible, to follow rules involving segments. (Fluent speakers of Pig Latin often fail to understand how difficult it is for those of us without an equivalent mis-spent youth to talk in this way.) Notions concerning phonemes size units have to be carefully learned - they are in no way naturally available.

Lindblom (1983) has suggested a nice analogy that can be extended to make this point clear. He has pointed out that termite nests appear to the outside observer to have a most intricate structure. There are great pillars and arches that rival those of medieval cathedrals. But it does not follow from this that individual termites know about arches. In fact they are simply following a very straightforward pattern of behavior, governed (in nest building) by a single rule: deposit grains of earth near other grains of earth that are scented with a termite secretion. At first this leads to random depositing of earth. But very soon the deposits are on top of other recent deposits and the pillars grow. As two pillars grow taller the scent is strongest on the sides closest to each other; and so those two sides grow together and form an arch. All from a single, simple, rule. Phonemes may be like arches in termite nests, visible to outside observers, but having no meaningful role in the activity of the individuals producing them. Speech *appears* to be composed of sequences of segments because of the interactions of the different systems of which it is composed. The complex gestures involved in producing syllables have diverse parts that look as if they are categorically distinct. We call these diverse parts vowels and consonants, but we must always remem-

ber that these are just names for readily distinguishable aspects of the stream of speech. Those of us who have been exposed to an alphabetic tradition may be influenced so that we are very conscious of the possibility of describing speech in terms of units of this kind. But illiterates may have little or no concept of speech segments (Morais et al, 1979). Similarly those involved in adult literacy campaigns report that the concept of the segment is far from self evident. Intelligent adults who have been taught to write a few words cannot perform tasks such as naming other words that begin with the same segment (Jackson, 1982). A language consultant who has been working extensively with a linguist will be able to learn the phonemic principle (Sapir, 1949), just as a child can learn to read and write. But this is hardly evidence for phonemic units in the normal process of speaking and listening.

We can carry the termite analogy a step further still. Just because the individual termite cannot be considered responsible for the design of the arches and pillars in a termite nest, it does not follow that it is not interesting to describe these pillars and arches. They are a necessary part of the termite nest; when a pillar or an arch is needed to support the edifice which the community requires, if the nest is to be one which survives, then the requisite structure will be present. In that sense, a termite nest is self-organized. In a similar way a language gets the sounds that it needs. The segments and phonemes are present in the structure of that abstract entity, the language, considered as a social institution. Indeed, as Halle commented several years ago, 'Almost every insight gained by modern linguistics from Grimm's law to Jakobson's distinctive features depends crucially on the assumption that speech [or, in my terms, language] is a sequence of discrete entities.' (Halle, 1964)

But despite the value of segments as descriptive units, it seems almost certain that the phonemic principle is not part of our genetic endowment (as it surely must be for those who view it as an innate ability). The manipulation of phonemes is an acquired ability. Evolutionists teach us that such things are properties of a culture, and not of an individual's physiology. The invention (not, for me, the discovery) of the alphabet occurred far too recently for it to have become part of our DNA. Indeed, as Gould (1981) puts it: '*Homo sapiens* arose at least 50,000 years ago and we have not a shred of evidence for any genetic improvement since then..... All that we have accomplished, for better or worse, is a result of cultural evolution.'

Our endeavors include building (like termites) social institutions such as language, morality, and economic systems. Each has, to a great extent, become its own thing, so that it is no longer entirely explicable in terms of outside forces. The evolution of language has involved its feeding upon itself, so that it must be described partly in terms of unique principles. The current accounts may not be correct. Descriptions of languages in terms of the fashionable metrical phonology (Halle and Vergnaud, 1980) are, after all, at least superficially very different from those of the older generative phonology (Chomsky and Halle, 1968). But the thread of the uniqueness of language

that runs through them is still valid. And they are all, despite their authors' claims, descriptions of social institutions and not explanations of mental activities. Like termites who do not know how to build an arch, ordinary speakers and listeners do not know the sound pattern of English.

References

Anderson S. (1981). Why phonology isn't 'natural'. *Linguistic Inquiry* **12**, 493-589.

Antilla, R. (1972). *An Introduction to Historical and Comparative Linguistics*. New York: Macmillan.

Bladon, A. and Lindblom, B. (1981). Modeling the judgment of vowel quality differences. *J. Acoust. Soc. Amer.* **69.5**, 1414-1422.

Chomsky, N. (1975). *Reflections on Language*. New York: Parthenon Press.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.

Dauer, R.M. (1983). Stress-timing and syllable-timing reanalyzed. *J. Phonetics* **11**, 51-62.

Disner, S. (1983). Vowel quality: the contribution of language particular and language universal factors. *UCLA Working Papers in Phonetics* **58**.

Fromkin, V. (1973). *Speech Errors as Linguistic Evidence*. The Hague: Mouton.

Gelb, I. (1952). *A Study of Writing*. Chicago: University of Chicago Press.

Gould, S.J. (1981). *The Panda's Thumb*. New York: Norton.

Halle, M. (1964). On the bases of phonology. In: *The Structure of Language* (Fodor and Katz, eds), Englewood Cliffs, New Jersey: Prentice Hall, 604-612.

Halle, M. and Vergnaud, J-R. (1980). Three dimensional phonology. *J. Linguistic Research* **1.1**, 83-105.

Jackson, E. (1982). Personal communication.

Kisseberth, C. (1970). On the functional unity of phonological rules. *Linguistic Inquiry* **1**, 291-306.

Ladefoged, P. (1972). Phonetic prerequisites for a distinctive feature theory. In: *Papers in linguistics and phonetics to the memory of Pierre Delattre*. (Valdman, ed), The Hague: Mouton, 273-285.

Ladefoged, P. (1979). Review of J.C. Catford: Fundamental Problems in Phonetics. *Language* **55.4**, 904-907.

Ladefoged, P. (1980). What are linguistic sounds made of? *Language* **56.3**, 485-502.

Lindblom, B. (1983). Economy of speech gestures. In: *The Production of speech* (MacNeilage, ed): New York Springer Verlag.

Linker, W. 1982). Articulatory and acoustic correlates of labial activity in vowels: a cross-linguistic study. *UCLA Working Papers in Phonetics* **56**, 1-154.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

Morais, J. et al. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* **7**, 323-331.

Nartey, J. (1982). On fricative phones and phonemes. *UCLA Working Papers in Phonetics* **55**, 1-181.

Ohala, J. (1981). The listener as a source of sound change. *Chicago Linguistic Society*. Papers from the parasession on language and behavior. 178-203.

Sapir, E. (1949). The psychological reality of phonemes. In: *Selected Writings of Edward Sapir*. (ed. D.G. Mandelbaum) Berkeley: University of California Press. 46-50.

Singh, S. and Black, J.W. (1966). Study of twenty six intervocalic consonants as spoken and recognized by four language groups. *J. Acoust. Soc. Amer.* **39**, 639-656.

Smith, C. and Tager-Flusberg, H. (1980). The relationship between language comprehension and the development of metalinguistic awareness. Paper presented at the Fifth Annual Boston University Conference on Language Development.