JOHN J.OHALA

Phonological Notations as Models*

Phonological notations need major overhaul. I would like to suggest how this can be succesfully done. My argument has four parts.

First, I would like to make a simple distinction — well known in semiotics — between what I will call notations as symbols and notations as models. Notations which merely symbolize the thing they stand for are in all respects arbitrary in their form except insofar as they must be differentiated from other symbols used in the same context. Notations which model the thing they stand for are, in at least one respect, non-arbitrary in their form, that is. in some way they are or purport to be isomorphic with the thing they represent. A few simple examples of these two types of notations are given in Figure 1.

In the case of the notation under the heading ‘symbols’, there is no recognizable part of the symbols that is isomorphic with any part of the entities they stand for. Only in *Playboy* cartoons and for humorous purposes is any' isomorphism suggested to exist between the scientific symbols for 'male' and 'female' and actual males and females. Likewise the graphic symbol '7' does not have seven distinct parts, although the tally mark for '7', on the other side, does. Words, of course, are well known as arbitrary symbols for the things or concepts they stand for. Thus, we do not find it inappropriate that the word 'big' is so small nor that the word 'microscopic' is relatively large. Most mathematical notations are symbols in this sense. However, as for the notations under the heading 'models', there are obvious respects in which they are isomorphic with the entities they represent. Architects' floor plans are isomorphic with the geometry and proportions of the building they represent. The electrical circuit diagram is isomorphic with the topological arrangement and connectedness of the various components. And so on.

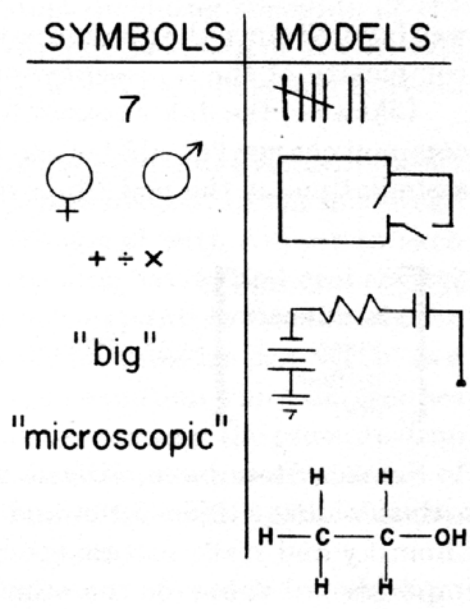

Fig. 1

* This work was supported in part by the National Science Foundation. [This paper originally appeared in 1978 in Proceedings of the Twelfth International Congress of Linguists. Vienna, August 28—September 2, 1977. Ed by W. U. Dressler, W. Meid, O. E. Pfeiffer, T. Herok. Innsbrucker Beiträge zur Sprachwissenschaft. Pp 811-815. This pdf version does not preserve original pagination.]

What we require of a symbol, then, is very simple. What we require of a model — by common practice — is much more, especially in the case of the more elaborate, more complex models used in modern scientific work. Specifically, we expect a model to represent all and only the facts and relations present in the system it stands for so that some of the behavior of the system will be self-evident, i.e., explained.

'Facts' and 'relations' refer to the usual two parts of a model, substantive primitives and formal primitives, respectively. The substantive primitives or 'ingredients' of the electrical circuit, for example, would be the constituent resistors, capacitors, power sources, etc. The formal primitives or 'rules of operation', not represented in Figure 1, would be the rules of current flow, and so on, e. g., Ohm's law, which may have a graphical or mathematical expression.

'All and only' means the representation should be minimally redundant. Naturally primitives irrelevant to the particular task should be omitted from the representation. In addition, all primitives that are mentioned should be completely independent of each other, i. e., 'orthogonal' with respect to each other. A simple corollary of this last point is that primitive entities and derived or predicted entities should be kept strictly apart (cf. Vennemann and Ladefoged 1973).

The ultimate purpose of the model, of course, is to explain or make self-evident some behavior of the system by reference to these primitive entities. Ideally the model should also suggest how it may be tested by making predictions about behavior or events that we have not yet observed but could observe given the proper circumstances.

The second point in my argument is to show that today's phonological notations are generally judged by how well they meet the criteria for models.

In an early paper Halle (1962) quite rightly rejected the use of IPA-type notation in the representation of phonological processes because they would not render self-evident the greater naturalness of (1), for example, as opposed to (2). Feature notation of the same processes (3, 4)

$$\text{(1)}\quad \text{a} \rightarrow \text{æ} \ / \ \_ \ \{\text{i, e, æ}\} \qquad\qquad \text{(2)}\quad \text{a} \rightarrow \text{æ} \ / \ \_ \ \{\text{i, p, z}\}$$

$$\text{(3)}\quad \text{a} \rightarrow [-\text{back}] \ / \ \_ \begin{bmatrix} \text{V} \\ -\text{back} \end{bmatrix}$$

$$\text{(4)}\quad \text{a} \rightarrow [-\text{back}] \ / \ \_ \left\{ \begin{bmatrix} \text{V} \\ +\text{high} \\ -\text{back} \end{bmatrix}, \begin{bmatrix} \text{C} \\ -\text{sonorant} \\ +\text{anterior} \\ -\text{coronal} \\ -\text{voice} \\ -\text{continuant} \end{bmatrix}, \text{etc.} \right\}$$

was favored since the greater naturalness of the one process was reflected in the lesser graphical complexity of the representation and by showing (1) to be assimilatory in character.

Likewise, the Jakobsonian feature system was criticized because it suggests that the very common change [si]->[ši] is one of dissimilation (cf. 5), whereas it is well-known to be a case of assimilation, as the new Chomsky and Halle (1968) features accurately indicate (see 6).

$$\text{(5)}\quad \begin{bmatrix} \text{C} \\ +\text{continuant} \\ -\text{voice} \\ +\text{diffuse} \\ \text{etc.} \end{bmatrix} \rightarrow [-\text{diffuse}] \ / \ \_\_ \begin{bmatrix} \text{V} \\ -\text{grave} \\ +\text{diffuse} \end{bmatrix} \qquad \text{(6)}\quad \begin{bmatrix} \text{C} \\ -\text{high} \\ +\text{anterior} \\ \text{etc.} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{high} \\ -\text{anterior} \end{bmatrix} \ / \ \_\_ \begin{bmatrix} \text{V} \\ +\text{high} \\ -\text{anterior} \\ \text{etc.} \end{bmatrix}$$

Some writers have criticized various of the current feature systems because of their non-orthogonality (Liljencrants and Lindblom 1972). The vowel features [low] and [high] of the Chomsky and Halle feature are highly non-orthogonal because a ' + ' value for one necessarily implies a ' -' value for the other. The features [sonorant] and [voice] are non-orthogonal for other, more interesting, reasons (Ohala 1972).

Other examples could be given but I hope this brief list will suffice to show that phonological notations are expected to model not merely symbolize phonological processes. None of the above criticisms of feature systems would make sense if the notations were only symbols of phonological phenomena.

The third point is to demonstrate that the potential of current phonological notations to adequately model and thereby explain a vast amount of interesting phonological behavior is very poor. Chomsky and Halle (1968: 400ff.) note this themselves when they admit that there is nothing in a representation such as (7) which reveals it to be much more natural than that in (8).

(7) [- sonorant]->[- voice] (8) [-sonorant]->[+ voice]

Unfortunately their solution to this is to posit marking conventions such as (9). But, again, there is nothing in the representation of (9) which suggests in a self-evident way that it is any more natural than (10), its opposite.

(9) [u(nmarked value) voice]→ [-voice] / [____, sonorant]

(10) [u(nmarked value) voice] → [+ voice] / [_____, sonorant]

So we are back where we started: without a representation of the process which makes its greater expectedness understandable. The marking conventions are nothing more than an admission that the existing phonological notation is inadequate to the tasks it is used for.

What is the basic fault of current phonological notations, i.e., the various feature systems ? Wrong choice of features? There certainly have been some bad choices for features (e.g., 'heightened subglottal pressure') but I don't believe that is the most serious problem facing such systems. In fact, the choice of many of the Jakobsonian features show considerable insight into the way speech sounds are structured (cf. Jonasson 1971). Many of the features proposed are non-orthogonal with respect to each other and this is a shortcoming but not a very serious one I would think. The most devastating failure of the various distinctive feature systems, in my opinion, is that they have only a very rudimentary set of formal primitives. The list of features themselves, the substantive primitives, are not a bad start for the representation of sound patterns but to constitute a good model they must be supplemented by the rules of interaction of the parameters. It must be possible to show, for example, in a principled way *why* [-sonorant] segments, i.e., obstruents, tend to be [-voice].

Is there any solution to this problem? Yes, lots of them and this is the final point of my argument. Fant (1960), Flanagan, Ishizaka, and Shipley (1975), Stevens (1971, 1972), Liljencrants and Lindblom (1972), Lindblom (1975) and others have published many interesting and empirically justified models of the speech production system including accounts of the relation between speech articulation and speech acoustics. These models explain or are capable of explaining a great variety of sound patterns (cf. Ohala 1974a, b, 1975a, Ohala and Lorentz 1977).

One quite humble contribution in this area is my own model of aerodynamic events in speech (Ohala 1975b, 1976). Skipping details which can be found in the papers referred to, the model can be characterized as having such substantive primitives (all orthogonal with respect to each other) as: mass of air in the lungs and oral cavity, volumes of the lung cavity and oral cavity, area of the glottis and the oral constriction, magnitude of pulmonic expiratory force, and such formal primitives as a version of Boyle's law which determines the air pressure in a cavity as a function of the mass of air involved, and the volume of the cavity, an equation which determines air flow through a constriction as a function of the area of the constriction and the pressure drop across it, and so on. From such parameters ('features', if one insists) presence or absence of voicing and noise and their intensities can be predicted (see also Scully 1975, Flanagan, Ishizaka, and Shipley 1975). All of this is incorporated in a computer program.

The model can explain certain universal tendencies in the behavior of speech sounds, e. g., that obstruents tend to become voiceless unless something is done to maintain oral pressure at a lower level than subglottal pressure, that is, to maintain a sufficient transglottal pressure drop. (In this case the model simply derives in a formal way what was well known previously.) It also explains why stops before high close vowels have a greater tendency to be aspirated or affricated than those before non-high vowels. This latter tendency is manifested in Japanese, among other languages. It also predicts that high close vowels or glides, especially if palatal, may become noisy (fricated) as is now happening in modern Stockholm Swedish.

The model also predicted something which at the time of its making I had not yet observed, namely, that vowels or glides should not become fricated if there is accompanying nasalization. The reason for this is obvious: frication requires a high rate of air flow through a narrow channel which in turn requires a fairly high pressure drop across this channel; the high pressure needed in back of the oral constriction will not occur if the air vents through the nasal cavity. Is this prediction borne out ? Yes. In Japanese /h/ is manifested as [ç] before the palatal vowel /i/, e. g., /hikari/ 'light' is [çikari], except when it appears in a heavily nasalized environment, e.g., /hiNhiN/ 'horse's neigh' (onomatopoetic) is [çĩŋhĩŋ]. Likewise, in my dialect of American English /h/ is rendered as [ç] before the palatal glide /j/ except when flanked by nasalized segments, in which case it is [h] (or possibly a voiceless nasal), thus: 'ill-humor' is [ɪlçjumɚ] but 'unhuman' is [ə̃nhj̃ũmn̩]. Naturally, we should not expect nasalization to interfere with glottal or pharyngeal frication since the air pressure built up behind their constrictions could not vent through the velopharyngeal opening. This is explicable only if the physical structure of the vocal tract is taken into consideration.

The model just discussed can only account for sound patterns which arise due to aerodynamic (and some simple anatomical) factors in speech. For other sound patterns other models are needed, e. g., those referred to above. Of course, a model can only tell us if certain behavior of the system is logically derivable from what we know about the makeup of the system (the primitives). Inevitably there will be some behavior due to as yet unknown factors; no model can or should predict behavior not due to the primitives built into it. In such cases, however, models serve the very useful function of revealing to us exactly how much we don't know about the system.

My message, then, is this: Existing feature systems are very poor at doing what we expect of them: to make the naturalness of certain sound patterns self-evident. The solution to this problem does not lie in the kind of piecemeal revisions that fill the current phonological literature, e. g., adding a feature here, deleting a feature there, using unary or n-ary instead of binary features, relaxing the requirement of simultaneity of features within a segment, A major overhaul is needed. We should start making use of the many already available physical models of the speech system.

Someone may object that what I have advocated as regards representation of phonological phenomena is suitable for purely physical phonetic facts but modern phonological representations reflect what is in the speaker's head, i.e., what he 'knows' about sound patterns. Nonsense. There is no evidence or plausibility for the claim that naturalness and universality of certain sound patterns is due to psychological as opposed to physical factors. Universal sound patterns are universal because physics and human anatomy and physiology are universal.

John J.Ohala, Dept. of Linguistics, Univ. of California. Berkeley, CA 94720, USA

Bibliography

Chomsky, N. and Halle, M. 1968. The sound pattern of English. New York: Harper and Row.

Fant, G. 1960. Acoustic theory of speech production. The Hague: Mouton.

Flanagan, J. L., Ishizaka, K., and Shipley, K. L. 1975. Synthesis of speech from a dynamic model of the vocal cords and vocal tract. Bell System Technical Journal 54. 485—506.

Halle, M. 1962. Phonology in generative grammar. Word 18. 54—72.

Jonasson, J. 1971. Perceptual similarity and articulatory reinterpretation as a source of phonological innovation. Royal Institute of Technology, Stockholm. STL—QPSR 1/1971. 30—41.

Liljencrants, J. and Lindblom, B. 1972. Numerical simulation of vowel quality systems: the role of perceptual contrast. Language 48. 839—862.

Lindblom, B. 1975. Experiments in sound structure. Plenary address, 8th International Congress of Phonetic Sciences, Leeds.

Ohala, J. .J. 1972. How to represent natural sound patterns. Project on Linguistic Analysis Report (Berkeley) 16. 40—57.

Ohala, J. J. 1974a. Experimental historical phonology. In J. M. Anderson and C. Jones (eds.) Historical linguistics II. Amsterdam: North Holland. Pp. 353—389.

Ohala, J. J. 1974b. Phonetic explanation in phonology. Papers from the parasession on Natural Phonology. Chicago Linguistic Society. Pp. 251—274.

Ohala, J. J. 1975a. Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman, and J. J. Ohala (eds.) Nasalfest. Papers from a symposium on nasals and nasalization. Stanford: Language Universals Project. Pp. 289—316.

Ohala, J. J. 1975b. A mathematical model of speech aerodynamics. In G. Fant (ed.) Speech Communication. Vol.2: Stockholm: Almqvist and Wiksell. Pp. 65-72.

Ohala, J. J. 1976. A model of speech aerodynamics. Report of the Phonology Laboratory (Berkeley) 1. 93—107.

Ohala, J. J. and Lorentz, J. 1977. The story of [w]: an exercise in the phonetic explanation for sound patterns. Proceedings of the Annual Meeting of the Berkeley Linguistics Society 3. 577—599.

Scully, C. 1975. A synthesis study of aerodynamic factors in speech segment durations. In G. Fant (ed.) Speech Communication. Vol. 2. Pp. 227 234.

Stevens, K. N. 1971. Airflow and turbulence noise for fricative and stop consonants: static considerations. Journal of the Acoustical Society of America 50. 1180—1192.

Stevens, K. N. 1972. The quantal nature of speech: evidence from articulatory-acoustic data. In E. E. David and P. B. Denes (eds.) Human communication: a unified view. New York: McGraw—Hill. Pp. 51--66.

Vennemann, T. and Ladefoged, P. 1973. Phonetic features and phonological features. Lingua 32. 61—74.