PHYSICAL MODELS IN PHONOLOGY

JOHN J. OHALA

Phonology studies the patterns or regularities to be found in the sounds of languages, including sound changes and their present-day result, the sound alternations between words. Over a century of categorizing and documenting such sound patterns by numerous phoneticians and linguists has given the field a keen sense of what is ordinary or expected among sound patterns and what is extraordinary or unexpected. For example, it is quite common that [ki], would change to [si, ʃi, or tʃi], however it would be quite exceptional for [pi] to undergo the same change.

Generative Phonology, which seeks to maintain an inverse relationship between the generality of a phonological process and the complexity of its notational representation, requires a way of formally differentiating 'expected' from 'unexpected' sound patterns. For this purpose, the labels 'marked-unmarked' were revived, the foımer applying to unexpected sound processes, the latter to expected ones. It was decided that only 'marked' entities or processes would add to the complexity of the representation and thus it was possible to maintain the most general rules as also the simplest notationally.

It should be clear that such terminological sleight-of-hand really accomplishes nothing. It provides no explanation as to WHY certain sound patterns are expected and others not, it merely provides a new LABEL for this distinction. Phonology has enough resources at hand to be able to progress beyond the essentially descriptivist stage represented by the 'theory' of markedness and to seek instead an explanation for the common tendencies in sound patterns.

It has been a frequent observation that common sound changes found in languages which are typologically, chronologically, and geographically distant must be due to the common physiological and perceptual mechanisms employed by human speakers. An undeıstanding of these sound changes and their resultant sound alternations then, must be gained by studying the physical aspects of the speech process.

For example, vowels frequently become nasalized in the environment of nasal consonants. This is a reasonably well understood process and is due to assimilation by the vowel of the state of the lowered velum for the nasal. However there seem to be many cases of vowels, particularly low vowels, becoming nasalized in the environment of glottal and, possibly, pharyngeal consonants. Witness the frequently observed nasalization in the British R.P. word [hãf] 'half'. Other languages exhibiting this process are Lahu, a Lolo-Burmese language (James Matisoff, personal communication), the Amoy dialect of Chinese, as well as other Sino-Tibetan languages (Matthew Chen, personal communication), East Gurage (Hetzron 1969), etc.

To see if the soft palate behaved in some special way during glottal consonants, a new device was used, the 'nasograph', named in imitation of the photoelectric glottograph since it operates on the same principle. The nasograph (see Figure 1) consists of a light and light sensor encased in a flexible transparent plastic tube which is inserted into the subject's nose and pharynx such that the light is in the pharynx and the light sensor in the nasal cavity. Greater or lesser velar elevation allows less or more light to impinge on the light sensor and thus develop relatively a greater or lesser voltage which can be recorded and related to other speech events.

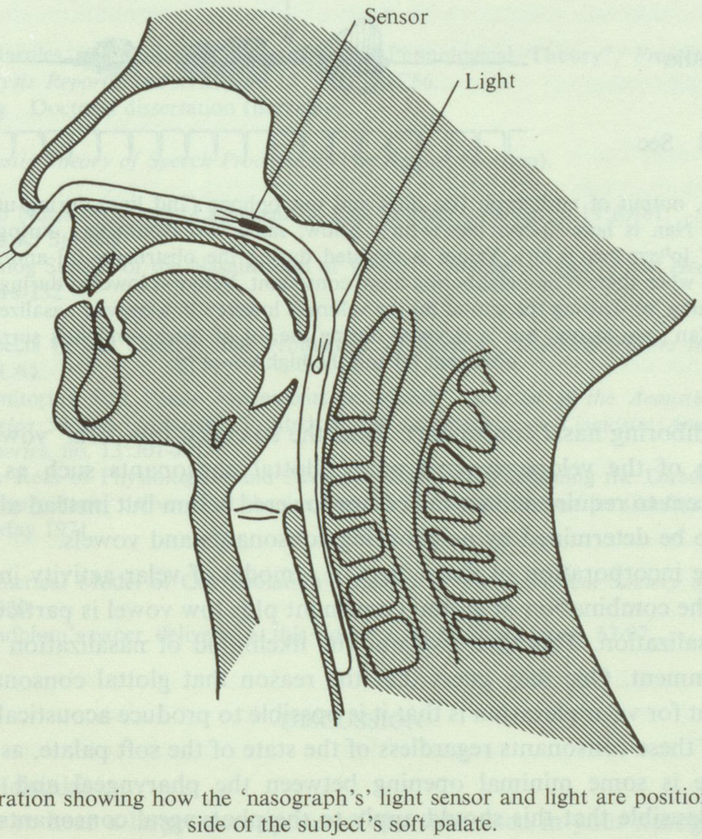

Fig. 1. Illustration showing how the 'nasograph's' light sensor and light are positioned on either side of the subject's soft palate.

As expected (see Figure 2), it was found that the velopharyngeal port must be closed for all obstruents and it must be open for nasal consonants. Also, as has been known, velar height for vowels varies directly with the 'height' of the vowel (in the absence

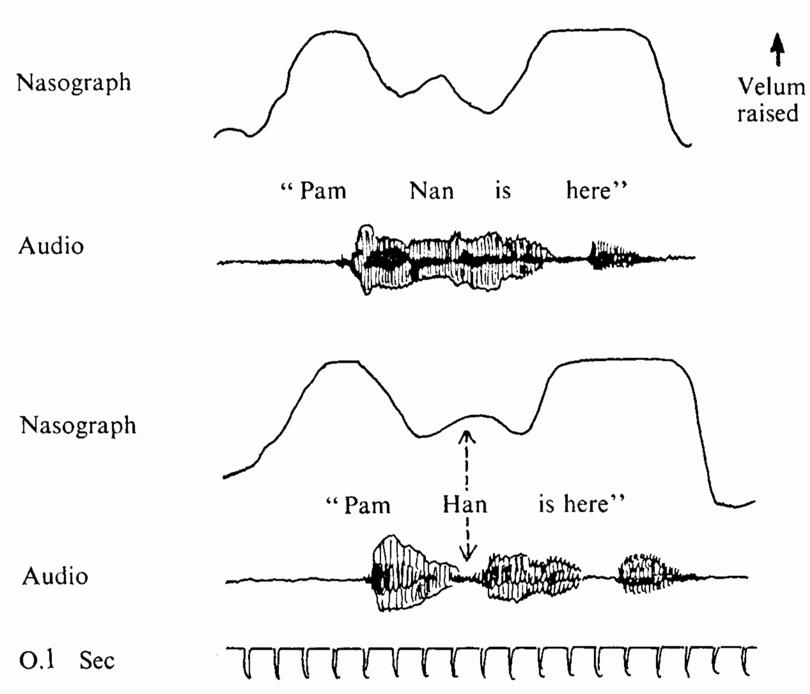

Fig. 2. Top, output of nasograph (1st line) and microphone (2nd line) during utterance of the phrase 'Pam Nan is here' [p^{h}æm næn ɪz hɪr]; below, same two parameters during phrase "Pam Han is here" [p^{h}æm hæn ɪz hɪr]. Velum is elevated during the obstruents [p] and [z] and during high vowel [ɪ] when there is no neighboring nasal consonant. Velum is lowered during [n] and during vowels adjoining [n]. During the glottal [h] the velum is lowered in a heavily nasalized environment (cf. [h] of 'Han', see arrow) but is elevated during the [h] of 'here', which is surrounded by the obstruent [z] and the high vowel [ɪ].

of any neighboring nasal consonant). Thus the so-called low 'oral' vowels may have an opening of the velopharyngeal port. Glottal consonants such as [ʔ] and [h], however, seem to require neither raised nor lowered velum but instead allow the velar elevation to be determined by neighboring consonants and vowels.

From the incorporation of these facts in a model of velar activity in speech, it is clear that the combination of glottal consonant plus low vowel is particularly vulnerable to nasalization. This then explains the likelihood of nasalization appearing in this environment. One may guess that the reason that glottal consonants have no requirement for velar elevation is that it is possible to produce acoustically acceptable versions of these consonants regardless of the state of the soft palate, as long as, perhaps, there is some minimal opening between the pharyngeal and oral cavities. It seems plausible that this should apply to the pharyngeal consonants as well.

Many more examples could be offered of the utility of physical models in providing explanations for the commonness of certain sound changes and sound patterns (House and Stevens 1956; Fant 1960; Öhman 1967; Ohala 1970, 1971a, 1971b; Chen 1971, forthcoming, and the paper presented by Professor Lindblom at this Congress,

see p. 63-97, etc). Such models have a clear advantage in explanatory power over those models implied by the various distinctive feature representations. The latter attempt to account for all sound patterns with the same small set of features (parameters), many of which are complexly interdependent but, which, given their placement in a two-dimensional matrix, are not easily recognized as such. Further they use features which reflect only the articulatory or only the acoustic-auditory aspect of speech. The physical models referred to have no such arbitrary limitations and are free to incorporate in themselves all and only those physiological and acoustic facts which are relevant to a particular phonological pattern.

Department of Linguistics
University of California, Berkeley

REFERENCES

Chen, M.
1971 "Metarules and Universal Constraints in Phonological Theory", *Project on Linguistic Analysis Reports, 2nd series*, no. 13: MC1-MC56.
forthcoming Doctoral dissertation (Berkeley).
Fant, G.
1960 *Acoustic Theory of Speech Production* (The Hague, Mouton).
Hetzron, R.
1969 "Two Notes on Semitic Laryngeals in East Gurage", *Phonetica* 19:69-81.
House, A. and K. Stevens
1956 "Analog Studies of the Nasalization of Vowels", *Journal of Speech and Hearing Research* 21:218-232.
Ohala, J.
1970 "Aspects of the Control and Production of Speech", *Working Papers in Phonetics* 15 (UCLA).
1971a "Monitoring Soft Palate Movements in Speech", *Journal of the Acoustical Society of America* 50:140 (Abstract). Printed in full in: *Project on Linguistic Analysis Reports, 2nd series*, no. 13:J01-J015.
1971b "The Role of Physiological and Acoustic Models in Explaining the Direction of Sound Change", Paper delivered at the 1st Annual California Linguistics Conference, Berkeley, 1-2 May 1971.
Öhman, S.
1967 "Numerical Model of Coarticulation", *Journal of the Acoustical Society of America* 41: 310-320.
See also B. Lindblom's paper, delivered at this Congress, to be found on pp. 63-97.

DISCUSSION

SOVIJÄRVI (Helsinki)

I would like to hear a larger description of the tube used in your nasography. Why did you not apply a smaller tube?

OHALA

The tubing is 4 mm. (O.D.) but it should be possible to make it even smaller. The end is sealed and is swallowed into the esophagus for stabilization and subject comfort. In general there is little discomfort experienced by subjects using this device.

SMITH T.S. (San Diego, Calif.)
Can the records from the 'nasograph' which you describe in this paper be calibrated? That is, can you determine the actual amount of velopharyngeal opening using this device?

OHALA
Tell me what other measure of velopharyngeal opening you would accept as valid, e.g., nasal air flow, distance between velum and nasopharyngeal wall as measured from cine-X-rays, etc., and that technique can be used simultaneously with the nasograph to calibrate it. I have not yet done this, but it should be done, of course. I doubt very much that the signal is linearly related to velopharyngeal area; however I have 'faith' that the relation is at least a monotonic one, although, again, my faith should be tested. It does give a good indication of the TIMING of velar movements and this is quite valuable.

SUSSMAN (Austin, Tex.)
There are strain-gauge techniques available now that monitor the movements of the velum; have you ever compared the output of your nasograph with strain-gauge records?

OHALA
No, I have not compared the output of this device with that obtained by a strain-gauge device.

VANDERSLICE (New York)
Would you comment on the possibility that some changes of light transmission may not indicate velic opening at all, but rather a change in the pressure or place of the velic closure such that light transmission is affected?

OHALA
In fact, it is quite likely that there are degrees of closure of the velopharyngeal part, e.g., tighter closure for [i] than for [e] although both would have an air-tight closure, as such, and that a varying amount of light would impinge on the light sensor as a result.

GRAHAM STUART (Silver Springs, Md.)
I am much impressed with Mr. Ohala's experimental work and its results as he has reported them. His concluding remarks, however, awaken in me a certain anxiety that important basic principles may be in danger of confusion. He proposes that progress in phonological science requires physiological and acoustic models providing continuous detail rather than discrete distinctive feature models. I am convinced that there can be no question of 'rather than': both are needed. Physical, neuro-physiological, and acoustic models describe in its various aspects the signalling variable, the values of which are the signalling elements available for encoding message categories. The message categories themselves (and ten distinctive features are the simplest of these) are at the CHOICE of the speaker. The phonological model must deal both with the structure of message category choice and the way in which that choice is encoded in the speech signalling variable.

OHALA
I did not say and did not mean to imply that the various distinctive feature notations or any system representing the facts of speech via a two-dimensional matrix are completely useless. Obviously they have their uses, but these do not extend to attempting to account for the naturalness of certain phonological processes. Phonologists genuinely interested in explaining common sound tendencies would waste less time and energy if they turned to speech models better able to incorporate the known facts of the physiology and perception of speech.