63

Frontiers of Speech Communication Research

Edited by

B. LINDBLOM

Department of Phonetics, Stockholm University, Sweden

and

S. ÖHMAN

Department of Linguistics, Uppsala University, Sweden

1979

ACADEMIC PRESS
LONDON NEW YORK SAN FRANCISCO
A Subsidiary of Harcourt Brace Jovanovich, Publishers

The Contribution of Acoustic Phonetics to Phonology

J. J. OHALA

Phonology Laboratory,
Department of Linguistics,
University of California,
Berkeley, Calif., USA

Introduction

In the nineteenth century the life sciences themselves came to life, commencing a tremendous acceleration in scientific accomplishment that has not slowed to this day. This came about due to a judicious injection into the field of experimental techniques and the application of the principles of the "hard" sciences, physics and chemistry, to the problems of living organisms (Bernard, 1865; Helmholtz, 1877).

There are signs today that phonology may soon experience the same type of explosive growth and for the same reasons. Experimental techniques have been successfully applied to classical phonological problems in the three traditional domains or spheres of language: the sociological domain (Labov, 1966), the psychological domain (Greenberg and Jenkins, 1964), and the physical domain. It is in the latter area that the application of the principles of physics, especially acoustics, has provided satisfying explanations for a variety of widely-attested sound patterns (Durand, 1956; Fant, 1973; Stevens, 1972; Liljencrants and Lindblom, 1972; Lindblom, 1975; Jonasson, 1971; Ohala, 1974, 1976, 1977; Wright, forthcoming).

In order to further demonstrate the promise of this approach I will briefly review here two cases in which common, frequently-observed behaviour of speech sounds may be explained by reference to principles of acoustics. All the cases will involve speech sound alternations manifested as allophonic variations, sound changes and/or morphophonemic variations. In fact, these three all represent the same process caught at different stages: today's allophonic variation leads to tomorrow's sound change which, in turn, may lead to morphophonemic variation. I assume that all these sound alternations came about due to what Sweet (1888: 238) calls "[false] acoustic imitation" i.e. the substitution of sound X by one speaker, the imitator, for original sound Y in the speech of another. The problem, then, is to find out why sounds articulated differently may nevertheless be similar acoustically.

Labial Velars are Both Labial and Velar*

There is extensive evidence that although labial velar sounds such as [w, ɥ, k͡p g͡b] act sometimes as labials and sometimes as velars, with little apparent systematicity, there are certain cases in which their behaviour is rather strictly determined.

(a) When becoming nasals or determining the place of articulation of assimilating nasal consonants, labial velars generally act as velars, but
(b) when becoming fricatives (or stops with burst releases) or determining the place of articulation of an assimilating fricative, they generally behave as labials.

Although some exceptions can be found, data of the sort in (c) and (d), which exemplify these patterns, can be found over and over again in the phonological descriptions of languages all over the world.

(c) Tswana passive verb stem formation (Cole, 1955).

verb root	+ passive suffix	→ passive verb stem	
-bala	+ wa	→ -balwa	"to read"
but:			
-roma	+ wa	→ -roŋwa	"to send"
-akaɲa	+ wa	→ -akaŋ:wa	"to think"

(d) Sentani [h] is optionally realized as [f] or [ɸ] after [w] (Cowan 1965) e.g. kew + hike → (optionally) kewfike "he threw away".

The explanation for (a) requires reference to some of the basic factors which create resonances and anti-resonances in the vocal tract. Resonances are determined by the geometry of those airways in the vocal tract which constitute the most direct route from the sound source to the point where the sound radiates. Anti-resonances are contributed by any airways which are cul-de-sacs branching off from this main airway. As shown in Fig. 1, where schematic vocal tract shapes for [m], [n], [ŋ], and [w̃] are given, the most direct route from the sound source (the glottis) to the radiation point (the nostrils) is

* Further data and discussion of this topic may be found in Ohala and Lorentz (1977a, b) and Ohala (forthcoming).

via the pharyngeal and nasal airways and is largely identical for all nasal consonants. This path is marked by the filled circles in the figure. The primary distinguishing characteristic of the different nasals, then, is the variation in the geometry of the oral cavity which is a cul-de-sac branching off from the main (pharyngeal-nasal) airway and which therefore creates anti-resonances whose frequencies, to a first approximation, depend on the effective length of this cavity. The effective length, of course, is the distance measured from the pharyngeal airway to the point of constriction in the oral cavity. In the case of

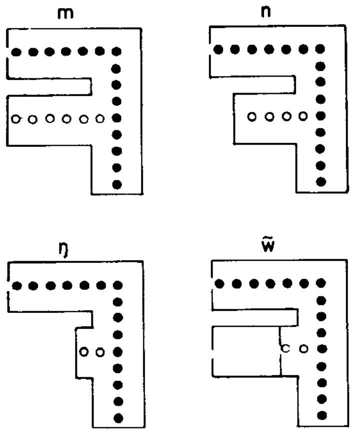

Fig. 1. Schematic representation of the vocal tract shapes for the sounds [m], [n], [ŋ] and [w̃]. Filled circles indicate airways contributing resonances of the sounds; open circles indicate airways contributing anti-resonances.

multiple constrictions in the oral cavity, as in [w̃], it will be the rear-most constriction, provided it is small enough, which will mark the boundary of the cul-de-sac. The cavity lengths contributing the anti-resonances of these nasals are marked in Fig. 1 by open circles. As can be seen, the acoustically relevant configuration of the vocal tract for [w̃] is most like that for [ŋ], least like that for [m].

For similar reasons [mʲ] or the sequence [mi] are disposed to change to (be misheard as) [n] or [ni], respectively (see below). In this case there are constrictions at both the labial and the palatal regions but it is the rear-most one, the palatal constriction, which determines the effective length of the oral side cavity. Thus [mʲ] will sound like a nasal with a constriction in a region further back than the labial position.

To explain (b) we must determine why, when there are two noise sources in the vocal tract, as in a fricativized labial velar, the noise from the rear source is perceptually less salient than that from the front source. The most probable reason, it seems to me, derives from two acoustic facts: first, that fricative noise contains primarily high frequencies and, second, that the oral cavity and small labial opening in front of the velar constriction in a labial velar constitute an effective low-pass filter. Thus the noise produced at the velar constriction will be severely attenuated by the resonating cavity in front of it whereas the noise produced at the labial constriction will not be subject to any significant attenuation.

Palatalization of Labials*

A survey of the phonological literature turns up many independent cases of the sound pattern exemplified in (e).

(e) [pj, pʲ] → [t, ts, ʧ]
[bj, bʲ] → [d, dz, dʒ]
[mj, mʲ] → [n, ɲ]
etc.

Some examples of this are given in (f) through (h).

(f) Czech (data from Bělič, 1966; Andersen, 1973).

Standard Czech	East Bohemian	English gloss
[mjɛstɔ]	[nɛstɔ]	"town"
[pʲɛt]	[tɛt]	"five"
[pʲi:vɔ]	[ti:vɔ]	"beer"
[pʲɛknʲɛ]	[tɛknʲɛ]	"nicely"

(g) Italian (data from Jaberg and Jud 1928–1940; transcription simplified).

Roman dialect	Genoese and neighboring dialects	English gloss
[pjeno]	[tʃena]	"full"
[pjanta]	[tʃanta]	"to plant"
[er fjato]	[uʃa]	"breath"
[bjaŋko]	[dʒaŋku]	"white"

(h) Classical Greek (data from Meillet and Vendryes, 1924).

Pre-Classical Greek	Classical Greek	English gloss
*gʷam-yo	βαίνω (cf. Latin venio)	"I come"
*kom-yo- (cf. Latin cum)	κοινός	"common"
*χαλεπ-yω	χαλέπτω	"provoke"
*θαφ-yω	θάπτω	"bury"

* Further data and discussion of this topic may be found in Ohala (1978 and forthcoming).

To gain a preliminary understanding of why such alternations occur, it is instructive to look at spectrograms of palatalized labials and compare them to those of plain labials and dentals. Figure 2 shows tracings of spectrograms published in Fant (1960) of the Russian CV syllables [ba], [b^ja] and [da]. In examining these spectrographic patterns it is necessary to keep in mind the fact that place of articulation cues for consonants reside in both the second formant (F_2) transition and in the noise burst. This being the case, it is interesting to note in Fig. 2 that the F_2 transition for the palatalized labial is more similar to that for the dental than it is to that for the plain labial. No doubt in most instances the noise burst from the release of the stop is a sufficient cue to the labiality of the palatalized labial in spite of the dental-like

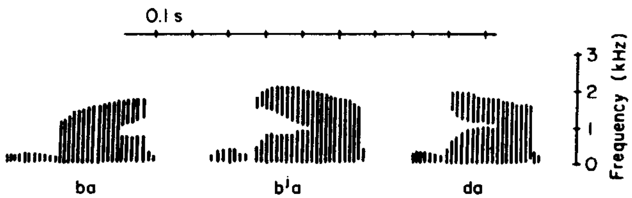

Fig. 2. Tracings of spectrographic patterns for the Russian syllables [ba], [b^ja] and [da] (from Fant, 1960).

F_2 pattern. If a listener missed the noise burst cue, however, the consonant would very likely be taken for a dental. Moreover, the impression that such stops were dentals or palatals would be reinforced by any fricative noise generated from the rush of air through the narrow palatal constriction.

However why should the palatal constriction, a secondary articulation, have a greater influence on the consonantal F_2, than the labial constriction, the primary articulation? The beginnings of an answer to this question can be gleaned from the nomogram in Fig. 3 (from Fant, 1960). The nomogram shows the formant frequencies that would result as one varies both the position of the constriction in the vocal tract and the accompanying labial constriction. As revealed in Fig. 3, even though F_2 is generally susceptible to change due both to variations in place of constriction within the vocal tract and to variations in the labial constriction, F_2 frequency is largely independent of the degree of labial constriction in the dental-palatal region (see arrow). A palatal constriction, then, although a secondary articulation will be the primary determinant of the F_2 frequency.

The acoustic similarity of palatalized labials (or labials followed by or coarticulated with palatal vowels) and dentals is further evidenced by the results of various speech perception studies e.g. Lyublinskaya (1966), Winitz *et al.* (1972), House (1957) and Gay (1970). Lyublinskaya, for example, in research on the confusability of VC transitions (i.e. where there were no consonant releases to aid identification) found palatalized labials were up to 30% more likely to be confused with dentals than were plain labials.

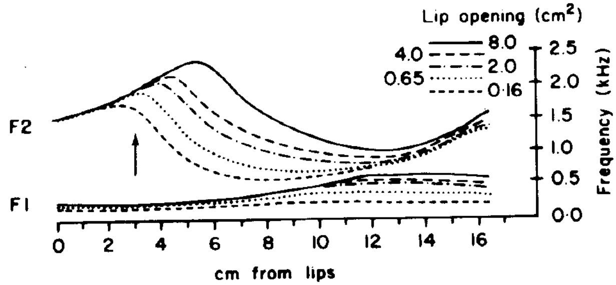

Fig. 3. A nomogram (from Fant, 1960) showing the frequencies of the first two formants (F_1 and F_2) that would result from constrictions in various places in the vocal tract (horizontal axis) and from variations in the lip aperture (the parameter). The arrow marks the approximate position of a palatal constriction.

Conclusion

Many other examples could be provided of questions about the origin and directionality of widespread sound patterns which have yielded to the application of principles of acoustics, aerodynamics, or other areas of physics. Of course, the explanations provided here are mere sketches of the more complete, more formal explanations which would ultimately have to be couched in quantitative terms and integrated into an expirically verifiable model of the speech processes. Fortunately, such models already exist. For this purpose, there is more of lasting value to the phonologist in Gunnar Fant's "Acoustic Theory of Speech Production" than there is in any other single work.

Of more importance to the practice of phonology, though—or to linguistics as a whole—is not the subset of problems that can be successfully dealt with through physics—for physics is likely to be of little help in solving problems of a sociological or psychological character—but rather what the solution of these problems teaches us about the notion of explanation and about the role of experiment in arriving at explanations.

In the current phonological literature common sound patterns are "explained" by statements of the sort "obstruents tend to devoice because it is normal (literally, "the unmarked case") for obstruents to devoice". Or, the fact that, diachronically, contrasts in initial position are more likely to be preserved than those in final position is accounted for by statements such as "initial position is stronger than final position", where "stronger" simply means "has the property of preserving relatively more contrasts". Whatever value such statements may have in systematizing phonological data, they are worthless as explanations since they reduce to the tautology "X because X". In contrast, the explanations offered above and those in the literature referred to, have the logically acceptable form "X because Y" where Y, the *explanans*, is distinct from X, the *explicandum*.

Finally, it matters how one establishes that, in fact, X follows from i.e. is a consequence of Y. As the Renaissance scientists demonstrated, one cannot rely on tradition, fashion, majority vote or the authority or eloquence of the person or persons making the claim. Rather, repeatable controlled experiments, which are nothing more than careful observations, provide the means to do this. By finding or creating the circumstances where Y exists and all other potential confounding factors are eliminated, one can observe whether or not X does in fact appear as predicted. No further justification needs to be given for this, the experimental method, than the fact that it has proved to be successful in numerous fields over the past 3 or 4 centuries. Throughout history practitioners in various fields and subfields have claimed that their subject matter was so different from others that they were exempt from using experiment to advance knowledge in their discipline. However, in the seventeenth century Newton proved that the forces governing the motion of celestial bodies was the same as those which held on earth, thus demonstrating that even theories in astronomy must be constrained by experimental results. Also, in the nineteenth century, Bernard, Müller, Helmholtz and others showed that the life sciences could no longer profitably ignore experimental techniques.

It is to be hoped that in this century, the implications of this trend will not be lost on phonologists, or linguists in general.

Acknowledgements

This research was supported by the National Science Foundation. The study of labial velars was originally done in collaboration with James Lorentz. I also gratefully acknowledge helpful advice from Madison Beeler, Cecil Coker, Carol Justus, Yakov Malkiel and the members of the Phonology Laboratory, University of California, Berkeley.

References

Andersen, H. (1973). Abductive and deductive change. *Language* **49**, 765–793.

Bělič, J. (1966). K otázce zaniklého litomyšlského *t*, *d*, *n* za někdejší *p*', *b*', *m*'. Slavica Pragensia 8, Acta Univ. Carolinae, Philologica, 1–3.

Bernard, C. (1865 [1957]). "An Introduction to the Study of Experimental Medicine" (Transl. by H. C. Green). Dover, New York.
Cole, D. T. (1955). "An Introduction to Tswana Grammar." Longmans, Green, London.
Cowan, H. K. J. (1965). "Grammar of the Sentani Language." Martinus Nijhoff, The Hague.
Durand, M. (1956). Du rôle de l'auditeur dans la formation des sons du langage. *J. de Psych. Norm. et Path.* **52,** 347–355.
Fant, C. G. M. (1960). "Acoustic Theory of Speech Production." Mouton, The Hague.
Fant, C. G. M. (1973). "Speech Sounds and Features." MIT Press, Cambridge, USA.
Gay, T. (1970). Effects of filtering and vowel environment on consonant perception. *J. Acoustic Soc. Am.* **48,** 993–998.
Greenberg, J. H. and Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word* **20,** 157–177.
Helmholtz, H. (1877). Das Denken in der Medicin. English translation: Thought in medicine. *In* "Selected Writings of Hermann von Helmholtz" (Kahl, R., ed.). Wesleyan University Press, Middletown, pp. 340–359.
House, A. S. (1957). Analog studies of nasal consonants. *J. Speech Hear. Disorders* **22,** 190–204.
Jaberg, K. and Jud, J. (1928–1940). Sprach und Sachatlas Italiens und der Südschweiz. Zurich.
Jonasson, J. (1971). Perceptual similarity and articulatory reinterpretation as a source of phonological innovation. Speech Transmission Laboratory (Stockholm) 1971/1, pp. 30–41.
Labov, W. (1966). "The Social Stratification of English in New York City." Center for Applied Linguistics, Washington, USA.
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, 8 Int. Congr. Phon. Sci., Leeds.
Lyublinskaya, V. V. (1966). Recognition of articulation cues in stop consonants in transition from vowel to consonant. *Soviet Physics-Acoustics* **12,** 185–192.
Meillet, A. and Vendryes, J. (1924). "Traité de grammaire comparée des langues classiques." Librairie Ancienne Edouard Champion, Paris.
Ohala, J. J. (1974). Experimental historical phonology. *In* "Historical Linguistics II: Theory and Description in Phonology" (Anderson, J. M. and Jones, C., eds). Proc. 1 Int. Conf. hist. Ling., Edinburgh. North-Holland, Amsterdam, pp. 353–389.
Ohala, J. J. (1976). A model of speech aerodynamics. *Report of the Phonology Laboratory (Berkeley)* **1,** 93–107.
Ohala, J. J. (1977). Phonological notations as models. To appear in: Proc. 12 Int. Congr. Ling. (Reprinted in: *Report of the Phonology Laboratory (Berkeley)* **2,** 38–46.)
Ohala, J. J. (1978). Southern Bantu vs. the world: the case of palatalization of labials. *Proc. Annual Meeting, Berkeley Ling. Soc.,* **4,** 370–386.
Ohala, J. J. (forthcoming). On the origin of sound patterns in language.
Ohala, J. J. and Lorentz, J. (1977a). The story of [w]: an exercise in the phonetic explanation for sound patterns. *Proc. Annual Meeting, Berkeley Linguistics Society,* **3,** 577–599.
Ohala, J. J. and Lorentz, J. (1977b). Acoustic reasons why labio-velars are both labial and velar. *J. Acoust. Soc. Am.* **61,** S32.

Stevens, K. N. (1972). The quantal nature of speech: evidence from articulatory-acoustic data. *In* "Human Communication; a Unified View" (David, E. E. and Denes, P. B., eds). McGraw-Hill, New York and London, pp. 51–66.

Sweet, H. (1888). "A New English Grammar." Oxford.

Winitz, H., Scheib, M. E. and Reeds, J. A. (1972). Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *J. Acoust. Soc. Am.* **51**, 1309–1317.

Wright, J. T. (forthcoming). The effect of nasalization on vowel perception.