Explanation in Phonology: Opinions and Examples

John J. Ohala



INTRODUCTION

There has recently been much commentary on the questions of whether it is possible to explain the sound patterns in language and, if so, what form those explanations will take — in particular, what facts and principles will be referred to from which the sound patterns follow; see Dinnsen (1980), Ladefoged (1980, 1983). Lass (1980), and Anderson (1981). I appreciate these writers' views and have learned much from them, but I find them much too pessimistic. In this paper I will offer what I believe to be a more hopeful alternative.

ARGUMENTS AGAINST EXPLANATION IN PHONOLOGY

Ladefoged questions our ability to account for certain language-specific phonetic facts by reference either to abstract psychological or physical phonetic entities. He rather endorses the notion that 'many interesting linguistic observations can be made in formal terms, independent of any other data' (1983:89).

Anderson allows that some aspects of phonology can be and have been explained by reference to phonetic and other facts, e.g., psychological, social, but doubts that all matters of interest to the phonologist, especially the very important ones, can be satisfactorily understood by reference to data which he regards as "external" to phonology.

Dinnsen accepts that phonetic explanations may be possible in principle but does not accept that any of those offered in the current literature meet the criteria for "true" explanations, which, according to him, must be deductive nomological (D-N) explanations, that is, law-like statements of cause and effect deduced from previously known facts and principles.

Lass goes farther than the others by arguing that D-N explanations, which every self-respecting scientific discipline is said to strive for, are not possible in *principle* when it comes to accounting for language change. This hits phonology 'right where it hurts' because sound change is one of the major areas in which explanations have been offered. Dinnsen and Lass present similar arguments for their positions: given a language exhibiting a certain sound change, is it possible, they ask, that a D-N explanation can be constructed, according to which the sound change *had* to happen? That is, can all

¹ Or, for Dinnsen, a phonological rule, which amounts to the same thing since rules are generally necessitated by alternations created by sound change.

the necessary and sufficient conditions for the sound change be identified? The answer, of course, which no one would dispute, is 'no'. Under what seems to be identical circumstances, one language will undergo a certain sound change and another will not.

COUNTERARGUMENTS

I agree with Ladefoged that there are facts about speech behavior which

are not yet explainable in any way and especially not by many of the currently fashionable phonological models. I also grant that we should not cease to make purely descriptive statements about any aspect of language structure even though we don't yet have an explanation for it. What strikes me as unfortunate about Ladefoged's views is that he seems to take the present primitive state of affairs in phonology as a reason to avoid attempts at explanation and instead to focus almost exclusively on pure description. I think that current research in phonology gives us reason to be optimistic that explanatory accounts for

sound patterns not only are possible but are presently at hand.

Neither Ladefoged nor Anderson have demonstrated — à la Gödel and Heisenberg — any inherent block to our ultimate understanding of sound patterns in physical, psychological, and/or social terms. Rather, they simply present various examples — case studies — which they believe resist explanation by reference to things outside of phonology². There is no proof that the cases they cite will resist such explanation forever. There is therefore no basis for debate here, except, perhaps, to show that some of their cases can be explained by principles they would consider 'outside' phonology and to point out certain weaknesses in argumentation.

Anderson, for example, cites the case of alternations in Fula, such as those in (1), as providing 'the crucial evidence... that phonological distinctions do not always correspond directly to phonetic observables' (1981:502).

(1) /war/ "kill"
$$\sim$$
 /-bar/ \sim /mbar/
/war/ "come" \sim /-gar/ \sim /ngar/

He claims that these alternations justify setting up two underlying /w/'s, one [+anterior] and the other [-anterior]. He dismisses as irrelevant the account which claims that these alternations are the residue of historical processes which did have a phonetic basis at their origin and which are now maintained in the language as fossilized morphological variants (i.e. where original /b/ and /g/ in word initial position both shifted to /w/ due to their being 'grave' — having a low second formant; see Ohala & Lorentz, 1977; Ohala, 1979). In support of his contention he points to the 'internal coherence'

 $^{^2}$ I do not accept this division of the universe into 'inside' and 'outside' phonology but I will not make an issue of it here; see Ohala (1983b, in press a).

of these alternations and their apparent stability (although no independent evidence is cited for this point). The same, of course, could be said about virtually any alternation, e.g. those which resulted from Grimm's Law, Verner's Law, etc., as in English father/paternal, foot/pedal, brother/fraternal, tooth/dental, heart/cardiac, horn/cornea, etc. This is just what is meant by the term 'alternation', i.e. a regular correspondence of sound in different environments. Thus Anderson's argument boils down to this: the evidence for his analysis of the alternation is the fact of the alternation itself. In other words, he begs the question; this and similar cases he covers are therefore without any logical force.

The error in Dinnsen's and Lass' reasoning is quite fundamental: they believe that D-N explanations exist. Outside of mathematics and logic, however, D-N explanations do not exist. Physics and chemistry are often held up to us as the paradigm disciplines capable of D-N explanations. Such, at least, is the impression created by philosophically naive text-book writers and by the physics and chemistry public relations effort. It does not affect my point that many physicists and chemists themselves have fallen prey to this same propaganda. But it is not true; these physical scientists do not literally discover the 'laws' of nature. I do not begrudge physics and chemistry and related disciplines the credit they richly deserve for penetrating the mysteries of the physical world and therefore taming it. But that accomplishment by itself in no way justifies characterizing them as 'the exact sciences'. Anyone who has spent some time doing physics or chemistry in the laboratory (even in secondary school) or reading the primary literature in these disciplines (as opposed to others' 'digested' accounts of scientists' work) will realize that strictly law-like accounts are not possible. As a way of arguing this point let us imagine the following Gedankenkampf between a linguist and a physicist.

The linguist is presented with an isolated island populated with a few hundred monolingual speakers and is challenged to predict how their language will change within a given period of time. As suggested above, the linguist will not be able to do this with any accuracy. At best, she could specify a range of possibilities for the language's future and rule out an even larger range of impossibilities. If we score the Kämpferinnen with a 0 if they are inaccurate and 1 if they are accurate, then the linguist would get a score of 0. Now the physicist is challenged to predict the path of a billiard ball on a billiards table. The physicist would insist on knowing beforehand a) whether the ball will be hit, b) if so, with what force and at what angle, c) that after the ball was hit it would not be subject to any 'outside' forces, e.g., movements of the table, strong air currents, etc., d) that the ball will be perfectly round, the rim of the table of known and uniform elasticity, etc., etc. Naturally, none of this can be specified or guaranteed; this contest is to take place in the 'real world' not in the pages of a textbook. The physicist will also fail her challenges and

will also get a score of 0. So the contest ends in a draw. The moral of the story is this: neither one can insulate their problems from all the unpredictable and uncontrollable causal factors that impinge on their respective universes of observation.

Since D-N explanations (predictions) are impossible in any scientific discipline, phonologists need not feel discouraged that they cannot produce them. What phonology can achieve — and in this it does not differ from other scientific disciplines — is deductive probabilistic (D-P) explanations which include appropriate statements as to the limited degree to which they hold. Lack of perfect knowledge of the universe also means that we will not be able to explain everything, even in D-P terms. This may seem to be a trivial observation, but part of Dinnsen's and Lass's dissatisfaction with some of the recently offered phonetic explanations in phonology is that these accounts cannot explain everything. This is an unattainable requirement — by any science — and therefore in no way invalidates what has been done so far. Rather than focusing on what cannot be done in a given discipline, attention should be given to what it can do, in particular, the 'fruitfulness' of its explanations — in the sense suggested by Louis Pasteur:

"...the characteristic of erroneous theories is the impossibility of ever foreseeing new facts; whenever such a fact is discovered, those theories have to be grafted with further hypotheses in order to account for them. True theories, on the contrary, are the expression of actual facts and are characterized by being able to predict new facts, a natural consequence of those already known. In a word, the characteristic of a true theory is its fruitfulness." [From Vallery-Radot, 1911: 243].

This is precisely what has happened in the mature sciences, e.g. physics. Since pre-historic times, people had a great many questions about the workings of the physical world — the motion of bodies, the transfer of heat, the causes of weather, etc. But it was during the scientific revolution in the 16th century that the physical sciences experienced their first significant success in getting answers to any of these questions — in this case, those about the motion of bodies. This was done by a mixture of rational and empirical means. Their work was motivated by a hypothesis, i.e. an act of reason, but at the heart of their method were the great pains taken to get high quality evidence in support of their explanatory hypotheses; high quality in the sense of being minimally subject to the extraneous distortions which dilute its evidential value. Careful, controlled, observations — that is what experiments are — are the way to minimize such distortions.

The same methods and, as it turned out, some of the same physical concepts showed their fruitfulness by eventually leading to the answers to questions on the nature and behavior of heat. But by no means have all of the questions about the workings of the physical universe been answered. We still do not fully understand the origin of weather and everyone knows that it can't

be predicted very accurately. Are there critics within physics who say the whole discipline is bankrupt and not a respectable science because it hasn't solved the problem of weather? Obviously not. Yet this is the essence of some of Dinnsen's and Lass' criticisms of currently offered explanations in phonology. Whether in physics or phonology, the successful (partial) solution of even a few problems and the *fruitfulness* of these solutions is enough to show that appropriate logical and empirical methods are being used.

CANDIDATE EXPLANATIONS

I would now like to consider in some detail some candidate explanations for sound patterns. I say 'candidate' to emphasize that there is no claim that these explanations are 'correct' because, as outlined above, absolute certitude turns out to be an elusive goal. We know from the history of science that theories once thought to be 'true' have been replaced or revised by subsequent theories. There is no reason to expect that whatever theories we propose today will escape the same fate. It is, to repeat, the empirical support and the fruitfulness of the theories, not their 'truth', which mark valid explanations.

Voiceless laterals revisited. I would first like to reply to Lass' criticism of an explanation proposed in Ohala (1974). I tried to explain there why Norwegian /sl/ became / \int l/ whereas /sn/ remained unchanged. I suggested the reason was that [] is acoustically similar to [] (this was demonstrated by comparing their spectra) such that the following stages in the change were plausible: /sl/>[s]l>/ \int l/. If the /n/ had become voiceless in the same environment, the voiceless nasal that would result would have had very low intensity and be quite unlike any oral fricative.

Lass faults this explanation (39–42) by pointing out that the voiceless lateral in Welsh is regularly interpreted by non-Welsh speakers in Britain as [f], $[\theta]$, or even [x] but rarely as $[\]$. In Ohala (1974) my purpose was to cite phonetic reasons for the different behavior of $[\]$ and $[\]$ (and on the basis of this, the different behavior of [s] before these segments). Lass raises a completely different question in asking why the preferred re-interpretations of Welsh $[\]$ should usually be the weak fricatives [f] and $[\theta]$. There are a number of possible hypotheses one might plausibly entertain (and which could be systematically investigated if anyone deems the issue to be of sufficient importance). First it should be obvious that the measure of 'acoustic similarity' or 'distance' between sounds is a continuous one; [f], $[\theta]$, and $[\]$ are all more or less close to $[\]$; for that matter [s] is certainly closer to $[\]$ than, say, [r] or [s], although more distant than $[\]$. Also, which sound is confused with another depends on which of the many phonetic features which characterize the sounds the listeners take to be most important. $[\]$ and $[\]$ have similar spectral structure,

[1] and [0] or [f] are similar in having relatively low intensity. Listeners with different language backgrounds, e.g., Norwegian and English, may attend to different cues. There may therefore be no incompatibility between the facts of the Norwegian case and those cited by Lass. Second, the [1]'s in the two languages may be phonetically different by virtue of their appearing in different environments; it is not unreasonable to expect that the Norwegian [1] appearing after an [s] may have more intensity than the word-initial [1] in Welsh. Moreover, they may just be phonetically different (i.e. in any environment; see Ladefoged, 1980 for evidence of crosslanguage phonetic differences). Third, there may be English-specific phonotactic constraints which would bias English ears to interpret /1/ (phonetically [1] 1) as the permissible cluster [fl] and not [fl]; different constraints may have applied in the case of Norwegian. There are many avenues to explore if one has the patience and resources.

Lass also offers as counterevidence the fact that /sl/ clusters show no evidence of changing to /ʃl/ in English. As indicated above, this type of argument is based on the false premise that an empirical discipline can possess such perfect knowledge of the universe as to be able to predict (or post-dict) the course of events.

It is worth mentioning that the interaction of [] and [] or (what is phonetically much the same thing) [4], is not limited to Norwegian; it also shows up in Chadic (Newman, 1977).

Given the limits of our current knowledge, Lass' data do not contradict and therefore refute the account I presented 10 years ago (which was itself a repetition of the account given by Haugen, 1942).

Spontaneous nasalization. Normally, distinctively nasalized vowels derive from sequences of vowel+nasal consonant, as given in (2).

- (2) Latin ventus > French [vã]; Sanskrit danta 'tooth' > Hindi [dãt] On occasion, however, nasal vowels (or sometimes nasal consonants) appear in words which never had a nasal consonant at any point in their history. (3) gives a couple of examples.
 - (3) Sanskrit sarpa 'snake' > Hindi [sãp]; Prakrit pahuccai 'attain' > Hindi [pəh \tilde{o} t \tilde{o}]

These are cases of so-called spontaneous nasalization. As it happens, most such cases appear adjacent to consonants characterized by heavy airflow, e.g. voiceless fricatives, aspirated stops, affricates.

My colleague, Mariscela Amador, and I attempted to test a hypothesis that I had made earlier, that vowels produced with a slightly open glottis might mimic the acoustic and perceptual effects of nasalization (Ohala 1975.

1980). A slightly open glottis creates a branch off the main oral resonating cavity in much the same way that the nasal cavity constitutes such a branch. They create similar acoustic effects, namely lowering of the relative amplitude of the first formant and increasing its bandwidth (Fant, 1973: 8; Fujimura & Lindqvist, 1971). Several previous instrumental studies had shown that high air flow consonants like [s] or aspirated stops have greater-than-normal glottal opening and that this condition is assimilated to some extent by neighboring vowels (Slis, 1970).

To see whether physiologically oral vowels might sound nasalized on those portions immediately adjacent to voiceless fricatives, we used digital means to create a series of steady-state vowels by repeating single periods from the ends of vowels before an /s/ and, for the control cases, before the lateral /l/ and the nasal /n/ (to make sure that this method of creating vowels from single periods did not by itself introduce spurious nasalization or remove nasalization actually present). We presented these vowels to listeners to judge the degree of perceived nasalization on a 7-point scale, where 1 meant 'completely oral' and 7 'heavily nasalized'. We also recorded velic elevation to make sure that the vowels near /s/ and /l/ were not physiologically nasalized, and we sampled oral air flow, a rough indicator of glottal area, to verify that only the vowel before /s/ had significant glottal opening.

The perceived nasalization from vowel stimuli made in this way from the syllables /bal/, /ban/, and /bas/ produced by one of the four speakers was 3.2, 5.2, and 5.4, respectively, thus confirming the hypothesis that portions of vowels abutting voiceless fricatives may sound nasalized. We concluded that sound changes showing spontaneous nasalization came about when vowels that 'sounded' nasalized, even though they were not, were re-interpreted and spoken with actual nasalization. (For further details, see Ohala, 1983b.)

DISSIMILATION: THE LISTENER'S 'FAULT'. In a couple of previous papers I have presented a new account of dissimilation (Ohala, 1981, 1983a, in press b, c). Perhaps the most well known example of dissimilation is Grassmann's Law (so-called), as exemplified in (4), where the diacritic '..' marks breathy-voicing.

(4) Proto-Indo-European *bend- 'bind' > Sanskrit band-.

I suggest that *listeners* are responsible for dissimilation. Through experience they learn that certain speech sounds introduce extraneous distortions on adjacent sounds (i.e. due to assimilation). They therefore formulate corrective rules which factor out these distortions — so they can figure out the pronunciation intended by the speaker. Dissimilation occurs when a listener invokes such corrective rules inappropriately, that is, factors out

aspects of pronunciation that were intended features of the pronunciations. Dissimilation is thus a form of low-level hypercorrection. Given a form such as /bend/ I suggest that the breathiness spilled over onto the segments intervening, yielding [bend], and that the listener misinterpreted the breathiness at the beginning of the word as being a distortion caused by the segment at the end. The listener therefore factored it out.

That listeners have and use these corrective rules has been demonstrated experimentally. One such study that I conducted with some of my colleagues (Ohala, Kawasaki, Riordan, & Caisse, forthcoming, as reported in Ohala, 1981) involved requiring listeners to identify synthetic steady-state vowels on a continuum between /i/ and /u/. Somewhere in the middle of the continuum listeners stop hearing /i/ and start hearing /u/. However, this crossover point is further front when the vowels appear with flanking apical consonants (/s-t/) than with flanking labials (/f-p/). The listeners presumably know that apicals (but not labials) cause the vowel /u/ to be fronted and therefore accept as /u/ a more fronted vowel in the apical environment.

Other explanations of dissimilation often appeal to articulatory principles, e.g. Ladefoged (1983) (à propos of Grassmann's Law):

Aspirated consonants ... are very costly in that they use considerable respiratory energy. A word with two such sounds is ... an obvious candidate for pruning in any attempt to reduce the overall effort required for an utterance.

This conception of dissimilation, however, fails to explain why in the majority of cases it is the first of two similar sounds which suffers dissimilation, whereas I account for this by the fact that assimilation, which feeds dissimilation, is predominantly anticipatory (Javkin, 1979: 74ff). Furthermore, Ladefoged's account is difficult to reconcile with cases of dissimilation involving place of articulation — see (5) —, vowel quality, and tone, where it does not seem possible to identify the 'costlier' of two sounds.

(5) Ancient Chinese *pjam 'diminish' > Cantonese pin

My account would predict that the only features which should be subject to dissimilation would be those which are known to spread by assimilation over adjacent segments. This would include the features in (6a), which include place of articulation, but not those in (6b).

- (6) a. labialization
 pharyngealization
 retroflexion
 glottalization
 aspiration
 palatalization
 place of articulation
- b. fricative-ness stopped-ness affricate-ness

By and large, these predictions are borne out (see Ohala, 1981, 1983a, in press b, c, for further data, discussion, and references).

A potentially troublesome counterexample is the apparent dissimilation of voicing among the stops in IE roots, since although voicing spreads to adjacent segments, it is not known to spread to such an extent as to camouflage the distinctiveness of voicing of a whole CVC syllable. However, Gamkrelidze and Ivanov (1973) and Hopper (1973) have suggested that this series of stops was really glottalized — i.e. having a feature well known to dissimilate, e.g. in Quechua and Lahu. These reanalyses — mine of the nature of dissimilation and Gamkrelidze's of the IE stop system — are thus mutually supporting and, one may hope, will have fruitful consequences.

CONCLUSION

If one accepts the pessimistic views discussed above regarding the necessity or the possibility — now or ever — of trying to formulate explanatory accounts of sound patterns, if one accepts the view that phonology is different from other disciplines and is therefore exempt from the usual standards of scientific evidence, then one also must be content with the endless stream of new labels and new notations which either masquerade as explanatory constructs or deliberately forestall the search for them. One will then also have to be content with arguments supported by the most ambiguous evidence — ambiguous in the sense that it is open to many other interpretations. I believe this would be a grave mistake and that this would seriously retard progress in our field.

Looking at the history of various sciences, it seems that one discipline after another eventually recognizes the necessity of adhering to the standards developed in the 16th century — first in the discipline of physics. Chemistry joined the fold around the end of the 18th century, thanks to the efforts of Lavoisier, Priestley, Berzelius, Mendeleev, and others. Physiology — in the face of strong objections from the vitalists, who asserted that the mechanisms of living things would never be reduced to physical processes — nevertheless succeeded in doing just that in the first half of the 19th century, due to the efforts of Claude Bernard, Pasteur, Helmholtz, and others. Ivan Pavlov and others helped to make psychology a rigorous discipline in relatively recent times.

It seems to me that the more difficult the problems studied by a discipline, the greater the delay in its achieving some initial success which finally convinces its practitioners that it should conform to the accepted rigorous standards of empirical science. Physics was the first because it studied (then) the simplest and most easily controlled of problems. Psychology is the newcomer because it studies the most difficult problems. Nevertheless, the *same* basic methods work on all problems, even phonological ones.

Phonology undoubtedly deals with very difficult problems but none are so difficult or so different in character as to bar or exempt it from conducting its research as other sciences do.

REFERENCES

- ANDERSON, STEPHEN R. 1981. Why phonology isn't "natural." Linguistic Inquiry 12. 493-539.
- DINNSEN, DANIEL A. 1980. Phonological rules and phonetic explanation. Journal of Linguistics 16. 171-191.
- FANT, GUNNAR. 1973. The acoustics of speech. Speech sounds and features, by Gunnar
- Fant, 3-16. Cambridge: MIT Press.

 FUJIMURA, OSAMU and JAN LINDQVIST. 1971. Sweep-tone measurements of vocal-tract characteristics. J. Acous. Soc. Am. 49. 541-558.
- GAMKRELIDZE, THOMAS V. and V. V. IVANOV, 1973. Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse. Phonetica 27. 150–156. HAUGEN, EINAR. 1942. Analysis of a sound group: sl and tl in Norwegian. PMLA
- 57. 879-907. HOPPER, PAUL J. 1973. Glottalized and murmured occlusives in Indo-European.
- Glossa 7. 141–166.

 JAVKIN, HECTOR R. 1979. Phonetic universals and phonological change. Report of
- the Phonology Laboratory, Berkeley, No. 4.

 LADEFOGED, PETER. 1980. What are linguistic sounds made of? Language 56.
 485-502.
- LASS, ROGER. 1980. On explaining language change. Cambridge University Press.
- NEWMAN, PAUL. 1977. Lateral fricatives ("hlaterals") in Chadic. Papers in Chadic linguistics, ed. by Paul Newman and Roxanna Ma Newman, 107–119. Leiden: Afrika-Studiencentrum.
- OHALA, JOHN J. 1974. Phonetic explanation in phonology. Papers from the Parasession on Natural Phonology, ed. by A. Bruck, R. A. Fox, & M. W. LaGaly, 251–274. Chicago: Chicago Linguistic Society.

- Speech and language. Advances in basic research and practice, Vol. 3, ed. by N. J. Lass, 75–97. New York: Academic Press.
- 1981. The listener as a source of sound change. Papers from the Parasession on Language and Behavior, ed. by C. S. Masek, R. A. Hendrick, & M. F. Miller, 178-203. Chicago: Chicago Linguistic Society.

- 1983a. The direction of sound change. Abstracts of the Tenth International Congress of Phonetic Sciences, ed. by Marcel P. R. van den Broecke & Antonie Cohen, 253-258. Dordrecht: Foris.
- XIIIth International Congress of Linguists, August 29-September 4, 1982, Tokyo, ed. by Shirô Hattori and Kazuko Inoue, 232-243. Tokyo [Distributed by Sanseido Shoten].
- In press a. Consumer's guide to evidence in phonology. Phonology Yearbook. In press b. Phonological evidence for top-down processing in speech
 - perception. Invariance and variability of speech processes, ed. by Joseph S. Perkell et al. Hillsdale, NJ: Lawrence Erlbaum Associates. —— In press c. Linguistics and automatic speech processing. New systems and architectures for automatic speech recognition and synthesis (NATO Advanced
- Studies Institute, Bonas, France, July 1984), ed. by Renato De Mori & Ching Y. Suen. Berlin: Springer-Verlag. OHALA, JOHN J., HARUKO KAWASAKI, CAROL RIORDAN, and MICHELLE CAISSE. Forthcoming. The influence of consonant environment upon the percep-
- tion of vowel quality. OHALA, JOHN J. and JAMES LORENTZ. 1977. The story of [w]: an exercise in the phonetic explanation for sound patterns. Berkeley Linguistics Society, Proceedings, Annual Meeting 3.577-599.
- SLIS, IMAM H. 1970. Articulatory measurements on voiced, voiceless and nasal con-
- sonants, Phonetica 21, 193-210, VALLERY-RADOT, RENÉ. 1911. The life of Pasteur. [Transl. by R. L. Devonshire.]

London: Constable.