15

Around *Flat*

John J. Ohala

1. INTRODUCTION

In this chapter I first discuss some of the uses of the feature *flat* and then, using *flat* as an example, explore the role of phonological feature systems in general.

2. WHAT *FLAT* WAS AND IS

Although the term *flat* appears in phonetic descriptions as early as 1855 (Müller) as a characterization of nonfront vowels, it appears that Trubetzkoy (1939/1969:127ff.) was the first to employ the term *flat* to describe the auditory quality of retroflexes (and other sounds). Nevertheless, the concept itself can be found in impressionistic phonetic descriptions in many places and at many times in history, referred to by such colorful terms as "hollow," "dark," and "heavy." In its modern usage (in the Jakobsonian feature system), the feature *flat* is defined as sounds manifesting "a downward shift of a set of formants or even of all the formants in the spectrum (Jakobson, Fant, & Halle 1963:31) or, alternatively, as sounds "characterized by a downward shift or weakening of some of their upper frequency components" (Jakobson & Halle 1956: 31). In articulatory terms, the feature refers to sounds that are distinctively labialized, retroflexed, or pharyngealized (or, possibly, velarized or uvularized). In Jakobson's system all features are binary, and the opposite of *flat* is simply *plain* or *nonflat*. However, the exclusive binarity of features is imposed rather arbitrarily and artificially on this feature, and it is more realistic to view it as one end of a continuum: sounds having a downward shift of their higher resonances at one extreme and those having an upward shift, Jakobson's feature *sharp*, at the other. This continuum interacts with the features *grave–acute* (concentration of energy in low- versus high-frequency region) in interesting ways (detailed below), which suggests that in some cases it might be justified to regard all these features as being points on a single continuum.

3. SOME INTERESTING FEATURES OF THE FEATURE *FLAT*

Flat is interesting for a number of reasons. First, it is an example of a feature that started out as a completely impressionistic term, not even very rigorously defined, which has evolved to a point that plausible acoustic correlates can now be specified for it. Second, these acoustic correlates can to some extent be derived from, that is, shown to be natural consequences of, its articulatory correlates. Third, it demonstrates in an even more dramatic fashion than is possible with *grave*, another acoustically defined feature with discontinuous articulatory correlates (labial and velar but not the intervening apical and palatal places of articulation), how essential it is to keep not only the articulatory but also the acoustic correlates of speech sounds in mind when trying to figure out why speech sounds behave the way they do—a point insisted on by Ladefoged (1971a) but neglected by many other phonologists.

Fourth, there is an interesting sense in which *flat* is inherently nonorthogonal with respect to other features at the acoustic–auditory level. Assuming a feature system like the Jakobsonian one, *flat*, in general, does not need to be assigned to a segment until after values for *grave* and *diffuse* have been assigned. This, however, is not just an artifact of the Jakobsonian system; as Stevens (1980) has emphasized, there is a strong tendency for languages not to utilize distinctively labialized, retroflexed, or pharyngealized sounds until after they have gotten near-maximal use out of the features *grave* and *diffuse*, that is, the features that determine the primary places of articulation. One might at first think that this is a matter of definition: One cannot have distinctively labialized sounds unless there are nonlabialized ones to contrast with. But the tendency is phonetically true as well. In languages with small segmental inventories, the consonants present are not phonetically labialized or pharyngealized, and if a language has one apical stop, it is generally not retroflex. Stevens suggests that there is a relatively small number of ways in which speech sounds can differ from one another and that some of these ways are more robust than others; for example, the large abrupt amplitude modulations characterizing sounds labeled 'consonantal' are more salient than the rather slow and spread-out acoustic modulations typical of labialization, retroflexion, palatalization, pharyngealization, and voice quality. Languages, as it were, first utilize segments made from the robust set and only optionally use the less salient features. The same

principles guide the construction of a snowman: fingers, bracelets, and earrings, are optional, but if one chooses to have them, one must have first chosen to make arms and ears. The implication of this is that speech sounds—some of them, at any rate—do not occupy a perceptual space with orthogonal dimensions.

3.1. *Flat* Permits Useful Phonological Generalizations

Since *flat* has a single acoustic–perceptual correlate but multiple articulatory correlates, Jakobson et al. (1963) were able to predict that no more than one of the articulatory phonetic manifestations of *flat* could be used distinctively in any given language. Although this prediction is apparently not absolutely true since there are languages (e.g., Abkhaz) that have contrasts between labialized and nonlabialized uvulars and pharyngeals (Catford 1977b; Colarusso 1974), it certainly is statistically true.

On the other hand, it would also be predicted that these secondary articulations made in quite distinct parts of the vocal tract would often be used together to enhance the "flatness" of some sounds vis-à-vis non*flat* sounds. In agreement with this, it has been noted that in some dialects of Arabic the pharyngealized or "emphatic" consonants are accompanied by lip rounding (Lehn 1963; however, see Ghazeli 1977:74ff.).

It may be for this reason or because of the inherent auditory similarity of all *flat* segments that Arabic pharyngealized sounds are often changed to labialized segments when Arabic words are borrowed by other languages (Jakobson 1962). Representative data (from Leslau 1957) showing borrowing into Argobba, a Semitic language of Ethiopia, are given in Table 15.1.

One of the most remarkable co-occurrences of labialization, retroflexion, and pharyngealization can be found in the American English vowel [ɝ] or in some versions of its associated glide [ɹ]. These segments are somewhat unusual in that their retroflex tongue configuration can apparently be

Table 15.1
ARABIC PHARYNGEALIZED SEGMENTS BORROWED AS LABIALIZED SEGMENTS CONTRASTED WITH PLAIN SOUNDS BORROWED UNCHANGED[a]

	Standard Arabic	Argobba	Translation
Pharyngealized	*ṣabijj*	*s*ʷ*abijj*	'baby'
	ṣadaqa	*s*ʷ*adaqa*	'alms, charity; death commemoration'
Nonpharyngealized	*risala*	*risāla*	'letter'
	dʒism	*dʒism*	'body'

[a] Data from Leslau (1957).

substituted by an extreme bunching of the tongue near the palate in a region that is approximately halfway between the maximal constrictions of the vowels [i] and [u] (Uldall 1958; Delattre 1971). Not coincidentally, the most obvious acoustic correlate of [ɚ] or [ɹ] is an extremely low F_3 (Peterson & Barney 1952; Lehiste & Peterson 1961; Lehiste 1964).

Possible further evidence for the similar character of *flat* segments is that there is a certain amount of overlap in the effects they create on neighboring vowels. Labial or labialized consonants such as [w], often cause front vowels, especially [i], to become more or less backed, centralized, and/or rounded, for example, [i] → [y]; on the other hand, they have relatively little effect on back rounded vowels (Ohala 1979a, 1979b). Retroflex consonants can do the same: Kirton (1967) reports that retroflex consonants in Anyula shift /i/ to [e] but have no noticeable effect on /a/ or /u/. Hercus (1969) indicates that retroflex consonants in Wembawemba (and other languages of Victoria, Australia) shift /i/ and /e/ approximately to the vowels [y] and [ɜ], respectively, but cause only slight centralization to /u/. In Arabic, pharyngealization causes noticeable lowering of /i/ and backing of /a/ but has much less influence on /u/ (Ghazeli 1977; Card 1983).

The similar effect that *flat* segments have on vowels may be responsible in some way for the fact that in Ancient Chinese the retroflex initial consonants and labial final consonants both resulted in the merger of adjacent tense and lax vowels (Hashimoto 1973), although why labial initials and retroflex finals did not do the same is still to be explained.

The extreme effect that *flat* sounds have on segments like [i], that is, those with high F_2, is a reflection of the fact that *flat* segments are at the opposite end of the tonality scale from *acute* or *sharp* segments. Or, rather, to avoid circularity, this interaction is the motivation for opposing the phonological feature *flat* to these other features. This relation also gives rise to the prediction that the distinctiveness of *flat* would be most evident on or near *acute* segments. Accordingly, pharyngealization in Arabic seems to be distinctive primarily on apical consonants; claims that pharyngealized labials or postvelars exist in some dialects of Arabic are, at best, controversial (see arguments in Ghazeli 1977, chapter 7, against this latter view). For similar reasons, one would predict that if optimal segmental sequences should be made from sounds that are as different as possible from each other acoustically, then *flat* sounds should not be found adjacent to other sounds that have low tonality. Although this is not absolutely true, there is a strong tendency among languages in this direction. Thus, Kawasaki (1982) found that sequences of *labial* + [w] are often absent in languages that otherwise allow [w] as the second member of consonant clusters. English, for example, allows [w] after apicals and velars, for example, *dwarf, twin, swine, quality, Gwendolyn*, but not after labials (except in words that are obvious loans, e.g., *bwana)* (see also Ohala 1978; Ohala & Lorentz 1977).[1] High tonality segments can also induce *flatness* (phonetically, at least) in contrasting segments. In Russian (and other Slavic languages), as is well known, there are extensive contrasts of *sharp* versus non*sharp* segments. Phonetically, the non*sharp* segments are often velarized, that is, phonetically, if not phonologically, *flat* (Fant 1960). (The reverse can also happen: Nonpharyngealized consonants in Arabic have been shown to have a somewhat fronted tongue configuration, Ghazeli 1977:75–76.) High tonality segments also reveal their opposition to *flat* segments by inhibiting their assimilatory effects. In Arabic, the palatal or *acute* segments like /i/ and /j/ block the assimilation of pharyngealization (Ghazeli 1977; Card 1983). In the development from Proto-Dravidian to (modern) Koḍagu, the front unrounded vowels /i/ and /e/ underwent a backward shift due to the influence of following retroflexes or labials but not if the palatal consonant /c/ preceded (Emeneau 1970).

In the preceding cases, the underlying unity of the diverse articulatory manifestations of *flat* segments are revealed first by the fact that they are all uniformly opposed to high tonality segments (those characterized as *sharp* or *acute)* and second by the observation that they often act in concert with or alternate with each other. These are, on the face of it, rather remarkable generalizations and should stimulate as much curiosity about their causes as would a generalization that high airflow segments such as voiceless fricatives, or affricates, or aspirated stops produce phonological effects similar to nasal consonants (which happens to be the case; see Ohala 1980, 1983b; Ohala & Amador 1981). How does the use of the feature *flat* and its defined physical correlates help to satisfy that curiosity? Before attempting to answer that question, let us take several paces back and examine in a more general way the explanatory task of phonology.

4. THE TASK OF PHONOLOGY

The discipline of phonology seeks to understand the behavior of speech sounds. The behavior of interest includes variation in pronunciation due to speaking style, phonetic context, sound change (and the results of sound change: regional and social dialectal variation, morphologically conditioned variation, the patterns of selection of sounds in segment inventories, etc.), the order of acquisition and patterns of mistakes in first and second language acquisition, and so on. As Hermann Paul (1880) pointed out, given the comlex nature of language, we must look in three domains for answers to our questions: the physical, psychological, and social. In each of these domains we identify entities or features of these domains that help us to explain the behavior of interest.[2] These features are the substantive primitives of the theoretical system or explanatory model—henceforth, simply "theory"—that we construct. For example, in the physical domain, oral-cavity volume and the compliance of the vocal tract walls are good candidates for such substantive primitives. From these (and further theoretical machinery) one can explain to a satisfying extent why the back-articulated stops, for example, [G, g], are the voiced stops most often missing even in languages that possess their voiceless cognates (Ohala & Riordan 1979; Ohala 1983a). Substantive primitives by themselves, however, accomplish nothing. Alone, they are like the ingredients of a cake, flour, sugar, eggs, and so on, which, even if lined up side-by-side on the kitchen counter, will never assemble themselves into a cake. One also needs an indication of how these things are to be mixed together. These are the formal primitives. In chess, they are the rules of the game; in economics, the laws of how prices are a function of supply and demand; in physics, the equations or nomograms that show how one variable can be predicted from others. Thus the familiar equation $p_2 = p_1 (v_2/v_1)$, a version of Boyle-Mariotte's Law, shows how, if one takes starting pressure and starting and ending volume as primitives, one can predict the resulting pressure of a gas in an enclosure. This equation is also part of the theory the phonologist needs to predict some of the aerodynamic conditions that affect voicing, frication, and derivative phenomena such as voice onset time and fundamendal frequency (F_0) microstructure, which phenomena play an important role in precipitating certain sound changes (Hombert, Ohala, & Ewan 1979; Ohala 1983a). Ultimately, all and every detail of the physical world that we find to be essential to an understanding of the behavior of speech sounds will have to be incorporated in the theory we construct. Naturally this will include neuromotor details, details on the structure and workings of the articulators, how the articulations get transformed into sound, how the sound is processed in the auditory system, and so on. Although it is necessary to limit the primitives in such a system to the absolute minimum, it does not seem to be a realistic expectation that the substantive primitives, that is, features, needed to explain the workings of speech can be limited to a mere 12 or 25 or even 50. Attempts to impose too small a limit on features could retard phonology in the same way Western chemistry was hampered by the premature theory that the primitive ingredients of all matter were earth, air, fire, and water.

The notion of primitive needs further qualification, however. In a

complete, strictly logical system, for example Euclid's geometry, there is only one level of primitives, and all theorems or predictions are derived directly or indirectly from it. One might imagine, then, that in phonology we could simply set up primitives at one level, say the level at which the physical properties of the speech mechanism were represented, and then all the rest of speech behavior, including aerodynamic and acoustic, could be derived from that. In such a system, tissue compliance might not be a primitive but rather something derived by reference to more basic physical properties of the tissues. In principle one could do things this way, but it is not really a very practical thing to try. First of all, our knowledge of how speech sounds are produced is now and always will be incomplete. So if we had to put off making predictions of how sounds will behave due to the characteristics of the auditory system until we had worked out all relevant details back to the level of the physical properties of tissues, we might never venture any explanations at all. Second, even though we know how to explain the resonances of the vocal tract by reference to more basic physical principles, it is hardly practical to have to cite them, when all that is to be explained is why listeners tend to confuse [k^w] and [p]—namely, because their acoustic patterns are similar. In this case it is sufficient to prove the point to refer to nomograms (e.g., those in Fant 1960) or acoustic data derived from spectrograms. The level at which primitives are stated is therefore a matter of choice. In this we do not differ from other disciplines: Molecular biologists seldom refer to the principles of atomic physics in their work.

Aside from coming up with theories in the first place, one of the principal activities of scientists is the evaluation of theories. Some forms of evaluation can be done by examining the linguistic and logical structure of the theories; others require comparison of the predictions made by theories with observations of the behavior of the object the theory purports to represent. In the former category there is the requirement that a theory do what it was constructed to do: make familiar what was mysterious. This seems such an obvious requirement, but it is so difficult to achieve that scientists in all fields have shown great ingenuity in masking their failure to satisfy it. They may do this by positing as a primitive or input to their theory a disguised form of the behavior or output they are supposed to account for. Positing a "vital principle" to account for differences between living and nonliving things or attributing the mutual attraction of physical bodies to gravity are examples of this. Theories also have to be self-consistent and nonredundant. This latter requirement—violations of which are said to be subject to correction by Occam's Razor—means that not only must primitives not duplicate each other but they must also be independent or orthogonal of each other, that is, the value or presence of one must not depend on any other primitive. It would violate orthogonality to state that susceptibility to heart disease is a function of salt intake and incidence of high blood pressure, since the latter is dependent, in part, on the former. (Of course, if salt consumption contributed to heart disease independently of its effects on blood pressure, then it, too, could be used as a primitive.)

The ultimate and often the most difficult form of evaluation of a theory requires comparing its predictions with what happens in nature, that is, to evaluate a theory by experiment. As Popper (1959) has convincingly argued, experimental tests do not reveal whether a theory is true but only that competing theories do less well at predicting what one observes. Thus the history of science is replete with examples of theories once thought to be true but which eventually had to be revised or replaced entirely by newer, more "fit," theories.

5. FEATURE SYSTEMS AS THE MODE FOR EXPLANATION

How do the traditional feature systems offered as the formal mechanism for representing and explaining aspects of speech-sound behavior stand up under these forms of evaluation? An assessment of this sort is attempted here, although, given the limitation of space, it is rather cursory. Nevertheless, except where noted, the criticisms made apply equally to all the better-known features systems and to the less well known ones, too (Brücke 1863; Bell 1867; Techmer 1880; Jespersen 1889; Pike 1943; Jakobson *et al.* 1963; Ladefoged 1971b, 1980; Chomsky & Halle 1968).

A feature system consists of the substantive primitives, the features, and the definitions of the features. In all the cases I know of, the definitions are given in physical terms, either articulatory (including aerodynamic) or acoustic–auditory. The definitions for *flat*, cited above, are typical. Inherent contradictions in the definitions of features can occur but are not a major problem.[3] Lack of orthogonality, however, is a serious problem for all, as has been recognized previously. Because the Jakobsonian and the Chomsky and Halle system use two binary features for vowel height, it turns out, for example, to use the former system, a value of '+' for [diffuse] dictates a '−' value for [compact] and vice versa. Ladefoged's multivalued feature system avoids this awkwardness and as a result is capable of giving a more insightful representation to vowel-shift rules (Ladefoged 1970). Lack of independence is also seen when a feature like [strident] is irrelevant for nasals and vowels. Any system that uses a feature of [voice] in addition to manner of articulation features such as

[sonorant] or [continuant] violates orthogonality since, as is well known, the value for voicing depends to some extent on whether there is a complete blocking of the airflow. If manner of articulation features are specified (they need not be; the primitives could have been chosen at a higher level), then voicing should be derived. In Ladefoged's (1980) system, the features [grave] and [sibilance] are dependent on the features for place of articulation. Such lack of orthogonality can be corrected by adopting primitives at a more basic level; on the other hand, the lack of orthogonality between *flat* and the features *grave* and *diffuse* (mentioned above) is apparently inherent in the makeup of speech and does not count as a defect of the system.

Inconsistency and redundancy are rather trivial imperfections, however. At most, they make the theories inelegant. A more important requirement is that these theories achieve their primary purpose of explaining speech-sound behavior. Most of the feature systems were designed to account for only a limited range of speech-sound behavior: the fact that speech sounds are different from each other and so can create semantically distinct utterances, and to give a common name to speech sounds that behave similarly. Although most feature systems need patching up as they are applied to a greater variety of languages, there does not seem to be any inherent flaw in them that would prevent them from achieving the first goal. The second goal, perhaps first articulated by Jespersen (1889:80–81), offers problems to virtually every feature system. By defining features in acoustic–auditory terms, the Jakobsonian system corrected an obvious deficiency of previous systems that were unable to deal with sounds that behaved similarly due to their shared acoustic–auditory properties. The feature *flat* that, as illustrated above, shows the similar phonological properties of labialization, retroflexion, and pharyngealization, is a prime example of this. But problems still remain. Some of these problems arise from failure to choose features at a sufficiently primitive level. For example, the glottal consonants [h ʔ] pattern like nonobstruents [j w l] in not blocking spreading nasalization (Schourup 1973), but they pattern like obstruents in creating F_0 perturbations (Hombert *et al.* 1979). Labial–velars such as [w k͡p g͡b] have a greater affinity to labials when they become or interact with fricatives or when they perturb the quality of adjacent vowels; however, they have greater affinity to velars when they become or interact with nasals (Ohala & Lorentz 1977; Ohala 1979a, b). Such problems will not be solved without an almost complete overhaul of most feature systems.

Feature systems used in generative phonology have adopted goals that go beyond the two just discussed; they are expected to reveal the naturalness of phonetically natural or expected sound patterns (Halle 1962;

Chomsky & Halle 1968:306–308, 400ff.). It is recognized that the feature systems in use fail to do this, that, as Chomsky and Halle note, their system fails to reflect the "intrinsic content" of the features. For example, there is nothing in the representations below to show that (1a) is more expected than (1b). Their proposed solution is a metarule, a marking convention (1c), which is used to interpret and evaluate the differences in naturalness of (1a) and (1b) (*u* is to be read as "the unmarked or expected value of").

(1) a. [−sonorant] → [−voice]
b. [−sonorant] → [+voice]
c. [*u* voice] → [−voice] / [______]
[−sonorant]

Unfortunately, the deficiency recognized is no more solved by the marking conventions (Chomsky & Halle 1968:chap. 9) than the introduction of further epicycles corrected the fundamental defect of the Ptolemaic model of the cosmos. The phonetic naturalness of sound patterns should be a product of the theory, that is, be derived from it; the marking conventions, on the contrary, decree it. This is an example of the fallacy whereby a restatement of a problem in a disguised way may be offered as a solution to the problem (Ohala 1971, 1972). None of the other feature systems would fare any better if adopted for the same task.

The problem is that there is more to creating a workable explanatory phonological theory than simply putting square brackets around adjectives. Stripped of the square brackets, arrows, and other typographical shorthand for ordinary verbal expressions such as "becomes," "in the environment of," these formal representations of phonological processes are virtually identical to the kinds of expressions used by Paṇini for the sandhi rules in Sanskrit, exemplified in (2).

(2) "The letters [s] and dentals in contact with [ṣ] and retroflexes are changed to [ṣ] and retroflexes [with the same manner of articulation], respectively" (Vasu 1891/1962:1671; "cerebrals" in original changed to "retroflexes").

There have been some advances made since Paṇini's time—one would hope so after 2300 years!—in the selection and definition of features, especially in picking ones that adequately reflect the natural classes of speech sounds and in defining features that are based on acoustic–auditory factors, but the theoretical matrix these terms are inserted into has not evolved accordingly.

It is the failure to develop an appropriate set of formal primitives for these feature systems, that is, relations that show how the features interact to produce observed sound patterns, that hinders progress. This is not because we are ignorant of the factors responsible. True, there are many important gaps in our knowledge about the workings of speech, for example, why assimilation occurs, but the more crucial hindrance is that the feature systems used are incapable of representing what we do know. The basic factors that affect the voicing of obstruents, for example, have been common knowledge for a long time (Passy 1890:161; Chao 1936; Chomsky & Halle 1968:325ff.) but the feature systems used simply are not adapted to incorporate this knowledge.

There are, however, several very promising formal theories of the speech process or parts of it that meet all the criteria discussed so far, including orthogonality of features (or what is equivalent, a strict differentiation of things that are primitive from those that are derived), full attention to substantive and formal primitives, and the ability (demonstrated in some cases, potential in others) to explain facts of speech of interest to phonologists; see, for example, Fant (1960); Liljencrants and Lindblom (1972); Stevens (1971, 1972); Mermelstein (1973); Flanagan, Ishizaka, and Shipley (1975); Catford (1977a); Ladefoged, Harshman, Goldstein, and Rice (1978); Wright (1980); Muller and Brown (1980); Kawasaki (1982); Ohala (1983); Keating (1984); Lindblom (1984). If we take seriously the goal adopted by generative phonologists of creating a formal system that will reflect the naturalness of sound patterns, then I suggest that we invest our time and energies in creating systems such as these that can help us to achieve that goal. The effort spent in creating and refining feature systems will never give us the same return.

6. WHAT ARE FEATURE SYSTEMS GOOD FOR?

At this point one might draw the conclusion that feature systems play no useful role at all in the task of phonology. This conclusion may be countered in the following ways.

First, it might be claimed that the task of representing sound patterns in a way that reveals their naturalness is not something that should be attempted using features grounded in phonetic substance. Such, at least, seems to be Ladefoged's view (Ladefoged 1980; Ladefoged & Traill 1980). He would limit the task of the phonetic features to their traditional task of accounting for the differences in sounds used to convey differences of meaning. A loose mapping would relate phonetic features to a separate set of phonological features, the task of the latter being to reveal naturalness of sound patterns. As should be clear from the preceding discussion, I agree with him that a simple list of phonetic features by themselves has very little explanatory value. However, I believe that a phonetically grounded theory should offer explanations for natural sound patterns. My solution to the dilemma is not to abandon the work but to use the right tools, to augment our list of features (substantive primitives) with a full range of formal primitives that show how the primitive elements interact. There is no need to develop logical arguments that this is the way to proceed; there are numerous existence proofs, those cited earlier, that demonstrate this. Ladefoged seems rather to emphasize nonexistence proofs for his position, that is, phonological data that cannot (yet) be explained by any known phonetic principles. Of course he is right. There will always be things in phonology that cannot be explained by reference to known phonetic principles—either because we have not discovered these phonetic principles yet or because they do not have a phonetic causation. As Paul (1880) indicated, some aspects of language are due to psychological and social factors. But Ladefoged has not identified (as Gödel and Heisenberg have in their disciplines) a logically or empirically necessary constraint on our phonological understanding. So, let us continue to try as best we can to explain phonological data in terms of phonetics. We will never know the limits of our understanding until we press those limits.

Second, it might be claimed that the set of features that reflects the naturalness of sound patterns has primarily (or even exclusively) a psychological reality. Thus, marking conventions may very well be represented in speakers' brains in a way not unlike a list so that metarules of the sort in (1c) are the appropriate way to indicate phonological naturalness. There is little evidence that native speakers are aware of all the sound patterns that linguists are able to discover and even less evidence that the native speakers have a special skill in differentiating natural and unnatural patterns. Until further evidence is produced, there is no reason to give serious attention to this position. I do agree, though, that it is important to investigate the psychological aspect of speech to discover the mental factors that help us to explain the behavior of speech sounds, but this is not to be done by the kind of facile psychologizing that has characterized the field of phonology for the last 30 years or so. Several techniques for exploring what native speakers know about the sound patterns in their language have been developed, and they give results of far more validity than those developed in the armchair (Ohala 1974, 1984; Ohala & Ohala 1978; Jaeger 1980).

Third, and this is my own position, it can be claimed that the features that have been offered over the past two millennia play a very useful role in phonology in that they are the terms we use in an informal, quite ad hoc, way to talk about speech sounds and classes of sounds. Every

scientific discipline makes use of such an informal, or less than rigorously defined, vocabulary. The terms "noble gas," "metal," and "semiconductor" have (or used to have) this status within chemistry. Terms such as "predator" and "altruism" serve a similar informal function in ethology. There is no need to think that every term finding its way into scientific descriptions has to be a primitive in a theoretical system or strictly orthogonal to every other term. An ad hoc label can even be applied to a class (defined by enumeration) whose underlying common composition has not yet been discovered. In this case it acts as a kind of flag alerting us to the need for further investigation; the label is not "carved in stone" and may be revised or dropped entirely as more information becomes available.[4] Needless to say, terms of this sort that can be casually invented should not be included in formal explanations unless they have been made to pass the kinds of tests discussed above. Feature systems, as they have traditionally been constructed and used, have this very useful but very limited function; the mistake made since the rise of generative phonology is to saddle them with a task that they are inherently incapable of accomplishing.

7. *FLAT* IS NOT AN ELEMENT IN A FORMAL SYSTEM

The feature *flat*, for example, is still primarily an informal term. It is defined by enumerating the class of segments that are included under it. Although some of the acoustic correlates of this class of sounds can be derived from the articulatory specification of members of the set, that is, labialization, retroflexion, and pharyngealization, not all of them can, and, more to the point, the common auditory effect of these sounds cannot yet be accounted for.

7.1. Where *Flat* Comes From

An appreciation of how vocal tract configurations give rise to some of the acoustic characteristics of *flat*, especially as this involves a lowering of F_3, may be obtained from the following oversimplified account (adapted from Chiba & Kajiyama 1958:149ff.; Fant 1960:86–87; Lieberman 1977).

Figure 15.1 shows, on the left, the standing wave patterns[5] of the volume–velocity waveforms of the lowest three resonances, R_1, R_2, and R_3, of a uniform tube closed at one end and open at the other. This is an analogue of the vocal tract, with the closed end representing the glottal region and the open end, the lips. The location of the minima and maxima in the waveform are marked by the letters "N" and "A,"

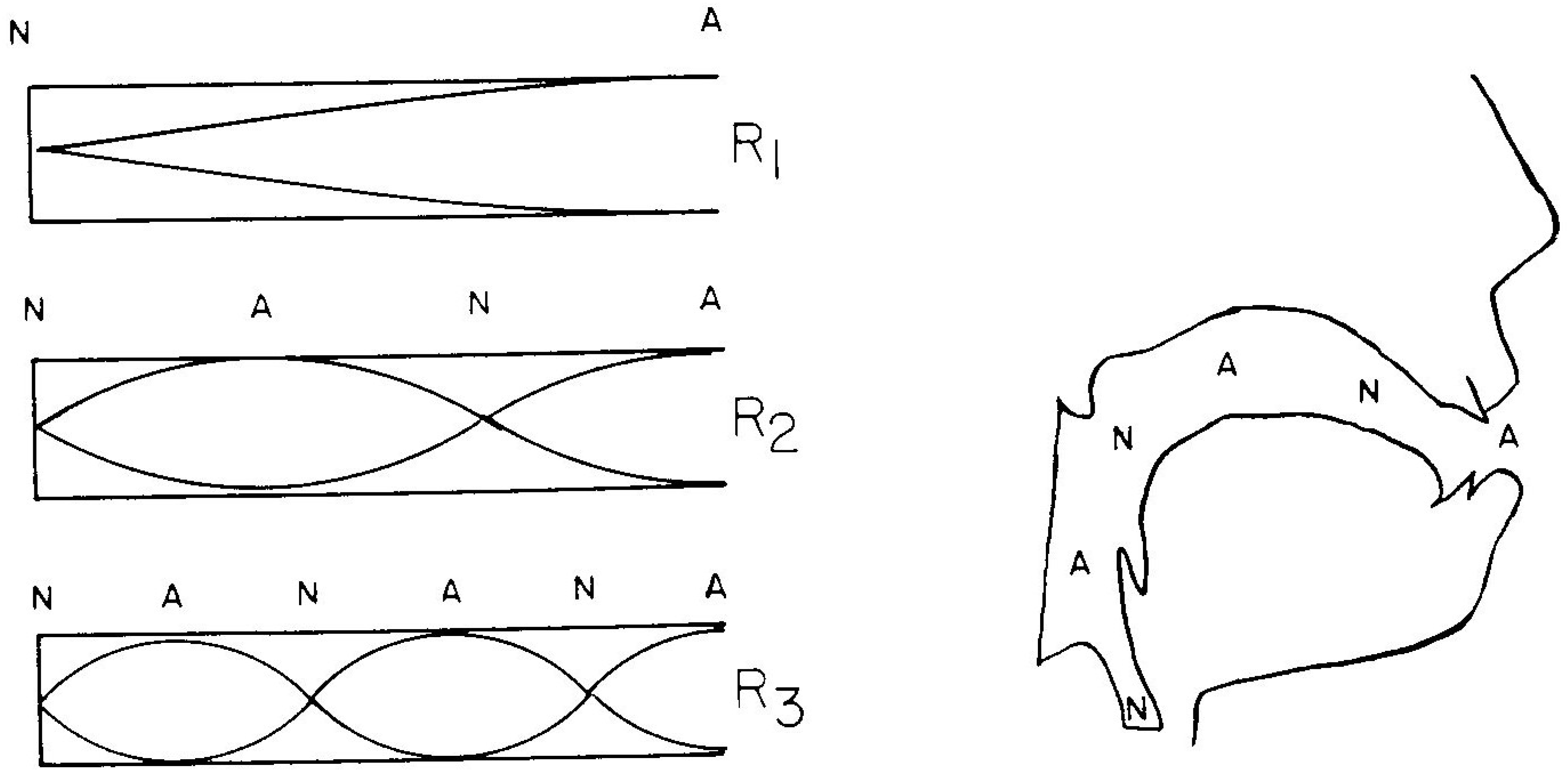

Figure 15.1 Left: schematic representation of the volume–velocity standing wave patterns of the lowest three resonant frequencies of a uniform tube closed at one end and open at the other. Right: the position of the nodes and antinodes of the standing wave pattern in the vocal tract of the third resonant frequency.

respectively (for "node" and "antinode"). The resonant frequencies of a vibrating body (in this case the air confined inside the tube) are those frequencies whose standing wave patterns optimally conform to the boundary conditions of the body. In this case the boundary conditions demand a velocity minimum at the closed end—where the movement of the air is most constrained—and a velocity maximum at the open end—where the movement of the air is least constrained. The standing waves shown are the lowest three that meet these conditions. The lowest resonant frequency has a wavelength equal to 4 times l (l = the length of the tube), the second to $\frac{4}{3}$ times l and the third to $\frac{4}{5}$ times l. Given that frequency = speed of sound/wavelength, it follows that for a 17.5 cm long tube (a typical value for an adult male speaker), and given the speed of sound as 35,000 cm/sec, the lowest three resonances will be 500, 1500, and 2500 Hz, which are roughly those of [ɜ], the vowel most closely approximating a uniform tube. Resonant frequencies higher or lower than these are obtained by introducing one or more constrictions. A constriction at a velocity maximum for a given resonant frequency's standing wave will lower the frequency, and a constriction at a velocity minimum will raise it. Conversely, a larger-than-normal dilation of the vocal tract at a velocity maximum will help to raise the frequency and at velocity minimum to lower it. In the right of Figure 15.1 is rough indication of where the velocity maxima and minima for the third resonance's standing wave are located in the vocal tract. Thus, a constriction

at the antinodes, that is, at the lips, in the midpalatal region, and/or in the pharyngeal region, will help to lower the third resonance. It should be emphasized that this is a considerably oversimplified account of the conversion from articulation to sound, but it does serve to indicate very roughly how the acoustic characteristics of *flat* might ultimately be related in a systematic way to its articulatory correlates.

This helps to make less mysterious the fact that articulations as distant as those at the lips and in the pharynx have similar effects. It also suggests why one variant of American English [ɝ], the vowel with the maximally low F_3, has three constrictions, one at the lips, one in the pharynx, and one in the midpalatal region. This still leaves unexplained, however, how the midpalatal constriction for this vowel can be substituted by a retroflex tongue shape, what is usually thought of as the canonical configuration for [ɝ]. Also unexplained is why a lowering or a weakening of the higher formants gives rise to the same auditory effect and why a lowering and/or weakening of either F_3 and/or F_2 or (in the case of labialization) all three of the lowest formants have the same auditory effect. Until we have answers to all these questions through a formal theory of the speech apparatus, the use of the feature *flat* is the only way—an informal way—we have to call attention to the natural class constituted by the segment types to which we attach that label. Nevertheless, it should not be assumed that the informal terminological system that includes *flat* is any substitute for the eventual explanatory formal theory of the speech process that is the ultimate goal of phonology.

NOTES

1. It might seem, at first glance, that the many instances of co-occurrence of labialization with velars or of sequences of velars plus [w], for example, [k^w] or [kw], are obvious counterexamples to the above prediction, since velars as well as labials are [+grave]. Velars, however, can be front or back; only back velars properly should be considered [+grave], that is, near the low end of the tonality scale. Front velars, that is, those at the beginning of English, *key*, *cape*, and *cash*, actually have quite high center frequencies in their noise burst and high F_2 (Lehiste & Peterson 1961; Liberman, Delattre, Cooper, & Gerstman 1954). In accord with the prediction given, back velars do not often exhibit a labialized–nonlabialized contrast. In fact, labialization is frequently found to accompany back velars, conceivably to enhance a contrast between them and front velars.

2. It should not be supposed—but unfortunately it often is—that these explanations can be completely lawlike, that is, nomological. They will instead be partial, probabilistic, explanations. So, in fact, are all explanations in every scientific discipline with the exception of mathematics, which has the advantage of dealing with manmade structures rather than the real world (Ohala forthcoming).

3. An inherent contradiction appeared, for example, in Halle's (1964) definitions of the Jakobson, Fant, and Halle featural characterizations of the major class segments (pointed out by Peter Ladefoged, personal communication). Halle recognized four degrees of narrowing in the vocal tract, from greater to lesser: contact, occlusion, obstruction, and constriction. Segments that are [+consonantal] have occlusion or contact; [+vocalic] segments have no narrowing closer than constriction. Given this, the characterization of liquids as [+consonantal] and [+vocalic] was inherently contradictory.

4. A good example of this is the replacement in many cases of the poorly defined terms "tense" versus "lax," "fortis" versus "lenis," and so on, by terms describing voice onset time (Lisker & Abramson 1964).

5. A standing wave is created when two waves having the same frequency and amplitude pass through each other going in opposite directions, as happens when a wave reflects off some acoustic interface, for example, the glottis end of the vocal tract. The summation of the two waves forms a wave that seems to oscillate in one spot in such a way that in certain regions (nodes) there is no net movement (these are found at intervals one-half wavelength apart) and at one-fourth wavelength away from the nodes there are areas where there is maximum movement of the wave (antinodes).

REFERENCES

Bell, A. M. (1867). *Visible speech: The science of universal alphabetics*. London: Simpkin, Marshall.

Brücke, E. W. von. (1863). Über eine neue Methode der phonetischen Transcription. *Sitzungsberichte der Philosophische-Historischen Classe der Kaiserlichen Akademie der Wissenschaften Wien, 41*, 223–285.

Card, E. A. (1983). *A phonetic and phonological study of Arabic emphasis*. Unpublished doctoral dissertation, Cornell University.

Catford, J. (1977a). *Fundamental problems in phonetics*. Bloomington: Indiana University Press.

Catford, J. (1977b). Mountain of tongues: The languages of the Caucasus. *Annual Review of Anthropology, 6*, 283–314.

Chao, Y. R. (1936). Types of plosives in Chinese. In D. Jones & D. B. Fry (Eds.), *Proceedings of the 2nd International Congress of Phonetic Sciences* (pp. 106–110). Cambridge: Cambridge University Press.

Chiba, T., & Kajiyama, M. (1958). *The vowel, its nature and structure*. Tokyo: The Phonetic Society of Japan.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Colarusso, J. (1974). Consonants with advanced tongue root in the Northwest Caucasian languages. *Papers from the Annual Meeting of the North Eastern Linguistic Society, 5*, 153–161.

Delattre, P. (1971). Pharyngeal features in the consonants of Arabic, German, Spanish, French, and American English. *Phonetica, 23*, 129–155.

Emeneau, M. B. (1970). Koḍagu vowels. *Journal of the American Oriental Society, 90*, 145–158.

Fant, G. (1960). *Acoustic theory of speech production*. 's-Gravenhage: Mouton.

Flanagan, J. L., Ishizaka, K., & 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.

Ghazeli, S. (1977). Back consonants and backing coarticulation in Arabic. Unpublished doctoral dissertation, University of Texas at Austin.

Halle, M. (1962). Phonology in generative grammar. *Word, 18*, 54–72.

Halle, M. (1964). On the bases of phonology. In J. Fodor & J. Katz (Eds.), *The structure of language* (pp. 324–333). Englewood Cliffs, NJ: Prentice-Hall.

Hashimoto, M. J. (1973). Retroflex endings in Ancient Chinese. *Journal of Chinese Linguistics, 1*, 183–207.

Hercus, L. A. (1969). *The languages of Victoria: A late survey* (Australian Aboriginal Studies No. 17; Linguistic Series No. 6). Canberra: Australian Institute of Aboriginal Studies.

Hombert, J.-M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language, 55*, 37–58.

Jaeger, J. J. (1980). Testing the psychological reality of phonemes. *Language & Speech, 23*, 233–253.

Jakobson, R. (1962). Mufaxxama, the 'emphatic' phonemes in Arabic. In *R. Jakobson, Selected Writings, Vol. 1: Phonological studies* (pp. 510–522). 's-Gravenhage: Mouton.

Jakobson, R., Fant, C. G. M., & Halle, M. (1963). *Preliminaries to speech analysis* (sixth ed.). Cambridge: MIT Press.

Jakobson, R., & Halle, M. (1956). *Fundamentals of language*. 's-Gravenhage: Mouton.

Jespersen, O. (1889). *Articulation of speech sounds, represented by means of analphabetic symbols*. Marburg in Hesse: N. G. Elwert.

Kawasaki, H. (1982). An acoustical basis for universal constraints on sound sequences. Unpublished doctoral dissertation, University of California, Berkeley.

Keating, P. (1984). Aerodynamic modeling at UCLA. *UCLA Working Papers in Phonetics, 59*, 18–28.

Kirton, J. F. (1967). Anyula phonology. *Pacific Linguistics, Series A, 10*, 15–28.

Ladefoged, P. (1970). An alternative set of vowel shift rules. *UCLA Working Papers in Phonetics, 17*, 25–28.

Ladefoged, P. (1971a). The limits of phonology. In L. L. Hammerich, R. Jakobson, & E. Zwirner (Eds.), *Form and substance. Phonetic and linguistic papers presented to Eli Fischer-Jørgensen* (pp. 47–56). Copenhagen: Akademisk Forlag.

Ladefoged, P. (1971b). *Preliminaries to linguistic phonetics*. Chicago: University of Chicago Press.

Ladefoged, P. (1980). What are linguistic sounds made of? *Language, 55*, 485–502.

Ladefoged, P., Harshman, R., Goldstein, L., & Rice, L. (1978). Generating vocal tract shapes from formant frequencies. *Journal of the Acoustical Society of America, 64*, 1027–1035.

Ladefoged, P., & Traill, T. (1980). The phonetic inadequacy of phonological specifications of clicks. *UCLA Working Papers in Phonetics, 49*, 1–27.

Lehiste, I. (1964). Acoustical characteristics of selected English consonants. *International Journal of American Linguistics*, Publication 34.

Lehiste, I., & Peterson, G. E. (1961). Transitions, glides, and diphthongs. *Journal of the Acoustical Society of America, 33*, 268–277.

Lehn, W. (1963). *Emphasis in Cairo Arabic*. *Language, 39*, 29–39.

Leslau, W. (1957). The phonetic treatment of the Arabic loanwords in Ethiopic. *Word, 13*, 100–123.

Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. (1954). The role of consonant–vowel transitions in the perception of stop and nasal consonants. *Psychological Monographs, 68*, no. 8 (whole no. 379), pp. 1–13.

Lieberman, P. (1977). Speech physiology and acoustic phonetics: An introduction. New York: Macmillan.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language, 48*, 839–862.

Lindblom, B. (1984). Can the models of evolutionary biology be applied to phonetic problems? In M. P. R. v. d. Broecke & A. Cohen (Eds.), *Proceedings, 10th International Congress of Phonetic Sciences* (pp. 67–81). Dordrecht: Foris.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Mermelstein, P. (1973). Articulatory model for the study of speech production. *Journal of the Acoustical Society of America, 53*, 1070–1082.

Muller, E. M., & Brown, W. S., Jr. (1980). Variations in the supraglottal air pressure waveform and their articulatory interpretation. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 4) (pp. 317–389). New York: Academic Press.

Müller, M. (1855). *The languages of the seat of war in the East* (2nd ed.). London: Williams & Norgate.

Ohala, J. J. (1971). The role of physiological and acoustic models in explaining the direction of sound change. In *Project on Linguistic Analysis Reports* (No. 15, pp. 25–40). Berkeley: University of California.

Ohala, J. J. (1972). Physical models in phonology. In A. Rigault & R. Charbonneau (Eds.), *Proceedings, 7th International Congress of Phonetic Sciences, Montreal, 1971* (pp. 1166–1171). The Hague: Mouton.

Ohala, J. J. (1974). Experimental historical linguistics. In J. M. Anderson & C. Jones (Eds.), *Historical linguistics II: Theory and description in phonology* (pp. 353–389). Amsterdam: North Holland.

Ohala, J. J. (1978). Southern Bantu vs. the world: The case of palatalization of labials. *Proceedings, Annual Meeting of the Berkeley Linguistics Society, 4*, 370–386.

Ohala, J. J. (1979a). The contribution of acoustic phonetics to phonology. In B. Lindblom & S. Ohman (Eds.), *Frontiers of speech communication research* (pp. 355–363). London: Academic Press.

Ohala, J. J. (1979b). Universals of labial velars and de Saussure's chess analogy. *Proceedings of the Ninth International Congress of Phonetic Sciences* (Vol. 2) (pp. 41–47). Copenhagen: Institute of Phonetics.

Ohala, J. J. (1980). The application of phonological universals in speech pathology. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 3) (pp. 75–97). New York: Academic Press.

Ohala, J. J. (1983a). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 189–216). New York: Springer Verlag.

Ohala, J. J. (1983b). The phonological end justifies any means. *Proceedings of the XIIIth International Congress of Linguists, Tokyo* (pp. 232–243). Tokyo: ICL Editorial Committee.

Ohala, J. J. (1984). *Consumer's guide to evidence in phonology*. Berkeley (Cognitive Science Report No. 17). Berkeley, University of California. Forthcoming in R. Rhodes (Ed.), *Evidence in phonology*. Ann Arbor: Karoma Press.

Ohala, J. J. (Forthcoming). Explanation, evidence, and experiment in phonology. *Proceedings, Fifth International Phonology Meeting, Eisenstadt*.

Ohala, J. J., & Amador, M. (1981). Spontaneous nasalization. *Journal of the Acoustical Society of America, 69*, 954–955.

Ohala, J. J., & Lorentz, J. (1977). The story of [w]: An exercise in the phonetic explanation for sound patterns. *Proceedings, Annual Meeting of the Berkeley Linguistics Society, 3*, 577–599.

Ohala, J. J., & Ohala, M. (1978). Experimental methods in phonology. *Acta Linguistica Hafniensia, 16*, 229–230.

Ohala, J. J., & Riordan, C. J. (1979). Passive vocal tract enlargement during voiced stops.

In J. J. Wolf & D. H. Klatt (Eds.), *Speech communication papers* (pp. 89–92). New York: Acoustical Society of America.

Paul, H. (1880). *Prinzipien der Sprachgeschichte*. Halle: Max Niemeyer.

Passy, P. (1890). *Étude sur les changements phonétiques*. Paris: Librairie Firmin-Didot.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175–184.

Pike, K. L. (1943). *Phonetics*. Ann Arbor: University of Michigan Press.

Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.

Schourup, L. C. (1973). A cross-language study of vowel nasalization. *Working Papers in Linguistics* [Ohio State University], *15*, 190–221.

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 & P. B. Denes (Eds.), *Human communication: A unified view* (pp. 51–66). New York: McGraw Hill.

Stevens, K. N. (1980). Discussion during symposium on phonetic universals in phonological systems and their explanation. *Proceedings, 9th International Congress of Phonetic Sciences* (Vol. 3) (pp. 185–186). Copenhagen: Institute of Phonetics.

Techmer, F. H. H. (1880). *Phonetik; zur vergleichenden Physiologie der Stimme und Sprache*. Leipzig.

Trubetzkoy, N. S. (1969). *Principles of phonology* (C. A. M. Baltaxe, Trans.). Berkeley: University of California Press. (Original work published 1939.)

Uldall, E. T. (1958). American 'molar' *r* and 'flapped' *r*. *Revista do Laboratorio de Fonetica Experimental* (Coimbra), *4*, 103–106.

Vasu, S. C. (1962). *The Ashṭādhyāyī of Pāṇini* (Vol. 2). Delhi: Motilal Banarsidass. (Original work published in 1891.)

Wright, J. T. (1980). The behavior of nasalized vowels in the perceptual vowel space. *Report of the Phonology Laboratory* (Berkeley), *5*, 127–163.