

## The Phonological End Justifies Any Means

John J. Ohala  
University of California, Berkeley

For well over a century linguists have declared that theirs is a largely autonomous discipline which, with the exceptions of a few excursions into such areas as history, physiology, or acoustics, owes little to other scholarly fields (cf. Whitney 1867:53). The subject matter of autonomous linguistics was said to be the patterns and the structure that exists in language and the method —there really has been only *one*— used to find this structure was the *comparative method* which is basically just a rigorous way of demonstrating relations between units of language. So defined, linguistics probably *was* an autonomous discipline and it got a lot of mileage out of its single method. Bloomfield's remark in 1922 still holds:

It was Grimm's merit... that by the strength of a method [the comparative method], he conquered for science a body of facts so vast that the generations since have worked well within the bounds he reached and scarcely ever gone beyond.

The comparative method when applied to a single language yielded the method of *internal reconstruction* and this technique is the basis for the structuralist phonologists' criterion of complementary distribution used in finding phonemes and the generative phonologists' positing of common "underlying" phonological forms for related morphemes.

The autonomy of linguistics and the use of the comparative method were suited to the task of describing and cataloguing the units and patterns in language. But in the 20th century linguists have started to ask *why* these units and patterns exist, i.e., they seek explanations for, not just descriptions of linguistic structure. We have reached the point where the continued re-working of the patterns in language no longer gives us answers to our questions.

Accordingly, one of the most promising trends in modern linguistics has been the bridges built between linguistics and other disciplines: sociolinguistics, psycholinguistics, neurolinguistics, etc. The data and methods of these other disciplines have already greatly enriched our field and promise to continue to do so. In this paper I will discuss four problems in phonology and attempt to show, briefly, how their solution requires us to throw open the gates of our science to the data and methods from any source that proves useful.

### Spontaneous Nasalization

In most cases, distinctively nasal vowels derive from sequences of *vowel* +

*nasal consonant* (or sometimes NC+V), e.g., French [vã] "wind" < Latin *ventus*, Hindi [dāt] "tooth" < Sanskrit *dant-*. On occasion, however, nasal vowels appear in words which never had a nasal consonant at any point in history, e.g., Hindi [sāp] "snake" < Sanskrit *sarpa*, [pəhuč] "attain" < Prakrit *pahuccai*. These are cases of so-called "spontaneous nasalization" which seems to have been systematically investigated first by Grierson (1922).

As it happens, in the majority of cases the spontaneously nasalized vowel appears adjacent to consonants characterized by heavy airflow: the glottal fricative [h], voiceless fricatives and affricates, and aspirated stops. In the case of [h] it is reasonable to suppose that since there is no aerodynamic requirement that the velum be raised during its production, it could be produced with a lowered velum and this state could be assimilated by adjacent vowels. (See Ohala 1975 for additional speculations.) This would not account for the involvement of oral obstruents, however, since they would definitely require an elevated velum.

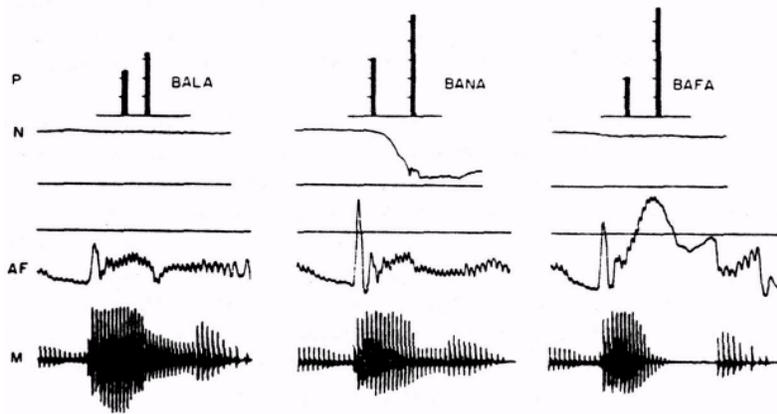
Ohala and Amador (1981) (henceforth O & A) attempted to test a hypothesis (Ohala 1975, 1980) that vowels produced with a slightly *open glottis* might have acoustic characteristics which would mimic the effects of nasalization. A slightly open glottis allows some coupling of the subglottal cavity to the oral cavity (comparable to the coupling of the nasal cavity to the oral cavity during nasal sounds) and results in anti-resonances which, when they interact with the resonances of the oral cavity, increase the bandwidth and lower the amplitude of the first resonance (Fant 1973:8, Fujimura & Lindqvist 1971). Such effects coincide with some of the acoustic cues for nasalization on vowels. That vowel margins immediately adjacent to high airflow consonants would have a slightly open glottis has been shown by several glottographic, fibrescopic, and airflow studies (e.g., Sawashima 1969).

To see whether physiologically oral vowels might sound more nasalized on those parts abutting voiceless fricatives, vis-a-vis other oral environments, O & A used digital methods to create a series of steady-state vowels each 500 msec long by iterating single periods from the relevant parts of CVC(V) speech samples spoken by 4 adult males (2 American English and 2 Mexican Spanish speakers). Test vowels were made from the last or second-to-last period of vowels before the voiceless fricatives [s], [f], and, for Spanish, [x]; control vowels came from the last period before [n] (to make sure that the cues for true nasalization would survive the iteration process) and from periods before the oral consonants [d], [l], and, for Spanish, the trill [r], as well as from periods equidistant between the 2 C's in the CVC(V) utterances, that is, where the effect of the consonants was expected to be minimal (to make sure that the iteration process itself didn't introduce distortions that would mimic nasalization).

These vowels, with normalized amplitude (but not normalized pitch), were randomized and presented to 14 phonetically trained American English listeners who judged the degree of nasalization of each vowel on a 7-point scale, where "1" meant "completely oral" and "7" meant "heavily nasalized." In separate

recording sessions, velar elevation and oral airflow were sampled (using the nasograph and a pneumotachograph, respectively) as the same 4 speakers spoke the same words from which the iterated vowels were made.

As shown in Fig. 1, which presents representative data from one of the Spanish speakers, the physiological recordings revealed: (1) the expected lowering of the velum on vowels near nasals (see [bana]) but no detectable lowering during vowels next to oral consonants (see [bala] and [bafa]), (2) greater airflow—and by implication, greater glottal opening—during the latter part of vowels next to voiceless fricatives (see [bafa]). Some results of the perceptual study are represented at the top of the figure.



**Fig. 1. P: Listeners' judgements of degree of nasality, on 7-pt. scale, of iterated vowel (left bar of each pair taken from period in middle of vowel; right bar from period at the end); N: Nasograph signal, where elevation of line is correlated with elevation of velum; AF: Oral airflow measured by pneumotachograph; M: Microphone signal. Utterances from a speaker of Mexican Spanish. Temporal synchronization of parameters is approximate.**

The height of the vertical bars indicates the degree of perceived nasalization (tick marks at 1 unit intervals). The left bar of each pair corresponds to the iterated vowel made from the period excised from the middle of the uttered vowel; the right bar, that from the period just before the onset of  $C_2$ . As expected, stimulus vowels made from the period in the middle and those taken from periods immediately before control oral consonants, e.g., [l], were judged to be relatively non-nasal. However, stimulus vowels from periods before nasals and those before voiceless fricatives were heard as being heavily nasalized even though, in the latter case, there was demonstrably no physiological nasalization. Such "spurious" nasalization was strongest on the vowel [a] but very weak on higher vowels. The vowel [a] may have enhanced the spurious nasalization since its pharyngeal constriction leads to acoustic coupling between the vocal tract and the glottal volume velocity waveform in the region of F1 and thus increased

bandwidth of F1 (K. N. Stevens, personal communication, Fant 1980a, b).

O & A concluded that the sound changes manifesting spontaneous nasalization came about when vowels that "sounded" nasalized, even though they weren't, were reinterpreted by listeners as having actual nasalization and were thereafter pronounced with nasalization.

### Asymmetries in the Direction of Sound Change

It is a very old and, I think, quite correct notion that certain sound changes occur due to the sounds involved being acoustically and perceptually similar and thus confusable (Sweet 1874:15). Consider, for example, the following very common sound changes: kw > p (e.g., Proto-Indo-European *\*ekwos* > Greek *hippos* "horse"), pj > tj (e.g., Roman Italian [pjeno] > Genoese Italian [tʃena] "full", and ki > tʃi (e.g., *chicken* [tʃɪkən] < Anglo-Saxon *cocc* "cock, rooster" + diminutive ending) (Ohala 1979). Various speech perception studies have found parallel confusions (e.g., Winitz, Scheib, & Reeds 1972). However, to say simply that two sounds, A and B, are confusable would imply that A should change into B as often as B changes into A. This is generally not the case, though. In the examples given above, the change is almost invariably in the direction presented, rarely in the reverse direction.

We might be tempted to offer an articulatory explanation for these asymmetries (e.g., "the preferred direction of change results in the physiologically simpler sound") except for the fact that these asymmetries also show up in the laboratory-derived confusion matrices from tasks where listeners just had to identify, not articulate, the sounds they heard. For example, Winitz *et al.*, in one of the conditions of their study, obtained the following confusions (where ">" means "reported as" and the percentage given indicates the percentage of time the sound was confused in the way indicated):

[k] > [t] / _ [i] (32%) but	[t] > [k] / _ [i] (6%)
[p] > [t] / _ [i] (34%) but	[t] > [p] / _ [i] (6%)
[k] > [p] / _ [u] (27%) but	[p] > [k] / _ [u] (16%)

One might explain these asymmetries by "response bias" in the cases where the more frequent confusion yields the more frequently occurring sound [t], but this would not account for the confusion [k] > [p] since [k] is more frequent in English than [p] (Wang & Crawford 1960). Asymmetries in confusion must have something to do with the physical structure of the sounds themselves and how the human perceptual system processes them.

A clue to this problem may be found by examining the kind of confusions that occur in identification tasks involving stimuli presented to other sensory channels. Confusion matrices derived from identification tasks where the stimuli are the 26 capital letters of the Roman alphabet have been reported by Gilmore, Hersh, Caramazza, & Griffin (1979) for a visual presentation and by Craig (1979) for a vibrotactile (touch) display. Both show clear and similar

asymmetries. For example the following confusions were found more frequently than their reverse: Q>O, E>F, R>P, B>P, P>F, J>I, W>V. In all these pairs the first letter is structurally identical to the second *plus an extra feature*. If we assume that the major cause of these confusions is the incomplete perception of the ensemble of attributes that make up the letter, it follows that the failure to detect this extra differentiating feature will lead to the misidentification of the target letter as that letter which equals the target letter minus this feature. Adding (or hallucinating) an absent feature is less likely so the reverse confusion should have a lower probability of occurrence. As Garner (1978) has convincingly argued, such asymmetrical confusions should happen only between stimuli that differ by attributes which vary in an all-or-none way ("features" in his terminology), not by attributes which vary in a continuous way ("dimensions"). This principle should hold no matter what sensory channel is involved. Thus we should have a good chance to understand the asymmetry of the above-mentioned sound changes if we look for the "extra" feature which differentiates, say, [kw] from [p], etc.

Careful research is needed to identify these features but some preliminary speculations about them can probably be made. The sequence [kw] differs from [p] by the presence of a moderately sharp spectral peak in the low frequency region of the noise burst; otherwise they are largely identical, e.g., in F2 transition. The palatalized labials [bj, pj] have a brief rise in F2 following release which is lacking in plain apicals (Ohala 1979). Systematic manipulation of the acoustic waveforms of these sounds could confirm whether removal of these features leads to the confusions predicted.

Of course, asymmetries in the direction of common sound changes may have other causes, too, some of them non-perceptual. Elsewhere, I have discussed aerodynamic reasons why long voiced stops should become voiceless but not vice-versa (Ohala, 1982).

### Phonemes as Categories

The "phoneme theory" which was elaborated around the turn of the century consists of many semi-independent claims, e.g., that speech consists of a string of concatenated units (of "phoneme" size) and that physically distinct sounds such as the aspirated and unaspirated stops in the words *cool* [k<sup>h</sup>u:l] and *school* [sku:l] are the "same" sound. With Sapir and a few other adventurous phonologists, these became psychological claims. But except for anecdotal evidence such as that reported by Sapir (1933) from his attempts to teach his informants to write their language, or Chao (1934) from word games, these claims have not been experimentally verified, i.e., where some pains are taken to control factors which might distort the results and therefore render their interpretation ambiguous (Twaddell 1935).

In the Phonology Laboratory at Berkeley we have tried to use some standard techniques from experimental psychology to confirm (or disconfirm) claims such as those made about the grouping of allophones into phonemes. One of these

techniques, so-called "concept formation" (CF) (Deese & Hulse 1967, chap. 12) had been previously used in a linguistics experiment addressing a syntactic issue (Baker, Prideaux, & Derwing 1973). Jaeger (1980a) reported our first experiences with this technique where the question addressed was "do native speakers of English regard the unaspirated stop in words such as *school* to be in the same category as the aspirated stop in *cool*?"

Briefly, a CF experiment would proceed as follows. The subject (S), seated in a quiet room, dons earphones and hears instructions of the sort: "You'll hear a series of words over the earphones, some of which belong to a certain category due to the way they sound; the rest do not belong to the category. If the word is in the category, respond 'yes', if not, respond 'no'. We'll let you know after you respond what the right answer was. You'll have to guess on the first few words but eventually you should figure how to anticipate the right answer. When this happens we'll give you a test. In this part we won't tell you what the right answer was after you respond."

Table 1. Sample CF session; target category: words with [k<sup>h</sup>].

	Trial No.	Stimulus (over earphones)	Sample response from S	Correct answer (over earphones)
	1	keep	no	yes
	2	fan	yes	no
	3	cash	yes	yes
	4	ghost	yes	no
	5	lip	no	no
	6	knife	no	no
TRAINING SESSION	7	choir	yes	yes
	8	chip	no	no
	9	ceiling	no	no
	10	kerosene	yes	yes
	11	occur	no	yes
	12	cool	yes	yes
	⋮			
	75	gnat	no	no
	76	science	no	—
	77	kitchen	yes	—
TEST	78	square	yes	—
	79	step	no	—
	80	school	yes	—
	⋮			

Then the experimental session might proceed as in Table 1, which should be read from left to right, top to bottom.

As can be seen, except for focussing Ss' attention on the pronunciation of the stimulus words, no other hints are given about the defining attributes of the target category. Ss must figure this out on their own by pure induction.

It is possible in this way to teach linguistic concepts or categories to linguistically naive Ss without using any verbal mediation.

The first part of the experiment, where only clear, uncontroversial exemplars of category and non-category items are presented, with feedback, is the *training session*. As shown in Table 1, in order to make sure Ss don't inadvertently form some unwanted category using orthographic criteria (based on their own mental image of the spelled word), the category words have spellings which represent [k<sup>h</sup>] in diverse ways: *ch*, *c*, *k*, *qu*. Moreover, some of the non-category words may be spelled with those same letters representing different sounds as in, e.g., *knife*, *chip*, *ceiling*. The criterion of having learned the category was set at 15 correct trials with 2 or fewer errors. When this was achieved the S began the *test session* where, without warning, the items whose category membership is at question are introduced (along with items like those in the training session) and where feedback is withheld. Of interest is how Ss categorize these new items in comparison to the clear cases. For example, in Table 1, if Ss respond 'yes' to the words *square* and *school* as often as they do to *kitchen* and the other words with [k<sup>h</sup>] (corrected for the number of times they respond 'yes' inappropriately to non-category items), then we may conclude that they regard [k] and [k<sup>h</sup>] as belonging to the same category.

In fact, as Jaeger reported, this is exactly what American English Ss (all linguistically naive) do, in conformity with the traditional phonemic analyses of English. This result is of some interest because there is phonetic evidence that the [k] in *school* is perceptually indistinguishable—at least to American English listeners—from the [g̊] in *ghoul* (Lotz, Abramson, Gerstman, Ingemann, & Nemser 1960), which is in a different phoneme.

One might still ask whether there could be some small phonetic difference between [g̊] and [k] which influenced Ss to put them in different categories. Also, is it possible that Ss had some sort of response bias? That is, that they automatically put the [k] or any new sound remotely like the [k<sup>h</sup>] into the target category and that they would do the same if the target category was /g/ (i.e., included [g̊] and [g]).

These questions were addressed using the CF technique (Ohala, forthcoming). Twenty linguistically naive American English speaking Ss were assigned randomly to one of two groups. The overall format of the experiment was the same as that described by Jaeger with the following exceptions: although the target category for Group I was 'words containing [k<sup>h</sup>]', included in the non-category exemplars during the training session were the words *ghoul*, *gate*, *gold*, and *grape*. These words were created by splicing the [s] from the beginning of the words *school*, *skate*, *scold*, and *scrape*, i.e., they contained [k], not [g̊] (if one believes there is a difference). These latter four words appeared intact (i.e., with the [s]) in the test session presented to Group I. Would Ss assign these words to the target category even though they—or the crucial part of them—had been presented as non-category items in the training session? The target category for Group II was words containing [g] or [g̊] which was

exemplified not only by "genuine" /g/'s such as those in *glitter* and *together* but also the four words given above, *ghoul*, etc., which had been formed by splicing the [s] from *school* and so on. Would Ss reject these words from the category even though a fragment of them had been given in the training session as category items?

As it turned out, Group I included the test words *school*, etc. in the category with [k<sup>h</sup>] just as decisively as Group II excluded them from the category of /g/, in spite of what might be considered conflicting evidence during the training session. More likely, the evidence in the training session was *not* conflicting because [g̊] and [k] really are identical but the criterion for assigning allophones to phonemes is distributional, not purely phonetic. These results also show that the categorization of new items in the test session is not influenced by any obvious response bias.

To my knowledge, these results and those reported by Jaeger represent the first experimental verification of the traditional phonemic claims. In other studies using the CF technique English Ss demonstrated that they regard the affricates [tʃ] and [dʒ] in words such as *chip* and *gyp* to be single sounds not clusters, again, as has traditionally been claimed. Jaeger (1980b) also used this technique to show that English speakers' "knowledge" of the sound patterns subsumed under the labels "vowel shift" and "vowel laxing" (e.g., the pattern that relates *insane* and *insanity*) is mediated to a significant extent by their knowledge of English orthography.

### Size-sound Symbolism

From Jespersen (1922) to Jakobson and Waugh (1979) there has been extensive documentation and experimental verification (Sapir 1929) of a widespread, cross-language tendency to use certain specific speech sounds in words related to the semantic dimension of size and correlated notions, e.g., distance, age, etc. Specifically, high front vowels, such as [i ɪ y e ø], are used in words associated with SMALL and lower, backer vowels, especially [ɑ ɔ o] in words associated with LARGE. Examples from a variety of languages of SMALL words are: English, *teeny*, *wee*, *little*, *doggie*; Spanish, *chico*; French *petit*; Japanese [tʃi:sai]; Greek *mikros*. Examples of LARGE vocabulary from the same languages: English, *large*, *huge*; Spanish, *gordo*; French, *gros*, *grand*; Japanese [o:ki:]; Greek, *makros*. Westermann (1927) showed that in some African languages tone was also used systematically to convey size, high tone for SMALL and low tone for LARGE, e.g., Twi [kàkrà] "small", [kàkrà] "large". Nichols (1971) documented cases in North American languages of systematic association of certain consonant types with opposite ends of the size continuum, e.g., Tillamook [waqɑq] "frog" but [wu-wekek] "(small) frog"; Wiyot, [ditatk] "two round things", [ditsatsk] (ditto, diminutive), [ditʃatʃk] (ditto, augmentative).

There are, to be sure, exceptions to these tendencies. The English words *big* and *small* are prime examples of this. Nevertheless, Sapir (1929) and many

others after him demonstrated in psychological experiments that although native speakers may have these conflicting patterns in the existing vocabulary of size, when asked to assign nonsense names such as [gil] and [gɔl] to large and small objects, they almost invariably pick the word with the [i] for the small object and [ɔ] for the large object. This result has cross-language validity (Chastaing 1965). Also, a few quantitative studies of the relevant vocabulary have been conducted (Thorndike 1945, Chastaing 1965, Ultan 1978) and they demonstrate that the tendency noted is not significantly weakened by the exceptions.

There have been many attempts to find some articulatory dimension characteristic of these speech sounds which is iconic with the dimension of size, but none of them can account for the full range of the size-sound symbolism data, including those of tone and consonants. There is, however, one physical characteristic of speech sounds, whether vowel, consonant, or tone, which predicts fairly successfully how they will be used in size-sound symbolism, viz., their *acoustic frequency*. The vowels characterizing SMALL have high F2, those characterizing LARGE, a low F2 (or, more precisely, the difference between F2 and F1). The consonants used with SMALL have, in general, predominantly higher frequencies (either in F2 transition or in frication or noise burst) than those used with LARGE. With tone, it is quite simply the higher FO which is used with SMALL and low FO with LARGE.

But why should a correlation exist between frequency and size? It has been suggested that speakers would naturally associate high frequency sounds with small objects and low frequency with large ones because for physical reasons the natural frequency of the sound emitted by an animal or an object (e.g., bells, hollow logs) is inversely related to their physical dimensions (Jespersen, Chastaing). I believe this is essentially correct except that the association is much older than the individual speaker who recognizes it and even much older than human language or the human species. This, at least, is the lesson I derive from reading the ethological literature. For example, Morton (1977) documented the existence of an amazing cross-species (birds and mammals) similarity in the acoustic characteristics of those vocalizations which animals use in face-to-face competitive encounters. Confident aggressors emit harsh or staccato cries with a low FO; submissive or non-threatening individuals, tonelike cries with a high FO. The dog's aggressive growl and submissive whine or yelp are familiar examples. Morton suggested that these vocalizations, like many visual displays given in competitive encounters (e.g., erection of the hair), serve indirectly to convey an impression of the apparent size of the animal. As mentioned above, a large individual would naturally have large and more massive vocal cords (or, in birds, syringeal membranes) and these would, for physical reasons, tend to vibrate irregularly and at a low frequency; the smaller vocal cords of small individuals would tend to vibrate in a more regular way at a high frequency. The aggressor could exploit this and enhance its fearsomeness by emitting a cry with acoustic characteristics of a larger in-

dividual; a submissive individual benefits by giving the impression of being small, and therefore non-threatening, and so would produce a "small" cry. So consistent across species, so stereotyped in use and apparently unlearned is the communication of size by frequency, that this code, call it the *frequency code*, must be genetically specified—that is, it is maintained by a genetic not a cultural template.

I propose that the frequency code, which must be innate in humans as well as non-humans, is the basis for the phonetic patterns observed in size-sound symbolism. (I also believe the frequency code underlies universals of intonation — both linguistic and paralinguistic—and of certain facial expressions such as the smile which, when produced with a vocalization, could systematically alter the dominant frequencies of the resulting sound. Limitations of space prevent me from providing details on these points.) To be sure, the use of the frequency code in sound symbolism by humans differs in important ways from its use by animals. The speaker uttering the word *teeny* is not attempting to appear small or non-threatening. Rather, the intention is to refer to or denote something small. Nevertheless, in other respects the parallels are considerable: the selection of the frequency parameter to convey size (why not duration, bandwidth, spectral tilt?), the assignment of SMALL to high frequency and LARGE to low frequency (why not the reverse?). These parallels argue for a common origin of the frequency code.

It should not surprise us that the shape of speech, which is influenced by the physical environment, is also influenced by the ethological environment.

### Conclusion

In this paper I have tried to demonstrate that phonology can benefit by embracing the data and methods from fields as diverse as acoustics, psychology, and ethology. Phonology does not lose its identity by this; what sets phonology off from other disciplines is its *questions*, its *end*, not its *methods* or *means*. In phonology (if not in ethics), the end justifies the means.

### Acknowledgements

Supported in part by grants from the National Science Foundation, The Harry Frank Guggenheim Foundation, and the University of California (Berkeley) Committee on Research. My thanks for help and advice to: Mariscela Amador, Louis Goldstein, Jeri Jaeger, James Lorentz, and Steve Pearson. Responsibility for the data and opinions offered are the author's.

### References

- Baker, W. J., Prideaux, G. D., & Derwing, B. L. 1973. Grammatical properties of sentences as a basis for concept formation. *J. Psycholinguistic Res.* 2. 210-220.
- Bloomfield, L. 1922. Review of *Language, its nature, development, and origin*. by O. Jespersen. *Am. J. Philology* 43. 370-373.
- Chao, Y.-R. 1934. The non-uniqueness of phonemic solutions of phonetic systems. *Bull. Inst. of*

- Hist. & Philol.. Acad. Sinica 4. 363-397.
- Chastaing, M. 1965. Dernières recherches sur le symbolisme vocalique de la petitesse. *Revue philosophique* 155. 41-56.
- Craig, J. C. 1979. A confusion matrix for tactually presented letters. *Perception & Psychophysics* 26. 409-411.
- Deese, J. & Hulse, S. 1967. *The psychology of learning*. New York: McGraw-Hill.
- Fant, G. 1973. *Speech sounds and features*. Cambridge: MIT Press.
- Fant, G. 1980a. Vocal source analysis—a progress report. *Speech Transmission Lab. [Stockholm], Quart. Progress & Status Rep.* 3-4/1979. 31-53.
- Fant, G. 1980b. Voice source dynamics. *Speech Transmission Lab. [Stockholm], Quart. Progress & Status Rep.* 2-3/1980. 17-37.
- Fujimura, O. & Lindqvist, J. 1971. Sweep-tone measurements of vocal tract characteristics. *J. Acoust. Soc. Am.* 49. 541-548.
- Garner, W. R. 1978. Aspects of a stimulus: features, dimensions, and configurations. In E. Rosch & B. Lloyd (eds.), *Cognition and categorization*. Hillsdale: Lawrence Erlbaum Associates. 99-133.
- Gilmore, G. C., Hersh, H., Caramazza, A., & Griffin, J. 1979. Multidimensional letter similarity derived from recognition errors. *Perception & Psychophysics* 25. 425-431.
- Grierson, G. A. 1922. Spontaneous nasalization in the Indo-Aryan languages. *J. Royal Asiatic Soc.* 1922. 381-388.
- Jaeger, J. J. 1980a. Testing the psychological reality of phonemes. *Lang. & Speech* 23. 233-253.
- Jaeger, J. J. 1980b. *Categorization in phonology: an experimental approach*. Doc. diss., Univ. of Calif., Berkeley.
- Jakobson, R. & Waugh, L. R. 1979. *The sound shape of language*. Bloomington: Indiana Univ. Press.
- Jespersen, O. 1922. Symbolic value of the vowel *i*. *Philologica [London and Prague]* 1.
- Lotz, J., Abramson, A. S., Gerstman, L. J., Ingemann, F., & Nemser, W. J. 1960. The perception of English stops by speakers of English, Spanish, Hungarian, and Thai: a tape-cutting experiment. *Lang. & Speech* 3. 71-77.
- Morton, E. W. 1977. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Naturalist* 111. 855-869.
- Nichols, J. 1971. Diminutive consonant symbolism in Western North America. *Language* 47. 826-848.
- Ohala, J. J. 1975. Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman, & J. Ohala (eds.), *Nasálfest: Papers from a symposium on nasals and nasalization*. Stanford: Language Universals Project. 289-316.
- Ohala, J. J. 1979. The contribution of acoustic phonetics to phonology. In B. Lindblom & S. Öhman (eds.), *Frontiers of speech communication research*. London: Academic Press. 355-363.
- Ohala, J. J. 1980. The application of phonological universals in speech pathology. In N. J. Lass (ed.), *Speech and language. Advances in basic research and practice*. Vol. 3. New York: Academic Press. 75-97.
- Ohala, J. J. 1982. The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage. *The production of speech*. New York: Springer-Verlag. 189-216.
- Ohala, J. J. & Amador, M. 1981. Spontaneous nasalization. *J. Acoust. Soc. Am.* 69. S54-S55.
- Sapir, E. 1929. A study in phonetic symbolism. *J. Experimental Psychol.* 12. 225-239.
- Sapir, E. The psychological reality of phonemes. *J. de Psychologie Normale et Pathologique* 30. 247-265.
- Sawashima, M. 1969. Devoiced syllables in Japanese. *Ann. Bull., Res. Inst. of Logopedics & Phoniatics, Univ. of Tokyo* 3. 35-41.
- Sweet, H. 1874. *History of English sounds*. London: Triibner.
- Thorndike, E. L. 1945. On Orr's hypothesis concerning the front and back vowels. *Brit. J. of Psychol.* 36. 10-14.

- Twaddell, W. F. 1935. On defining the phoneme. *Language Monographs* No. 16.
- Ulan, R. 1978. Size-sound symbolism. In J. H. Greenberg, C. A. Ferguson, & E. Moravcsik (eds.), *Universals of human language*. Vol. 2: Phonology. Stanford: Stanford Univ. Press. 527-568.
- Wang, W. S.-Y. & Crawford, J. 1960. Frequency studies of English consonants. *Lang. & Speech* 3. 131-139.
- Westermann, D. 1927. Laut, Ton und Sinn in westafrikanischen Sudan-Sprachen. *Festschrift Meinhof*. Hamburg. 315-328.
- Whitney, W. D. 1867. *Language and the study of language*. New York: Charles Scribner & Co.
- Winitz, H., Scheib, M. E. & 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.