684

Journal of Phonetics (1989) **17,** 135–143

Feature Fitting: A comment on K. N. Stevens' "On the quantal nature of speech"

Michael Studdert-Kennedy

Department of Psychology, University of Connecticut, U.S.A., Department of Linguistics, Yale University, U.S.A. and Haskins Laboratories, New Haven, CT, U.S.A.

---

1. Introduction

Why are the sound patterns of language the way they are? To answer this question we must have a description of the sound patterns, a conception of the communicative function that they fulfil, and some notion of the physical and psychological conditions that constrain them. If we assume, as does quantal theory (QT), that the sound patterns are properly described by distinctive feature theory, then the task is to show how the human vocal tract and auditory system have shaped features to their function. QT apparently assumes, though does not explicitly state, that the function of phonological form is to assure maximum perceptual distinctiveness at least articulatory cost (cf. Lindblom, 1983, 1986). QT defines distinctiveness not in terms of contrast across the entire phonological system of a particular language (as we might have expected from an heir to Jakobsonian theory), but in terms of the "salience" of isolated features, and, to some extent, in terms of contrast between pairs of features, said to be neighbors in phonetic space. Similarly, QT defines the articulatory cost of individual features in terms of the precision required to execute each one: the cheapest (or easiest, or least effort) articulations are those that demand least precision.

QT then undertakes to demonstrate the existence of "quantal" regions of articulatory space where salient acoustic (and auditory) spectral patterns remain "stable" in the face of "small modifications" in articulatory configuration. For a persuasive demonstration that these regions exist, QT must provide at least: (1) a consistent definition of what constitutes a region of acoustic/auditory stability, and (2) a consistent criterion of articulatory precision (how small is a "small modification"?). In what follows, I shall argue that QT fails on both accounts, forfeiting coherence by repeatedly shifting the grounds on which quantal regions are claimed.

2. Acoustic stability

The word "stable" has two possible meanings in the present context: (1) unchanging over some period of time during the execution of a particular utterance, and therefore static; (2) unchanging over different utterances, and therefore invariant. The two meanings are not incompatible: a static pattern might be invariant, and an invariant pattern might be static. Nor are they mutually entailed: a static pattern might vary from one utterance to another, and an invariant property might be a recurrent pattern of change.

0095–4470/89/010135 + 09 $03.00/0

In its initial programmatic statement, QT unambiguously adopts the first meaning of stable. Thus, in the explication of Stevens' (1989)[1] Fig. 1 we read: ". . . as the articulatory state undergoes a continuous sequence of maneuvers toward and away from the target value, the acoustic parameter resulting from this articulatory gesture may remain relatively stable over some part of the sequence. . . In the acoustic signal, therefore, there will be an alternation between *temporal* regions where the acoustic parameters remain relatively steady, and narrow regions marked by acoustic events where there are rapid changes" (p. 5, my emphasis). A region of rapid change is said to be a "landmark in the acoustic stream" (p. 5), but is not itself the correlate of a distinctive feature. The feature correlates are rather ". . .the articulatory and acoustic attributes that occur within the plateau-like regions of the relations. . ." (p. 5).

In this formulation, the abscissa of Stevens' Fig. 1 represents articulatory variations over space and time, an idealized trajectory along a single dimension during the execution of a single utterance; and the acoustic correlates of features are static patterns, from which temporal variations are explicitly excluded. QT hints at a possible perceptual role for such variations by characterizing regions of rapid change as acoustic "landmarks". But, with the possible exception of the role assigned to "abrupt onsets and offsets" in the *continuant–non-continuant* contrast (p. 37), the hint is not elaborated. One reason for this is that in most subsequent formulations QT abandons the model of Fig. 1.

In fact, the only examples that conform to the model seem to be those that concern certain acoustic changes as a function of constriction size or degree. We can illustrate this with Stevens' Fig. 19, where the abscissa, like that of his Fig. 1, represents a trajectory in space and time along a single articulatory dimension, namely, the area of the supraglottal constriction during the production of an intervocalic fricative. The text describes the figure as follows: "As the active articulator moves from the relatively unconstricted position for the vowel to a constricted configuration for the consonant, the noise generated near the constriction increases to a maximum and its amplitude remains essentially constant as the constriction size decreases, the articulator reverses its direction of movement, and the constriction size begins to increase. The articulator can undergo a continuous smooth movement, and there is no requirement that it remains in a fixed position over the duration of the fricative consonant in order to maintain a constant amplitude" (p. 23). Notice that, in this example, the stable property is amplitude, not formant frequency pattern which is highly sensitive to changes in degree of constriction (see below).

For another example, consider the utterance, *east* ([ist]). Here, the tongue passes through three degrees of constriction with acoustically (and auditorily) distinct properties corresponding to features of the three segments (Stevens, 1972, pp. 53–54). Although the "landmark" regions of rapid change are not perceptually salient (as they presumably should be, if they have any perceptual function), continuous spatio-temporal changes in articulation along a single dimension do indeed give rise to discontinuous changes in the acoustic output, exactly as QT proposes. What we have, then, in his Fig. 1, is a model of aspects of the temporal process of speaking and, by implication, of listening.

However, when we turn to the effects of constriction position, QT abandons process and adopts the second meaning of "stable" given above (i.e. invariant across utterances), implicitly redefining the abscissa of Stevens' Fig. 1 to represent variations over space, but not over time. Thus, the abscissae of his Figs 3, 4, 8, 11, 13 and 22, on which QT rests

---

[1] All subsequent references are to this paper unless otherwise stated.

much of its case, represent the length of the back cavity as a purely spatial dimension. Accordingly, regions of acoustic stability corresponding to features of certain consonants and vowels are no longer ". . .temporal regions where the acoustic parameters remain relatively steady" (p. 5) during an articulatory trajectory through space and time. Rather, they are regions where the same (i.e. invariant) acoustic pattern will emerge from slightly different places of articulation in repetitions of a given consonant or vowel, either in the same or in different contexts. That this is the correct interpretation of these figures is evident from the fact that articulatory parameters which would necessarily also change if constriction position were changing over time (i.e. length and cross-sectional area of constriction) are held constant.

QT does not acknowledge the shift in its definition of "stable" from "static within an utterance" to "invariant across utterances". But it is evident that the shift was necessary because the model of Stevens' Fig. 1 simply does not apply to contrasts in place of articulation as it may to certain aspects of contrasts in manner. Consider, for example, the utterance, *noon* ([nun]). Here, there are no abrupt discontinuities in the spectral pattern as the tongue enters and leaves the vowel region; nor will the pattern remain static (as for the fricative of his Fig. 19), unless the tongue holds a fixed position for some portion of the vowel. Rather, the "continuous smooth movement" of the tongue will give rise to a continuous smooth trajectory through acoustic space.

Of course, such a pattern of change might still be invariant across utterances, and this seems to be what QT intends in its treatment of the converging formant patterns that characterize velar consonants (Stevens' Fig. 14) and pharyngeal stop consonants (Stevens' Fig. 17): "Perturbations in the place of articulation for the velar consonant are not expected to modify greatly the basic acoustic attributions of a midfrequency spectral prominence that is a consequence of the two converging formants" (p. 24). Here, the "perturbations" are clearly not analogous to the slight modifications of constriction area in the course of a single utterance (as for the fricative of Fig. 19). Instead, they are the slight differences in place of articulation that may occur across repetitions of an utterance.

In short, QT does not offer a coherent definition of its first key concept, acoustic stability. As I shall argue below, it fares no better in its attempts to define its second key concept, articulatory precision.

3. Articulatory precision

QT does not explicitly define "precision". But the treatment of constriction position outlined above evidently commits the theory to a definition of articulatory precision as the reciprocal of the variation in target attainment across repetitions of an utterance. Two classes of gesture may then be compared in terms of their relative ranges of variation across utterances: the larger the range, the less precise the articulation. However, as we shall see, QT does not maintain a consistent quantitative criterion of precision for different articulatory parameters. Nor does it hold to the implicit definition of precision in terms of range of variation across utterances for either consonants or vowels.

3.1. *Vowels*

For vowels, QT describes the effects of variation in constriction position for three pairs at, or close to, the points of the vowel triangle: /i,e/, /u,o/ and /a,æ/. For each pair QT specifies constriction positions, with respect to distance from the glottis, ranging over

1–2 cm, within which there occur stable spectral prominences due to the proximity of either F_2 and F_3, or of F_1 and F_2. The criterion for an acceptable range of articulatory variation in constriction position seems therefore to be 1–2 cm.

As already noted, these vowel regions of stability are defined entirely in terms of constriction position, with length and cross-sectional area of constriction and liprounding held constant. If the latter parameters were included in a full multidimensional description of the vocal tract configuration as it changes over time, regions of stability would disappear. In fact, QT acknowledges (Stevens', 1989, Section 2.5) that changes in the latter parameters do not give rise to regions of formant stability, but offers several extenuating arguments to discount the departures from the quantal model.

First, QT claims that "... formant frequencies are usually not strongly sensitive to constriction size under most conditions" (p. 15), and supports this claim by referring to the resonators of Stevens' Fig. 2(b), modeling the low vowels, and to the corresponding plots of his Fig. 5. However, for the non-low front vowels (his Figs 2(a), 3, 7 and 8), the situation is rather different. For these vowels, F_2 and F_3 values and spacing ". . .depend on the length and cross-sectional area of the constriction" (p. 11). In his Fig. 8, the cross-sectional area of the constriction (A_c) is set at 0.3 cm^2, corresponding to a constriction diameter ($d = 2r = 2\sqrt{A_c/\pi}$) of approximately 0.62 cm. If we allow A_c to vary from 0.2 cm^2, the minimum required to maintain sonorancy (p. 8), to a maximum of, say, 0.5 cm^2 (his Fig. 3) d will vary over a range of 3 mm, from about 0.50 cm to 0.80 cm; and the distance between F_2 and F_3 will increase from 200 Hz to 400 Hz, if the length of the constriction (l_c) is set at 2 cm, as in his Fig. 3, or to an even higher value, if l_c is set at 5 or 6 cm, as in his Fig. 8[(a), (b)]. Thus, QT proposes a range of acceptable articulatory variation for degree of constriction that is roughly one-fifth of the range proposed for position of constriction.

Quite apart from the disparity of the two ranges, the absolute value of the range for degree of constriction (3 mm) does not seem, intuitively, to allow much imprecision of articulatory action. On the contrary, the narrow range implies that the acoustic signal is highly sensitive to changes in degree of constriction, and that achievement of a constriction in the required range must call for appreciably greater articulatory precision than does achievement of an adequate constriction position.

This was, in fact, precisely the hypothesis of Perkell & Nelson (1982) in an X-ray microbeam study of multiple repetitions of the vowels /i, a, æ/in several consonantal contexts, by two speakers. They hypothesized, that, due to ". . .*location-dependent* acoustic *insensitivity* and *constriction-dependent sensitivity* in the region of maximum constriction" (p. 191, emphasis in the original), there should be a significant quantitative difference in the variability of the two parameters across repetitions. The results supported their hypothesis for /i/ and, less definitively, for /a/. The authors were not able to estimate absolute values of d (which would, in any case, be different for an actual vocal tract than for the idealized tube of QT), but from their Figures 6 and 7 we can measure, for /i/, a range (± 2 standard deviations from the mean position of a pellet on the tongue blade) of some 4–5 mm. Yet even this more generous estimate of the range of variation in degree of constriction is still less than one half of the range reported for position of constriction. This study, then, does not support the claim that". . .formant frequencies are usually not strongly sensitive to constriction size under most conditions" (p. 15), either for non-low front or for low vowels.

Finally, there is the matter of rounding. Here, to counter the acknowledged absence of regions of stability (for which see also Figure 3 of Stevens & House, 1955), QT treats rounding as a dichotomous variable. The theory speculates that a shift from unrounded to rounded increases the proximity of F_1 and F_2 in back vowels, and of F_2 and F_3 in front vowels, thus creating perceptually salient, single-peaked spectral prominences. That these are indeed the effects of rounding is not in question. But we should note that the argument circumvents the monotonic changes due to changes in degree of rounding by implicitly denying their importance.

Having thus discounted the nonquantal effects of continuous change in degrees of tongue constriction and liprounding, QT goes on to reintroduce quantal effects by positing non-linearities at quite another level, namely, that of muscle excitation. This is the point at which QT abandons the definition of precision in terms of variation across utterances, and switches to a definition in terms of a hypothetical mechanism that might, in principle, be applied to individual utterances. For degree of constriction, the theory invokes Fujimura & Kakita's (1979) model of the neuromuscular mechanism involved in assuring the precise degree of constriction size necessary to produce /i/: due to a "saturation effect" in the pattern of genioglossus contraction, variable degrees of contraction give rise to a stable constriction size. Similarly, for rounding QT proposes ". . .anatomical constraints that make it possible to reproduce a particular degree of rounding without requiring a great deal of precision in the degree of excitation of the appropriate muscles" (p. 16). Whether this is indeed a correct account of the mechanism is not at issue: presumably there is *some* mechanism by which continuously variable patterns of firing in neuromotor units are marshalled to achieve precise articulatory targets. Whatever the mechanism, QT can always claim that the mechanism uses imprecise means to achieve precise ends. But this is surely to stretch the meaning of "articulatory precision" beyond reasonable limits. Moreover, one's scepticism is aroused by an argument that first denies the importance of precision and then posits a mechanism for achieving it.

3.2 *Consonants*

Different criteria of articulatory precision for position and degree of constriction are proposed also for consonants. For example, the amplitude of turbulence noise in a fricative consonant is said to remain stable ". . .within 3 dB of its maximum value over a range of constriction sizes from 0.03 to 0.2 cm^2" (p. 21). These cross-sectional areas correspond to a range of constriction diameters of 3 mm from about 2–5 mm. On the other hand, the optimum position for a tongue blade constriction for /s/ or /t/ is said to vary over a range of about 2 cm (Stevens' Fig. 21).

These different ranges for position and degree of constriction are extended to position of constriction itself in the discussion of certain other consonants (Section 2.6). Thus, a pharyngeal consonant affords a spectral peak due to proximity of F_1 and F_2 with a range of articulatory variation of 1–2 cm, while a spectral peak due to proximity of F_3 and F_4 in a retroflex consonant is said to be ". . .insensitive to the exact position of the constriction at least to within 1–2 mm" (p. 19). Both patterns are then cited as examples of consonantal configurations that are ". . .only weakly sensitive to perturbations in the placement of the constriction" (p. 20). QT acknowledges the disparity between these two ranges of "weak sensitivity" and again appeals to the black box: "Presumably. . .the motor and orosensory system is capable of greater precision in positioning a constriction with the tongue blade in the region of the hard palate than with the tongue body in the lower pharynx" (p. 20). Here we have yet another measure of articulatory precision:

precision relative to the capacities of the articulator in question. Yet a presumptively greater innervation ratio for the tongue blade than for the tongue body was not invoked to account for the achievement of F_2–F_3 proximity in the formation of /i/, also with the tongue blade—presumably because the F_2–F_3 prominence extends across a 1–2 cm range of tongue positions (Stevens' Fig. 8), so that no special sensitivity seemed to be called for.

In short, we are offered three quite different definitions of articulatory precision, two of them highly speculative: (1) absolute precision, with a fivefold variation in spatial range from 3 mm to 1–2 cm; (2) precision arising from hypothetical neurophysiological constraints that permit precise articulation without precise excitation of the appropriate muscles; (3) relative precision defined with respect to the presumed capacity of the articulator in question. We surely need a more consistent definition of articulatory precision than whatever happens to be, or is presumed to be, the case for the parameter or articulator under consideration.

4. Quantal effects in auditory processing

Under this heading, QT groups sets of data gathered by the standard experimental paradigms of categorical perception (CP) studies. To assimilate these findings to the quantal model, the abscissa of Stevens' Fig. 1 is given its second (explicit) interpretation, as representing continuous variations in an acoustic pattern that give rise to discrete categories of auditory response, either in the auditory nerve or in psychophysical judgements. As in the (implicit) interpretation assigned to the abscissa for variations in constriction position, the variations form a series of mutually exclusive patterns that (it is assumed) might occur in different executions of the same utterance.

There are several reasons for questioning the validity of assimilating CP studies to the quantal model for constriction position. First, while variations in constriction position might be viewed as permitting the tongue some latitude in the execution of a given feature in different phonetic contexts, CP studies are confined to acoustic variations across repetitions of a feature correlate in exactly the same context. For example, a synthetic continuum from /bæ/ to /dæ/ comprises several exemplars such as might (perhaps) occur in multiple repetitions of the two syllables. The relation between perceptual categories formed from these essentially random variants, and categories formed from the presumably systematic variations that accompany executions of the same consonants before different vowels, or in syllable final rather than syllable initial position, is not known. Accordingly, CP studies have little bearing on a central issue of speech research that QT would seem to address, namely, the basis on which listeners assign the apparently variable acoustic correlates of a feature, executed in *different* phonetic contexts, to the *same* category.

A second reason for doubt arises from the fact that CP effects are often demonstrated on a synthetic continuum for which there is no underlying articulatory continuum. A striking example comes from "one of the basic consonantal distinctions in language. . . [namely,] the distinction between labial consonants produced with a constriction at the lips, and coronal consonants with a constriction formed by the tongue blade" (p. 37). QT displays concern with the legitimacy of drawing inferences from listeners' judgements of "unnatural stimuli" (p. 38), yet seems to attach the same importance to the "threshold phenomenon" (p. 38) observed on a labial–coronal continuum, for which there can be no underlying articulatory continuum, as to, say, the nasal–non-nasal vowel distinction

(p. 34) for which there is. The labial–coronal example throws into question the functional role of the various "natural perceptual boundaries" that QT documents, in determining the acoustic ranges of speech categories.

In fact, several of the papers in Harnad (1987), to which Stevens' refers (p. 30), make clear that the position of the boundary on a synthetic continuum may vary with phonetic, syntactic and semantic context, with speaking rate, with the language that listeners believe the synthesizer to be mimicking, and even with the psychophysical method used to measure it. A fair conclusion from many dozens of CP studies is that category boundaries are labile phenomena, sensitive to a variety of stimulus manipulations and therefore useful as a measure of their effects. Far from a boundary determining the categories that it separates, as QT implies by its invocation of "natural perceptual boundaries", it would seem that the categories determine the boundary.

In conclusion, by leaning on CP studies (to which roughly a third of Stevens' paper is devoted), QT abandons not only temporal process (CP boundaries are certainly not "landmarks in the acoustic stream"), as it does also in the treatment of constriction positions, but even the articulatory substrate that lends some weight to the earlier formulations. CP would appear to be largely a laboratory phenomenon, a function of currently available techniques for the analysis and synthesis of speech, with little relevance to the normal processes of speaking and listening.

5. Features as functional units

Let me turn finally and briefly to a broader aspect of QT: its commitment to distinctive feature theory. The difficulties that the commitment raises are most readily apparent in QT's approach to the evolution and development of language, although these topics are not explicitly addressed by Stevens' paper. "Language seeks out regions [of stability], as it were, and from them assembles an inventory of phonetic elements that are used to form the code for communication by language" (Stevens, 1972, p. 64). An analogous view of individual development is implied by Stevens & Blumstein (1978). They propose that the auditory system is endowed with specialized detection mechanisms and that these mechanisms give children access to certain primary and invariant attributes of the speech signal. From these, children learn, by processes of association, to construct the full pattern to which they will respond as adults.

There are several things wrong here. First, neither evolution nor ontogeny proceeds from the specific to the general, by constructing complex forms from pre-existing simpler components. The normal process, in both morphology and behaviour, is one of differentiation by which smaller structures emerge from larger. We have no reason to suppose that language reverses the process.

In fact, there is increasing evidence that the child's entry into language is mediated by meaning, and that the earliest units of contrast are neither features nor phonemes, but words and formulaic phrases (e.g. Ferguson, 1986). Of course, a word, even if it were perceived as a whole, could not be reproduced as a whole: even a simple CV syllable requires the co-ordinated actions of several more or less independent articulators. The child must therefore come to differentiate the sound pattern of a word into auditory pieces that correspond to natural units of articulatory action, or gestures. On this account, gestures rather than features are the ultimate functional units of spoken language (cf. Browman & Goldstein, 1986). The integral patterns of gesture that we term phonemes and their descriptive predicates, features, would then emerge from self-organizing processes in the growing lexicon (Lindblom, MacNeilage & Studdert-Kennedy, 1983; Studdert-Kennedy, 1987).

QT itself uses the concept of an articulatory gesture, but strips it of its dynamic properties, by defining it as "...a continuous sequence of maneuvers toward and away from ...[a] target value" (p. 5). Thus, a gesture is specified by a target vocal tract configuration, and the target configuration by the static acoustic property to which it gives rise. QT does not view the co-ordination of these gestures (each corresponding to a feature) as a problem: they are simply "simultaneously executed [as a] bundle of features, or a segment" (pp. 41–42). Thus, the temporal properties of speech are reduced to a serially ordered sequence of discrete segments.

Whether a child could learn to reproduce the subtleties of its native dialect from the auditory equivalent of a rapid succession of still photographs, we may doubt. Certainly, no one, so far as I know, has attempted to develop a theory of speech production—or even a program for the articulatory synthesis of speech—from nothing more than a sequence of static targets. QT, it seems to me, has mistaken the problem for its solution.

6. Conclusion

I have argued that QT is not coherent as a theory and that it is flawed by a premature attempt to fit phonetic fact into the framework of distinctive feature theory. Many years ago, Fant (1962) observed that the theory of Jakobson, Fant & Halle (1963) was limited because "...its formulations are made for the benefit of linguistic theory rather than for engineering or phonetic applications. Statements of the acoustic correlates to distinctive features have been condensed to an extent where they retain merely a generalized abstraction insufficient as a basis for the quantitative operations needed for practical applications" (p. 4). We may see QT as a response to Fant's implicit challenge to go beyond "generalized abstraction" to more precise specification of features. Whether QT has gone far enough to permit either practical applications in, say, machines for automatic speech recognition and synthesis, or even phonetic applications in an account of human perception and production, the reader must judge.

My own conclusion, of course, is that it has not. Nonetheless, QT has the rare virtue of adopting a functional rather than a purely formal approach to phonology. The theory is original and important because it draws attention to the sort of considerations that will have to go into any satisfactory account of the biological bases of phonological form.

My thanks to Cathe Browman for very useful comments on an earlier version. Preparation of the paper was supported, in part, by NIH grant HD 01994 to Haskins Laboratories.

References

Browman, C. & Goldstein, L. (1986) Towards an articulatory phonology, *Phonology Yearbook*, **3**, 219–252.

Fant, C. G. M. (1962) Descriptive analysis of the acoustic aspects of speech, *Logos*, **5**, 3–17.

Ferguson, C. A. (1986) Discovering sound units and constructing sound systems: it's child's play. In *Invariance and variability in speech processes* (J. S. Perkell & D. H. Klatt, editors) pp. 36–51. Hillsdale, NJ: Erlbaum Associates.

Fujimura, O. & Kakita, Y. (1979) Remarks on the quantitative description of lingual articulation. In *Frontiers of speech communication research* (B. Lindblom & S. Öhman, editors) pp. 17–24. London: Academic Press.

Harnad, S. (editor) (1987) *Categorical perception*. New York: Cambridge University Press.

Jakobson, R., Fant, C. G. M. & Halle, M. (1963) *Preliminaries to speech analysis: the features and their correlates*. Cambridge: MIT Press.

Lindblom, B. (1983) Economy of speech gestures. In *The Production of Speech* (P. F. MacNeilage, editor) pp. 217–245. New York: Springer-Verlag.

Lindblom, B. (1986) Phonetic universals in vowel systems. In *Experimental Phonology* (J. Ohala & J. J. Jaeger, editors) pp. 13–44. Orlando, FL: Academic Press.

Lindblom, B., MacNeilage, P. & Studdert-Kennedy, M. (1983) Self-organizing processes and the explanation of phonological universals. In *Explanations for Language Universals* (B. Butterworth, B. Comrie & O. Dahl, editors). pp. 181–203. The Hague: Mouton.

Perkell, J. S. & Nelson, W. L. (1982) Articulatory targets and speech motor control: a study of vowel production. In *Speech Motor Control* (S. Grillner, A. Persson, B. Lindblom & J. Lubker, editors) pp. 187–204. Oxford: Pergamon.

Stevens, K. N. (1972) The quantal nature of speech: evidence from articulatory-acoustic data. In *Human Communication:* (E. E. David, Jr. & P. B. Denes, editors). pp. 51–66. New York: McGraw Hill.

Stevens, K. N. (1989) On the quantal nature of speech, *Journal of Phonetics*, **17**, 3–45.

Stevens, K. N. & Blumstein, S. E. (1978) Invariant cues for place of articulation in stop consonants, *Journal of the Acoustical Society of America*, **64**, 1358–1368.

Stevens, K. N. & House, A. S. (1955) Development of a quantitative description of vowel articulation, *Journal of Acoustical Society of America*, **27**, 484–493.

Studdert-Kennedy, M. (1987) The phoneme as a perceptuomotor structure. In *Language Perception and Production* (A. Allport, D. MacKay, W. Prinz & E. Scheerer, editors). pp. 67–83. London: Academic Press.