Classifying allophones

Mark Tatham

Reproduced from *Occasional Papers* No. 3 (March 1969): The Language Centre, University of Essex.

Also published as: Tatham, M. A. A. (1971) Classifying allophones. *Language and Speech* 14, 140-145

This paper discusses the basis for the categories 'intrinsic' and 'extrinsic' allophones, that is articulatory effects which are neuro-physiologically determined and those which are language specific, and argues the need for introducing a third category to include effects which are the result of inhibiting an 'intrinsic' influence on articulation.

Following Wang and Fillmore (1961) an explicit distinction was made by Ladefoged (1965, 1967) between two types of allophone: *extrinsic* and *intrinsic;* these have been discussed in various papers, particularly by Kim (1966) and Tatham (1968). I want to continue here the discussion of the implications of this distinction between 'programmed' (extrinsic) and 'automatic' (intrinsic) allophones in the context of the development of what might be called a performance or production model in phonetics.

Kim (following a pre-print of Ladefoged 1967) does not agree with Ladefoged's view that like extrinsic allophones intrinsic allophones should be regarded as language-specific; he argues that they should be considered as universal. The basic properties of the human vocal apparatus are similar for all human beings, irrespective of their language; any physiological constraints on articulation operating on any one vocal tract will also operate on all other normal vocal tracts. According to Kim, conjoining rules of the type proposed by Lindblom (1964) and Ohman (1966) account adequately for the missed target phenomenon. Given similar circumstances and context (same temporal restriction, etc.) any series of identical target positions will be missed equally and similarly by any human being. Ladefoged's argument for language-specific intrinsic allophones is illustrated by the slight, yet noticeable, variations in the articulation of plosives in French as compared with English. I shall argue that the target positions of these 'same' phonemes in English and French are different (i.e. the extrinsic allophones are different in neuro-physiological terms) and that therefore the comparison is invalid. /p/ of French and /p/ of English are equatable in terms of *distribution* (i.e. on linguistic grounds); this does not give them phonetic similarity.

I prefer to conflate the view that intrinsic allophones should be regarded as language-specific and the view that they should be regarded as universal. Let us assume the example of two languages each with the morpheme /ka/, realised in L1 as [«ka] and in L2 as [»ka] [fronted [k] and retracted [k] respectively]. Lindblom's and Ohman's conjoining rules, based on universals of articulation, would indicate that /k/ will tend to be pronounced further back when co-articulated with /a/ than when co-articulated with /i/, for example. If we find that the /k/ in L1 is forward from the target position then L1 does not conform to the rule, or the rule is wrong, or the theory behind the rule is wrong. This gives rise to the language-specific view: since in L1 the /a/ does not influence the /k/ or influences it in a different way from L2, a different rule is required for each language. This automatically lifts the conjoining rules from the level of physiological constraint to the level of linguistic requirement and destroys the attempt to formulate some kind of statement of universal physiological constraints.

It is difficult to extrapolate from Kim's monograph and guess his attitude to the anomaly. My own view is that any attempt to formulate universal rules about physical constraints,

whether neurological, physiological or acoustic, is worthwhile. There are a good many limitations at these levels, and doubtless a number of things happen to the articulators that are nothing to do with the direct commands from the motor system.

Let us assume for the moment that often the central processing mechanism *intends* one event and another *occurs* because of these constraints. There may be, for example, a particular target gesture for /k/; on occasions however, 1 because of context the articulators miss the target by this or that amount. It is very important (but often very difficult) in the construction of any model to separate out at an early stage those phenomena which appear to occur for different basic reasons: therefore, it seems appropriate to talk of the gesture /k/ on which are superimposed certain co-articulation effects due to a predictable physiological constraint.

But this does not solve the anomaly of the occurrence of fronted /k/ before /a/. As mentioned the co-articulation rules based on observation would lead us to expect that the /k/ will be retracted; it seems therefore that this is an exception. Perhaps the best way of overcoming the problem would be to state that in a language of type L1l the natural physiological tendency to back the /k/ is counteracted — *it is this counteraction which I wish to regard as language-specific, not the co-articulation*. Thus I would prefer not to say that the intrinsic allophones are language-specific, but would prefer to talk in terms of the model itself rather than in terms of the output of the model.

If the rules are correct for predicting co-articulation phenomena, then a neutral /k/ *must* be retracted when preceding /a/; but in addition we must account for the observable fact that certain languages do not succumb to this retraction — I am suggesting that the natural tendency is counteracted. By positing that the fronting of /k/ in a context where one would normally expect a retraction *(not because most languages produce a retraction but because the physiology dictates this as the normal unchecked state of affairs — i.e. not as a* regularity *but as a* rule) is voluntary, I am placing this phenomenon in the category of extrinsic events (and therefore language-specific). I think that it may be the desire to account for this kind of phenomenon which forces Ladefoged to regard even intrinsic allophones as language-specific. If his definition is kept it will be necessary to admit a third level: that where events are predictable because of physiological constraints as well as language-specific intrinsic allophones. I prefer to say that all voluntary effects are the result of linguistic considerations and that these are defined as extrinsic, and that superimposed on these are constraints which will occur unless prevented as a result of not controlling the particular articulatory parameters.

The human vocal apparatus is fairly constant irrespective of race, etc. There are certain irrelevant differences: the average fundamental frequency of voiced sounds in women is higher than in men because their vocal cords are consistently shorter, etc. But by and large the available components and their manner of control are the same. There are limitations on the use of these components or organs. Some of these are

- physiological (e.g. the nasal cavity can only be blocked at one end [by the velum] — the other end of necessity remains open);
- physical (e.g. a voiced stop can only continue to be voiced for a certain period of time if the stop is not released (because the air pressure above the vocal cords will soon equalise with the air pressure below, thus causing the vocal cords to stop vibrating, all other things being equal);
- neurological (rate of neural transmission of impulses will severely limit the upper rate of control, or the possibilities of active proprioceptive feedback, etc.).

These limitations will operate for all normal people in all languages irrespective of any voluntary mechanism — their effects cannot be changed and can in principle be stated mathematically with defined limits.

In addition there will be effects which will occur or tend to *occur unless there is a specific command for them not to occur*. Into this category we can place the example of /k/.

Unless there is a specific command to the contrary, /k/ will retract under the influence of the following /a/.

There is every evidence that the division consonant / vowel may be a natural division and that control of consonants and vowels is effected somewhat differently. This assertion is made more plausible by the notion of superimposition of consonant gestures on a continuous vowel gesture and by the notion of speech posture (Ohman 1967, Gårding 1967, Chomsky and Halle 1968). Suppose that a command issuing from the phonological processor has been sent to initiate the syllable /ka .../. The basic speech posture is set up (tonic muscle activity is increased) well in advance of any articulation that might be recognised as /k/ or /a/. There is no sound for acoustic feedback correction.

The basic vowel gesture, according to the above model, covers the entire syllable and is established concurrently with the command for a superimposition of /k/ on this gesture and before it. In other words (to be more specific than the above-cited authors) the vowel gesture takes precedence; information concerning the state of the muscles is received at the motor centre (*via* the sensory loops) and processed for modifying commands *with respect to the vowel and not with respect to the consonant*. This notion of sensory feedback, essential to the present model, is an extension of the former co-articulation model. While the superimposed command for the /k/, relatively independent of influential sensory feedback, is being delivered, the control mechanism is adjusting the basic speech posture from a 'neutral' position to a target posture for the /a/ — assuming that vowels (or syllable nuclei) are not independent of sensory feedback. [In the case of vowels here, the gamma loops and other mechanisms provide 'starting-point' information for the motor system.] In adjusting the vocal tract configuration for the /a/ the place of articulation of the /k/ is changed.

The suggestion is that the preceding is what normally happens. But in those languages where there is a phonemic difference in different types of /k/ or where there is an importance to establishing variations in extrinsic allophones of /k/, then the influence of the /a/ over the /k/ is resisted.

It might be a good idea to index those physiological phenomena which cannot be resisted, and those which will happen unless resisted. It is necessary to make an important distinction between those events which happen because they are under direct control and those events which happen because something else has happened but not because they themselves were actually controlled. An event which co-occurs without control and under the mechanical influence of the dominant and controlled event I shall call an intrinsic event. The resistance to this intrinsic event I shall call extrinsic resistance whilst preserving the intrinsic status of the normal (uncontrolled) event. This use of 'extrinsic' and 'intrinsic' conforms with the original use by Wang and Fillmore (1961), but is somewhat different from Ladefoged's and Kim's use.

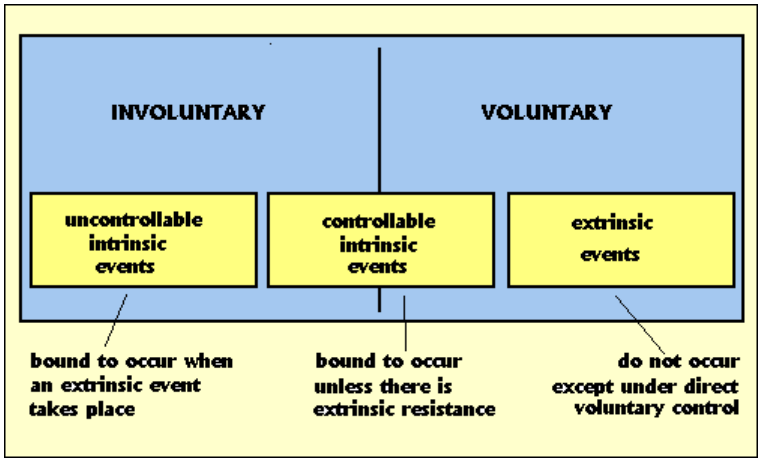

Figure 1 The relationship between different types of intrinsic event

Fig. 1 diagrams this procedure. The additional constraint which is extrinsic on the degree to which a particular intrinsic phenomenon is allowed to occur can easily be incorporated into the type of co-articulation rule proposed by Ohman. Kim mentions a 'freedom-factor' indexed to a particular gesture indicating the extent to which, for example, a muscle may move, or the extent to which is may be limited in its movement by other factors: this is an intrinsic feature. I am suggesting that a similar weighting factor might be added which is language-specific and which will considerably add to the weighting factor suggested by Ohman. For example, let us take the case of /ka/ again. In language L2 this is realised as [»ka] and in language L3 as [ka] where k without a diacritic means 'neutral target k'. In the case of L3 a language-specific weighting must be given to the co-articulation rule to nullify the retraction operation.

The addition here of a language-specific weighting or limiting factor is part of a more accurate model. I do not take the view that a simple indexing of the type: '/k/ before /a/ in language L3 is not subject to co-articulation rule R' is appropriate. The point is that co-articulation rule R will operate automatically and that this operation must be nullified. Only in this way will a truer picture of actual events (as postulated by the model diagrammed in Fig. 1) be presented. It is not the case in this model that there is some instruction to the articulators that the intrinsic co-articulation effect should not be operated on this occasion but that on the assumption that it will operate a complex of commands is sent out hindering or negating that operation. It is in this sense that the non-operation of a neuro-physiological determined (not just observed phenomenon in the majority of languages) tendency is called extrinsic: there is a voluntary restriction, precisely analogous to the voluntary setting up of any particular extrinsic allophone configuration.

By allophone we mean not variations in sound but a complex of commands sent to the articulators (in the case of extrinsic allophones) to produce a particular segment of sound which may be intricately involved in the syllable or other complex of commands. It is not the result of that command; there can be no utterance of an extrinsic allophone and in this sense the concept is an abstraction similar (but not identical) to the phoneme.

Kim has expressed the opinion that the general rule can be set up that extrinsic allophones are important in perception and intrinsic allophones are not. Given the caution of the preceding paragraph this may be interpreted as saying that in the perceptual process somehow the co-articulation effects actually present in the articulation or manifested in the quasi-continuous soundwave are negated and that the soundwave is reinterpreted in terms of the extrinsic allophone — i.e. in terms either of the intended allophone or of a target deduced from knowledge of the co-articulation rules and intrinsic limiting factors in articulation.

Kim's reason for aiming to separate completely the levels of extrinsic and intrinsic allophones and their associated processes is that he wishes to have the intrinsic allophone level determined entirely automatically.

Employing the term 'systematic synthesis' to stand for a model of production, he recognises Halle's level of systematic phonetics as the output of the phonological component of a transformational grammar and asserts that there follows another level (of extrinsic allophones) linked by the rules of systematic synthesis. This is the limit of generation in the strict linguistic sense (in my terminology, voluntary control of articulation); after this what 'happens' to extrinsic allophones (presumably these are now in terms of complexes of motor commands, although Kim does not say so) is dependent upon uncontrollable neuro-physiological factors. I tend to take the view for the present that the levels can only be a convenience of approach to the problem. Feedback mechanisms spanning levels and the complex sensory system involved in actual muscle control will quickly blur any attempt to separate levels in anything by a crude abstract fashion. In particular there may be considerable importance in the voluntary limiting of involuntary reactions. Let us take a very simple example to conclude: if I put my hand into a flame, after a very short period of time an involuntary reflex system comes into play causing me to withdraw my hand; however, it is perfectly possible for me to override this reflex action in favour of a voluntary decision to allow the hand to burn.

References

Chomsky, N. and M. Halle (1968) *The Sound Pattern of English.* Harper Row: New York

Gårding, Eva 1967) Internal Juncture in Swedish. *Travaux de l'Institut de Phonétique de Lund*

Kim, C-W. (1966) The Linguistic Specification of Speech. *Working Papers in Phonetics* 5, UCLA

Ladefoged, P. (1965) The Nature of General Phonetic Theories*. Georgetown University Monograph on Languages and Linguistics* No. 18: Georgetown University Press: Washington DC

Ladefoged, P. (1967) Linguistic Phonetics. *Working Papers n Phonetics* 6, UCLA

Lindblom, B. E. F. (1964) Articulatory Activity in Vowels. *STL-QPSR* 2/1964, KTH: Stockholm

Ohman, S. E. G. (1967) Numerical Model of Co-articulation. *JASA* 41

Tatham, M. A. A. (1968) Specifying Speech Phonetics. *Department Report* No. 1, University of Leeds

Wang, W. S-Y. and C. J. Fillmore (1961) Intrinsic cues and consonant perception. *J. Speech and Hearing Research* 4