Coarticulation and Phonetic Competence

Mark Tatham

Paper read at the Atlantic City meeting of the Acoustical Society of America, April 1970 — reproduced from University of Essex, Language Centre, *Occasional Papers* No. 8. This work was supported in part by grant B/SR/6733 from the Science Research Council.

If we consider the variables which must be incorporated in a formula or rule concerned with describing coarticulation, then we meet right away at least two weighting factors which must be taken into account. I am thinking here of rules of the form devised by Ohman (1967) a few years back, when he formalised a descriptive statement predicting the position at any moment in time of a given point on the surface of the tongue. Briefly, Ohman hypothesised that, considering only the tongue parameter, there is a continuous dominant vowel gesture on which any (in this case, lingual) consonants are superimposed.

The weighting factors I mentioned seem to fall roughly into two major categories:

1. those concerned with physical or mechanical constraints on movement or control of movement — further sub-divided into absolute constraints and constraints at a certain time or in a certain context, and
2. those concerned with linguistic constraints — once again further sub-divided into absolute constraints and constraints at a certain time or in a certain context.

My paper is concerned with taking a look at these two categories of constraints on articulation and saying something about how they might fit into the grammar's phonetic component or into some speech production model which is not itself initially constrained by the output of the 'previous' component.

My category 1. is those mechanical, sometimes neuro-mechanical, constraints which form the focus of observations on coarticulation: let me insert here that I think this focus is a little misdirected, since, for the linguist, once the observation is made that such constraints exist and have been duly noted, the interest is dulled. Mechanical constraints occur, we can guess, right through articulation: they have been noted with a certain amount of enthusiasm because they form an important key in overcoming the apparent problem of poor fit occurring between segmented-string levels and non-segmented-string levels or prosodic levels. Thus for many researchers concerned with the coarticulation phenomenon, description, based upon observation of the end effect, constitutes the goal: non-segmentation in speech, it has seemed, can be largely explained by the complex interaction of time and neural or mechanical inertia.

Supposing now we look at the accountable data from the point of view of the linguist: the phonetic component, for example, might continue the competence/performance distinction — it is not necessarily the case that phonetics equals performance. 'Knowing' about which sequence of articulations is required and, further, knowing about mechanical constraints, commands to the appropriate muscles can be computed: something of this kind is involved in setting up the computer control of speech synthesis as a *working model of speech production.*

Notice, however, that this kind of knowledge is inherently a-linguistic and probably should not, except peripherally, be included, as mere observation classified into tables, in the phonetic competence part of any linguistic model. However if it did eventually turn out to be the case that phonetics is a-linguistic anyway, then competence would now mean something different than, say, competence in phonology — and, as I have said, something linguistically trivial.

Nevertheless such knowledge there must be, and it would perhaps be better to regard it as an essential part of phonetic performance; it might be termed 'capacity' or something similar.

The second weighting factor in coarticulation formulae involves a *linguistic* constraint. Very often a positive or negative bias, as I suppose we could call it, is placed on the use made of mechanical constraint. By bias I mean a weighting superimposed on the mechanical constraint, indicating just how much yield or how much counteraction occurs,

Take an example: Ohman's rules that I mentioned earlier are about the variation of tongue position for an intervocalic consonant, dependent on the target values of the surrounding vowels and upon the time available. They quantify the range of values for the consonantal target. That a range will occur is a mechanical constraint, and it is further a constraint that the range will have certain, mechanically dependent, limits.

But, of course, that cannot be all there is to it. We are familiar with those graphs which plot areas corresponding to a statistical sample of missed target values for a particular phoneme. It must be the case, if perceptual stability is to be maintained, that these areas do *not* overlap too much — that is, that they do not overlap to cause phoneme differentiation problems when other perceptual forces are not recruitable. Thus if, considering front-back coarticulation in VCV utterances where the C is a fricative consonant a particular language has only two palatal fricatives, we would expect the system to succumb more readily to mechanical coarticulation than if three or four palatal fricatives were used in this language — thus the underlying psychological principle of maximal differentiation would limit the coarticulation.

Quite clearly the set of linguistic constraints or limits is contained *within* the set of mechanical limits: it cannot be the case that there are linguistic demands greater than the system can support.

So far, then, we have considered the observable phenomenon of mechanical constraint and the postulated phenomenon of further linguistic constraint — the latter limiting the former in the sense that certain otherwise occurring inertia points can be controlled to a certain extent.

If we take a look at a linguistically defined generative model of speech production, we can single out, among others, two important problems to be tackled: the form of the component and the form of the input (the output being possibly the articulatory level of speech). One reason for the 'phonetics-equals-performance' viewpoint is, of course, the dominance of action over knowledge at this level — that is the meaninglessness of the component when divorced from time. But conceptually there is no reason why a segmental input embodying only notional time in the form of gross sequencing of matrix columns should not be converted to *a finer time bound string*. There is no reason why some kind of differential or threshold statement of knowledge should not be introduced at certain points: by this I mean a substitution for the more usual 'if x, then y' by a statement of the form 'if enough of x, then y'.

It is up to the phonetic theory, which may have internal constraints determined by the articulatory system and external constraints determined by the all-important need to differentiate different speech acts, to determine the form of the phonetic component: while taking over from the phonology, it is above all the introduction of such constraints as time, which dominate, although these will have been all but absent from the phonological component. The phonetic component will be responsible for converting the phonological embodiment of linguistic knowledge of the structure of the speech act into a complex of commands or instructions to the peripheral muscles controlling articulatory movements. Knowing limitations of control takes over, and in addition so does knowing the use or *curbing* of those limitations for language in general ant for one language in particular.

To summarise: I have been concerned in this paper with the status and form of a phonetic component of a grammar. Using a current descriptive model of coarticulation phenomena, I have tried to show that time and neuro-physiological constraints provide the (at first sight) linguistically trivial setting for a complex psychologically determined array of strategies for controlling those constraints. I think the importance particularly to phonological theory should be underlined of the limits set by linguistic dominance of the system constraints.

Reference

Ohman, S. E. G. (1967) Numerical Model of Coarticulation. *JASA* 41