MOTOR CONTROL OF COARTICULATION: LINGUISTIC CONSIDERATIONS

R.A.W. Bladon, Department of Linguistics, University College of North Wales, Bangor, U.K.

Various orientations to the motor control of speech

An orientation advocated recently by Moll, Zimmermann and Smith (1977) calls for priority to be given, in research on speech motor control, to exclusively neurophysiologically-based studies. The need is, they argue, to determine the properties of the human neuromotor system based on investigations of movement, muscle contraction and motor unit activity, freed from any constraints or a priori constructs imposed by linguistic considerations. Any processes or units of neuromotor coding which such enquiry were to establish might or might not subsequently turn out to correlate with linguistic units such as the phone, the feature or the syllable.

Without wishing to deny that the approach of Moll et al. has value, we propose to offer to this symposium the opposite orientation, wherein aspects of the descriptive linguistic apparatus are of prime importance. This decision reflects partly the conviction of a linguistic phonetician that the physical facts of phonetics are at their most interesting when they serve to explain some aspects of phonology, to answer the question why the sound systems of human languages are the way they are. The decision is also derived from the evidence that a wide range of phenomena of coarticulation are not obviously explainable (as yet, at least) in terms of the neuromotor system such as motor unit activity or articulatory velocity and inertia, but are referable to linguistically-defined entities which they thereby can validate.

In addition, many models of the speech production processes seem to occupy the middle ground between these two extreme positions. Among the proposals which might be grouped together here are those of Kozhevnikov and Chistovich (1965), Henke (1966), MacNeilage (1970), Gay (1977a) and Perkell (1977). In very general terms, these models are of a basic "wedding-cake" form such as Figure 1: that is, they are arranged sequentially as tiered boxes, with a distinct top and bottom corresponding to mechanisms associated respectively with more central cortical functions and with more peripheral ones. The number of tiers, and the content of each one, is stylised and is not meant to be attributed specifically

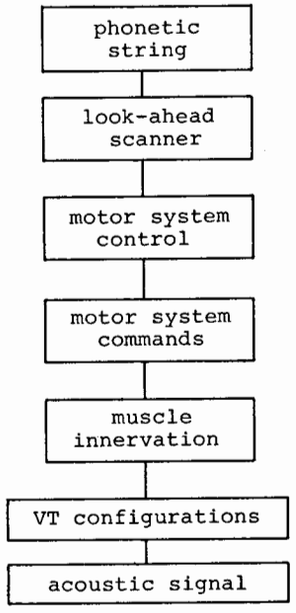

Figure 1

Typical "wedding-cake" structure of some speech production models

to any author. Varying types of feedback between tiers are postulated, but not shown in the figure. A possible objection to the sequential arrangement is that, while no doubt well motivated for the lowest tiers representing the transduction of speech between various fairly accessible transmission channels, such an arrangement is more speculative as a claim about the higher levels of the central nervous system.

The "wedding-cake" models share with our own orientation (and in contrast to Moll et al.) an interest in the linguistic nature of the input, which they normally state to be a string of discrete phonological units.

By contrast with the claims implied by Figure 1, however, our conception of the "upstream" processes feeding data to the motor control system for coarticulation is organised not in a unidirectional, tiered fashion, but as a set of satellites (many representing linguistic factors) linked ambidirectionally to a nuclear body, the CR (Coarticulation Resistance) compiler. The following formulation is a brief summary of this theoretical position. Justification of it, of an organisation which reduces the emphasis on the supposed sequential nature of the speech processors and which minimises the "more central/less central" distinction, and of the particular components recognised in the model, has been given elsewhere (Bladon and Al-Bamerni, 1976; Bladon, 1978).

Information relevant to the direction and domain of coarticulatory effects seems to derive from a wide variety of satellite sources, some quasi-universal (such as the boundary of an intonation-group, which is widely observed to impede the temporal spread of coarticulated features), some language-specific (such as the

report by Ladefoged (1967) that while French and English both show a /k/ coarticulatorily advanced before an /i/ vowel, only French shows the similar effect after /i/), and some speaker-sensitive. Our discussions of these factors have led Kent and Minifie (1977, 120) to write: "Perhaps the solution to coarticulation is as complex as this multiplicity of factors suggests, but ... (they) represent the contributions of many unknown, or poorly known effects". This comment is valid; but it is not a criticism of our position. It seems to be inescapable that the control of coarticulation in speech is indeed governed by a multiplicity of factors.

With a view to integrating these disparate factors in a theory, let us initially postulate the notion of coarticulation resistance (CR) as the central principle of articulatory control. The speech production mechanism is hypothesised to have continuous access to CR information, which can be considered to attach to each allophone and phonetic boundary. It is important to realise why an initial CR specification is tabulated for each allophone. A classic demonstration of this is afforded by the RP English /l/ allophones, of which dark syllabic [l̩] ('fiddle') is highly resistant to coarticulation, dark nonsyllabic [l] ('feel') is somewhat less so, and clear nonsyllabic [l] ('leaf') is very much less resistant. We are aware of no explanation of this behaviour in any terms, linguistical or neurophysiological: the idiosyncrasies can only be handled by assuming something like an allophone-specific assignment of a CR value. This numerical value is re-computed at a level of articulatory planning by the CR compiler to take account of the wide range of relevant coarticulatory constraints.

In what follows we presuppose a phonological apparatus broadly of the kind of Chomsky and Halle (1968). Within their phonological component, various kinds of linguistic construct will be examined, in conjunction with the control mechanisms related to them.

Phonological constructs related to units of articulatory planning

Many early coarticulation studies hoped to identify a single determinant regulating the control of the domain of coarticulatory effects: an invariant unit of production which would have a linguistic counterpart, such as the phone, syllable or phonetic feature. The hope was a vain one. The weight of evidence now available suggests that the coarticulatory control mechanism is sensitive not to any one invariant unit alone but, at different times, to (at least) all those three.

The phonetic feature is at the basis of coarticulation theory in that typical cases of coarticulation arise by definition from the asynchrony of events associated with different articulators. This is reflected phonetically in the temporal spreading of a feature. It is true that the speech control mechanism can be highly sensitive to the feature being coarticulated. Thus, for example, it has been shown that English /s/ occurring in CCC clusters blocks the spread of anticipatory jaw-opening before /æ/ (Amerman, Daniloff and Moll, 1970); and that /s/ resists any shift in its tongue-bladeness (towards a tip articulation) adjacent to /t d n l/ (Bladon and Nolan, 1977); but that this resistance to coarticulate is specific to the coarticulated feature in question, because /s/ freely allows coarticulated labialisation anticipating an /u/ vowel (Daniloff and Moll, 1968). It is equally true that to propose the feature as the sole unit of coarticulatory control would be unattractive, as it would not account for example for British English clear [l], which is quite free in its coarticulation in respect of any of the features vowel-quality, lateral-quality and voicelessness indiscriminately (Bladon and Al-Bamerni, 1976).

Numerous cases, such as the last-mentioned, argue for the phoneme (or perhaps better, the extrinsic allophone) as the unit of articulatory planning. Two further examples may be mentioned. In Italian, intervocalic consonants demonstrate an equal degree of coarticulated tongue-body movement with both a preceding and a following vowel, thus irrespective of syllable boundaries. In French, the anticipatory spread of velum lowering before a nasal, as revealed by EMG, is over a limited domain within a string of preceding oral vowels (Bladon and Carbonaro, 1978; Benguerel et al., 1977). Such arguments for the allophone-sized unit tend, however, to be of a "default" kind, postulated whenever coarticulation fails to coincide with syllable boundaries in some sense. Generalising the allophone, in the interests of proposing an invariant unit, to cases which the syllable could have successfully delimited, has led to an overall too weak hypothesis concerning coarticulatory domain, such as that of Henke's model (1966), which predicted coarticulatory activity whenever a segment showed no antagonistic specification. Our model avoids this problem by two expedients: first, by a segment-specific index of CR (referred to earlier) which inhibits coarticulatory spread in appropriate circumstances, and second, by recognising a plurality of articulatorily-relevant

units which will include the syllable as required.

The phonological syllable, neglected by Chomsky and Halle, has since 1968 enjoyed a revival. Syllable-structure rules in phonology would define the syllable differently for different languages; nevertheless, the structure CV has a claim to universal preference in that, first, there appear to be no languages without CV syllables, second, several languages have syllables of only the CV type, and third, CV is the attested structure in early language acquisition. We profoundly disagree, therefore, with Gay's opinion (1977a) that Kozhevnikov and Chistovich's notion of an articulatory syllable of the form C_oV (where C_o stands for any number of consonants) is "an unnatural and counterintuitive syllable that bears no simple correspondence to common linguistic or phonetic units." Within their articulatory syllable, it will be recalled, coarticulation was hypothesised to be maximal. A great deal of evidence supports this hypothesis, notably the labialization of a string of C before /u/ in Russian (Kozhevnikov and Chistovich, 1965), in English (Daniloff and Moll, 1968) and in Frénch (Benguerel and Cowan, 1974); and also the finding (Bladon, 1977) that even the relatively weaker lip-rounding accompanying English /r/ extended leftwards to the same C_oV boundary.

Other substantive constructs in phonology

Explanation of the control of coarticulatory behaviour in VC positions has remained elusive. Relevant data here include the anticipatory nasalisation of English vowels before nasals (Moll and Daniloff, 1971); American English /r/ which coarticulates with adjacent vowel quality more readily in the final position than in the initial CV position (Lehiste, 1964); or, in VCC sequences, the consonantal influence upon tongue apex position in V (Amerman and Daniloff, 1977). Current phonological theory suggests an explanation in terms of the phonological strength hierarchy. Based on a variety of evidence including sound-change, phonological segments, sequences and positions in the word are assigned a degree of phonological strength. VC positions are weak, since they show more phonological assimilations and elisions. It is reasonable to suggest that the coarticulatory control mechanism is sensitive to this, as to other linguistic properties.

A second such property is the lexical representation of the inventory of phonological items in a language. The degree to which a lateral, for instance, undergoes vowel-quality coarticulation

varies according to the number of laterals in a language's phonological system: in our data, Irish, with three laterals to be kept distinct, shows very little quality coarticulation in comparison with American English, with only one (but highly coarticulated) lateral; Swedish or Italian, with two laterals each, fall in between with respect to coarticulation. The need in such cases to maintain phonemic distinctions (short of the point of incipient sound change and phonemic restructuring) has widely been held to place an upper bound on the extent of coarticulatory behaviour. The principle appeared to make the wrong predictions in the data of Benguerel and Cowan (1974), however, who found that lip protrusion anticipating French /u/ could sometimes extend transconsonantally into the preceding vowel, despite the apparent threat to the lexical contrast /i - y/ in French.

It seems certain that rapid-speech variations are subject to a degree of coarticulatory control. Gay (1977b) showed that at a fast speaking rate a vowel F2 transition effectively begins at a point of greater overlap with the preceding consonant than at a normal speaking rate. Rapid or casual speech variants are coming under scrutiny by phonologists in order to validate their substantive hypotheses of rule ordering. In the derivation of the (ultimate?) rapid-speech form [dəˈvĩ̃ĩ̃] 'divinity', Stampe (1972) demonstrates fairly convincingly that phonological processes do not apply in a linear order, but whenever the configurations they would eliminate arise. Among the processes concerned is the coarticulatory one of intra-syllable vowel nasalization, which reapplies three times, as successively more rapid forms are derived. This cyclic manner of application has important implications for the operational design of the motor control component of the speech production processes, and strongly supports the notion of ambidirectional, on-line exchange of information between the CR compiler (or its equivalent) and the linguistic rule system.

The testing of these various elements of the coarticulatory model by predicting from them onto new data, turns out to be partly successful, but, as has been demonstrated for several cases, partly unsuccessful. Apparently, no one linguistically-related mechanism will explain all or even a majority of observed coarticulatory behaviour. How to assign a weighting to the separate contribution of each mechanism, and indeed how many such mechanisms there are, remain research questions for the future.

References

Amerman, J.D. and R.G. Daniloff (1977): "Aspects of lingual coarticulation", JPh 5, 107-114.

Amerman, J.D., R.G. Daniloff and K. Moll (1970): "Lip and jaw coarticulation for the phoneme /æ/", JSHR 13, 147-161.

Benguerel, A.-P. and H.A. Cowan (1974): "Coarticulation of upper lip protrusion in French", Phonetica 30, 41-55.

Benguerel, A.-P., H. Hirose, M. Sawashima and T. Ushijima (1977): "Velar coarticulation in French", JPh 5, 159-168.

Bladon, R.A.W. (1978): "Some control components of a speech production model", in Current Issues in the Phonetic Sciences.

Bladon, R.A.W. and A. Al-Bamerni (1976): "Coarticulation resistance in English /l/", JPh 4, 137-150.

Bladon, R.A.W. and E. Carbonaro (1978): "Lateral consonants in Italian", Italian Linguistics, forthcoming.

Bladon, R.A.W. and F.J. Nolan (1977): "A videofluorographic investigation of tip and blade alveolars in English", JPh 5, 185-193.

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

Daniloff, R.G. and K. Moll (1968): "Coarticulation of lip-rounding", JSHR 11, 707-721.

Gay, T. (1977a): "Articulatory units: segments or syllables?", Paper read at Symposium, Boulder, Colorado.

Gay, T. (1977b): "Articulatory movements in VCV sequences", JASA 62, 183-193.

Henke, W. (1966): Dynamic articulatory model of speech production using computer simulation, Ph.D. thesis, M.I.T.

Kent, R.D. and F.D. Minifie (1977): "Coarticulation in recent speech production models", JPh 5, 115-133.

Kozhevnikov, V.A. and L.A. Chistovich (1965): Speech: Articulation and Perception, JPRS 30, 543, US Department of Commerce.

Ladefoged, P. (1967): Linguistic Phonetics, UCLA Working Papers in Phonetics 6, Los Angeles: UCLA.

Lehiste, I. (1964): Acoustic characteristics of selected English consonants, Bloomington: Indiana UP.

MacNeilage, P.F. (1970): "Motor control of serial ordering of speech", Psych.Rev. 77, 182-196.

Moll, K.L. and R.G. Daniloff (1971): "Investigation of the timing of velar movements in speech", JASA 50, 678-684.

Moll, K.L., G.N. Zimmermann and A. Smith (1977): "The study of speech production as a human neuromotor system", in Dynamic Aspects of Speech Production, M. Sawashima and F.S. Cooper (eds.), Tokyo.

Perkell, J. (1977): "Articulatory modeling...", in Carre R. et al. Modèles articulatoires et phonétiques, G.A.L.F., 197-192.

Stampe, D. (1972): A dissertation on natural phonology, Ph.D. thesis, U. Chicago.