Forschungsberichte des
Instituts für Phonetik und
Sprachliche Kommunikation
der Universität München
(FIPKM) 35 (1997) 61-75

THE ORIGIN OF COARTICULATION[1]

Barbara Kühnert[1,2] and Francis Nolan[2]

[1]Institut für Phonetik und Sprachliche Kommunikation
Ludwig-Maximilians-Universität München
Schellingstr. 3
D-80799 Munich, Germany
kuehnert@phonetik.uni-muenchen.de

[2]Department of Linguistics
Cambridge University
Sidgwick Ave.
Cambridge, CB39DS, England
fjn1@cus.cam.ac.uk

ABSTRACT

The concept of coarticulation, i.e. the apparent variation of segments due to the influence of adjacent or nearby segments, is central to almost any area in phonetic research. The following text considers the 'origin' of this concept from three different perspectives. In the first section the reasons why coarticulation exists in speech are outlined and the phenomenon and its underlying assumptions are presented in more detail. The second part of the paper deals with the history of the concept. Firstly, an overview of the origin of coarticulation is given from a historical point of view. Secondly, the adoption of the concept of coarticulation as the basis of a major research paradigm in speech production is discussed. The latter includes a summary of the main models and experimental results presented since the late 1960s. In the third section, finally, the ontogenetic origin, i.e. the way in which children acquire coarticulatory behaviour, is considered.

1. WHAT IS COARTICULATION, AND WHY DOES IT EXIST?

The title of this chapter is deliberately ambiguous. 'Origin' refers both to the history of the scientific concept of coarticulation, and to the question of what causes the phenomena in speech which are known as coarticulation. The history of the concept will be dealt with in Section 2, while the reasons why there are phenomena in speech which we can characterize as coarticulation are dealt with explicitly below, as well as implicitly in the discussion of the history of coarticulation in Section 2. There is even a third sense of 'origin' which is dealt with briefly in Section 3, namely the way in which coarticulation develops as a child learns to speak.

Coarticulation, very broadly, refers to the fact that a phonological segment is not realized identically in all environments, but often apparently varies to become more like an adjacent or nearby segment. The English phoneme /k/, for instance, will be articulated further forward on the palate before a front vowel ([k̟iː] *key*) and further back before a back vowel ([k̠ɔː] *caw*); and will have a lip position influenced by the following vowel (in particular, with some rounding before the rounded vowel in [k̠ʷɔː] *caw*). As here, some instances of coarticulation are available to impressionistic observation, and constitute an important part of what has traditionally been thought of as allophonic variation. In many other instances, however, the kind of

[1]To appear in: "*Coarticulation: Theoretical and Empirical Perspectives*", W.J. Hardcastle & N. Hewlett (eds.). Cambridge, CUP.

variation which a segment undergoes only becomes apparent from quantitative instrumental investigation, either of the acoustic signal or of speech production itself.

It is essential to the concept of coarticulation that at some level there be invariant, discrete units underlying the variable and continuous activity of speech production. If this were not the case, and, for instance, the mentally stored representation giving rise to a production of the word *caw* were a fully detailed articulatory plan, then when that word was spoken (in isolation at least) there would be no question of a process of coarticulation - the word would simply correspond to a set of instructions for the time-varying activity of the articulators, and subword segments would not exist in any sense, and could therefore not undergo 'coarticulation'.

There are, however, good reasons to assume that the 'componentiality' which characterizes language (sentences made up of words, words made up of morphemes, and so on) extends down at least to the level of phoneme-sized segments. From the point of view of storing and accessing the mental lexicon, it would be massively less efficient if every entry were represented by its own idiosyncratic articulatory (or indeed auditory) properties, rather than in terms of some kind of phonemic code - a finite set of symbols abstracted from phonetic behaviour. Ironically, studies of coarticulation itself, based on the premise of phoneme-sized segments at some level of representation, lend independent support to the premise. For instance such studies have conspicuously not shown, to take a hypothetical example, that the onset of lip-rounding is consistently different between the words *caw* and *caught,* or that the degree of velar fronting is consistently different in each of the words *key, Keith,* and *keen.* If each word were represented holistically and independently, this result would have to be put down to coincidence. Repeated across the lexicon, the coincidence would be huge and extraordinary. On the other hand this kind of regularity across words is predicted by a view in which these sets of words are represented with common sequences of abstract elements, for instance /k/, /ɔː/ in the case of the first set and /k/, /iː/ in the case of the second, and in which the articulatory realisation of those sequences is governed by regular principles of integration - that is, by principles of coarticulation.

Accepting the role of segments we may then ask a complementary question: why, if a linguistic system operates in terms of discrete and invariant units (let us say phonemes), are these units not realized discretely and invariantly in speech? After all, there is a medium in which this does happen. When we use a typewriter, each letter is realized on the paper separately from the preceding and following ones, and is realized identically (as near as makes no difference) every time it is typed. English orthography has of course strayed somewhat from a phonemic analysis, but all alphabetic writing systems are in essence a way of representing the phonemic code visually. Why does the speech mechanism not behave like an acoustic typewriter?

One reason is perhaps that we do not have a separate vocal tract for each phoneme, in the way that an old-fashioned typewriter has a separate 'hammer' to produce each letter. Instead, a single vocal tract has to alter its shape to satisfy the requirements of all the sounds in a sequence. The vocal tract is governed by the laws of physics and the constraints of physiology, but (also unlike the typewriter) it is producing its communicative artefact in 'real time'. It cannot move instantaneously from one target configuration to the next. Rather than giving one phoneme an invariant articulation, and then performing a separate and time consuming transition to the next, it steers a graceful and rapid course through the sequence. The result of this is coarticulation. It is perhaps rather like a slalom skier, whose 'target' is to be to the left and to the right of successive posts, and who minimally satisfies this target with his skis as they zig-zag down the hill, but whose body pursues a more direct course from the top of the hill to the bottom. In the written medium, it is not typing but handwriting which provides the closer analogy to speech. Examine the occurrences of a given letter in any fluent handwriting, and its realisations will vary. Maybe the tail of the <y> will make a closed loop if a letter follows, but not when it is at the end of a word, and so on. The more fluent the handwriting, the less possible it is to pick out discrete letters, and concomitantly the more each letter's shape will be a product of its environment.

It would be misleading to think of coarticulation in speech as if it were an imperfection in the way language is realized. Speech and language have evolved under the influence of the constraints of the vocal mechanism, and there is no reason to suppose that the relationship between language and the vocal mechanism is not a satisfactory one. The phenomenon of coarticulation may in fact bring advantages beyond the efficient integration of the realisations of successive phonological units.

In particular, the fact that the influence of a segment often extends well beyond its own boundaries means that information about that segment is available to perception longer than would be the case if all cues were confined inside its boundaries. As pointed out by early perceptual theories, the possibility of 'parallel processing' of information for more than one phoneme probably allows speech to be perceived more rapidly than would otherwise be feasible. The possibility that the origin of coarticulation lies not only in the requirements of the articulatory mechanism, but in those of our perceptual system, cannot be discounted.

To recapitulate: the concept of coarticulation entails the hypotheses that at some level speakers make use of a representation in terms of abstract phonological segments, and that there are regular principles governing the articulatory integration of those segments in speech.

Given a coarticulatory standpoint, one way to conceptualize part of the variation in the realisation of /k/ in *caw* and *key* above is to think in terms of the velar stop having a 'target' place of articulation, which is then modified to facilitate the integration of /k/ with the tongue movement for the following vowel segment. From this perspective coarticulation involves a spatial or configurational modification of the affected segment. Alternatively we can break away from thinking in terms of spatial targets for successive segments, and regard coarticulation as the spreading of a property from one segment to a nearby one. For instance if we concentrate on lip activity in the example above, we noted that the lip-rounding always associated with the vowel of *caw* is also present on the consonant preceding it: [k̲ʷɔː]. In *key* there is no lip-rounding on the velar. From this alternative (temporal rather than spatial) perspective, what matters is when articulatory movements begin and end relative to each other. The rounding of [ɔː] has begun during, or even at the start of, the [k][2].

It might appear from this example that the spatial/temporal distinction depends on whether or not a property involved in Coarticulation is crucial to the identity of the affected segment. A velar stop involves a raising of the tongue dorsum, and it is merely the precise location of that raising which is affected by a following [iː]. On the other hand lip activity is not required for a velar stop, and so, in the word *caw,* the lip movement can be anticipated. It is unlikely, however, that a consistent division can be sustained between 'crucial' properties and other properties of a segment. It may be that the absence of lip-rounding on the /k/ of *key* is just as crucial to the perception of this word as the presence of lip-rounding on *caw* (cf. the 'trough' of rounding found on the fricative in /usu/ sequences - see Section 2.2 below, p.xx). So a simplistic linking of spatial coarticulation to crucial properties, and temporal coarticulation to inessential properties, is not valid.

In fact the very distinction between spatial and temporal coarticulation breaks down as soon as we take a more abstract view of articulation. Recent models of speech production (see Section 2.2) hypothesize that each segment is associated with an abstract control structure which is in tune with the mechanical properties of the vocal tract, and which defines that segment in terms of dynamic activity of the articulators. In such a view the distinction between space and time becomes less clear. The control structure for [k] would overlap that for [iː] in time in the phonetic plan of an utterance of *key,* but the competing demands of the two would result in a spatial compromise in the resultant articulation. A current hope, therefore, is that a definition of segments not in terms of superficially observable articulatory movements and positions, but in terms of more abstract articulatory control structures, may lead to a more general and unified description of the variety of coarticulatory phenomena.

This Section has summarized the origin, in the nature of language and its vocal realisation, of the phenomena which are conceived of as coarticulation. We now turn to the origin of the concept of coarticulation in the history of phonetics (Section 2.1), and its widespread adoption as the basis of a research paradigm (Section 2.2).

2. THE HISTORICAL PERSPECTIVE

2.1 The early history

The term 'coarticulation' dates from the 1930s when Menzerath and De Lacerda published *Koartikulation, Lautabgrenzung und Steuerung* (1933). However, the fact that speech sounds influence each other and vary, often substantially, with changes in the adjacent phonetic context had already been known for centuries, while the demonstration that the stream of speech cannot be divided into separate segments corresponding to 'sounds' (or 'letters') coincided with the establishment of experimental phonetics as an independent discipline.

Before experimental techniques were introduced in the study of speech sounds the main tools were 'direct observation' and introspection. Brücke's (1856) and Bell's (1867) insights for German and English, respectively, which laid the foundations for academic phonetics, were based upon such subjective observations. Not surprisingly, early phoneticians shared the assumption that alphabetical letters have corresponding physical realizations in the form of single sounds. The leading idea at the time was that every sound has a static positional (steady state) phase and that different sounds are connected by short transitional glides. The concept of such transitional glides ('Übergangslaute'), which allow the stream of speech to be continuous, was formulated most explicitly by Sievers (1876). For instance, he described the production of a syllable such as *al* in such a way that there exists neither a pure *a*-sound nor a pure *l*-sound during the linking movement of the tongue but a continuous series of transitional sounds which as a whole were referred to as a 'glide' (for an overview of early phonetics, see Tillmann (1994)).

There were, however, some indications in the early literature that the classical view might not capture the whole story. Sievers (1876) himself acknowledged the possibility that, in certain sound combinations, articulators which are not involved in the current sound production might anticipate their upcoming configuration as long as it does not compete with the requirements of the present sound. Examples are the rounding of the lips during the production of /k/ in /ku/ or the preparation of the tongue position for the vowel during the consonant in syllables such as /mi/. And from a more theoretical perspective, Paul (1898:48) wrote: 'A genuine dissection of the word into its elements is not only very difficult, it is almost impossible. The word does not correspond to a sequence of a specific number of independent sounds, each of which could be represented by a sign of the alphabet, but it is, in fact, always a continuous row of an

[2]In this example the influence is mainly of the vowel on the preceding consonant, and we have instances of what has been termed 'anticipatory' or 'right-to-left' or 'backward' coarticulation, as opposed to 'perseverative' or 'left-to-right' of 'forward' or 'carryover' coarticulation (when an earlier segement influences a later one). Both the direction an extent of coarticulatory effects have been regarded as crucial in testing hypotheses about coarticulation, as will emerge in Sections 2 and 3.

infinite number of sounds (...)'.[3]

Historically, phonetics moved towards experimental research during the last quarter of the 19th century. 'Kymography' allowed the mechanical recording of time varying signals, including the acoustic signal, air flow, and (via rubber pressure-sensitive bulbs in the mouth) tongue movements (for detailed descriptions of a number of such devices, see Scripture, 1902). Initially, instruments were introduced into the discipline in order to find objective measurements of what had been more or less taken for granted: individual speech events. As soon as technical means had been adopted for investigating spoken language, however, the premonitions just mentioned were confirmed and attempts to impose strict linguistic boundaries on articulatory events or the speech signal proved problematic.

Rousselot (1897-1901) was the first who made use of the newly available apparatus for the study of speech. On the one hand, Rousselot still shared the opinion that the obtained speech curves could, and should, be divided into separate sounds. As Tillmann (1994) points out, this assumption led in some instances to a rather questionable methodology. In most recording sessions, more than one speech curve was recorded, such as tracings of the lips, and the nasal and oral air flow. If segment boundaries could not be clearly located by comparing the different curves it was assumed that, in fact, a 'recording error' had occurred and the kymograms were then replaced. On the other hand, Rousselot clearly recognized that at a given point in time there is a superposition of several influences on the movements of the articulators, stressing that the study of sounds in isolation should be replaced by the study of sounds in context.

Rousselot's vast collection of data is full of various examples of phenomena which were later studied extensively within the coarticulation paradigm of phonetic research. For instance, in his analysis of consonant clusters, he observed that if two consonants are produced by two independent articulators, as in /pla/ or /fla/, they are prepared together and the movements uttering them may be simultaneous - a notion confirmed by Stetson (1951). In a related sense, he reported that in some CV sequences the tongue might take the appropriate position for the vowel at the beginning of the preceding consonant. Traces of the lips and the tongue during /ba/ and /bi/ or /za/ and /zi/ productions showed that the tongue lies much higher during the articulation of the consonant when preceding the vowel /i/ than when preceding /a/. On the basis of such evidence Menzerath and De Lacerda (1933) formulated their principle of 'Koartikulation' some 30 years later. Furthermore, Rousselot (1897-1901:947) pointed out that the lip rounding for the vowel /u/ in a VCV sequence may already start during the first vowel, noting that "the second of two vowels is the most influential one, as can be seen in *a tu* in the utterance *il a tourné.* At the middle of the *a* the lips start to raise due to the requirements for the production of the *u*".[4]

Scripture (1902) gives a broad and systematic survey of the experimental research carried out around the turn of the century. In many instances results like the ones outlined above are described, i.e., observations of parallel articulatory activity and modifications of one sound by another. The most noteworthy result for the concept of coarticulation comes from a study by Lacotte (in Scripture 1902:372; for a history of early investigations in lingual coarticulation, see Hardcastle, 1981). Lacotte found that a vowel may influence not only the preceding consonant but also the vowel before the consonant. Records of /eli/ and /ela/ or /ebi/ and /eba/ showed that the articulatory setting for /e/ was different according to the second vowel in the sequence. The tongue rose higher and nearer to the /i/ in /eli/ or /ebi/ than in tokens in which the last sound constituted an /a/. Essentially, this anticipates the form of vowel-to-vowel coarticulation which was postulated by Öhman (1966) and which became influential in the development of some coarticulation theories.

On the basis of this and other studies presented in his book Scripture (1902) presumed that the character of any movement depends on other movements occurring at the same time and that in speech there are hardly any static postures of the articulators. "The tongue is never still and never occupies exactly the same position for any period of time" (p.325). Thus, for the first time the understanding of speech as a sequence of single elements linked by fast transitions was rejected. "Speech cannot be considered as made up of separate elements placed side by side like letters. In the flow of speech it is just as arbitrary a matter to consider certain portions to be separate sounds as to mark off by a line where a hill begins and a plain ends" (p.446). On the same grounds, Scripture rejected the division of speech into any higher units, such as the syllable, which just would correspond to divisions of a landscape into hill-blocks and valley-blocks.

Thus, although early experimental phoneticians did not have any coherent concept of coarticulation, they were well aware of its existence. The use of the terminology in the field, however, was still rather vague. The terms 'assimilation' and/or 'adaptation' have been loosely applied to all sorts of phenomena, ranging from the treatment of historical sound changes to the articulatory mechanisms described previously.

Jones (1932) introduced a distinction between 'similitude' and 'assimilation', depending on whether two sounds influence

[3]"Eine wirkliche Zerlegung des Wortes in seine Elemente ist nicht bloss sehr schwierig, sie ist geradezu unmöglich. Das Wort ist nicht eine Aneinandersetzung einer bestimmten Anzahl selbständiger Laute, von denen jeder durch ein Zeichen des Alphabetes ausgedrückt werden könnte, sondern es ist im Grunde immer eine kontinuierliche Reihe von unendlich vielen Lauten (...)".

[4]"Entre les deux voyelles, la plus influente est la seconde, comme cela se montre très bien pour *a tu* dans *il a tourné*. Dès le milieu de l'*a*, la ligne des lèvres s'élève à la sollicitation de l'*u*".

each other in such a way that they become more alike but do not change their phonemic identity (similitude - as for instance when the normally voiced /l/ phoneme of English is partially devoiced in [pl̥iːz] *please*), or whether they influence each other in such a way that one is replaced by another phoneme (assimilation - as when /s/ changes to /ʃ/ in *horse-shoe*). Jones' concept of similitude therefore covers only some cases which previously had been subsumed under the notion of assimilation, but does include the major part of what was later studied under the heading of coarticulation. The formulation of similitude was, though, still rather static in nature, and seems not far removed from Wickelgren's theory of context-sensitive allophones (1969; see p.xx) which implies that numerous fixed variants of a phoneme are available to the speaker for use when required by the context. The distinction between similitude and assimilation is also reminiscent of Keating's (1990) distinction between phonological and phonetic coarticulation (see p.xx).

The term 'coarticulation' as such, and perhaps the most explicit rejection of the old view of positional sounds, was put forward by Menzerath and De Lacerda (1933). By means of kymograms including air flow measurements they investigated the production of a series of German labial consonant and vowel sequences. Failing to find stable articulatory positions, Menzerath and De Lacerda proposed that articulation is governed by two major principles, 'Koartikulation' and 'Steuerung'.

'Koartikulation' (or 'Synkinese') indicates that articulators already prepare for following sounds during the production of a preceding segment. Moreover, as proposed afterwards, it was hypothesized that this preparatory activity begins as early as possible. The evidence for the concept of coarticulation was primarily based upon the observation that the articulatory movements for the vowel in tokens such as /ma/ or /pu/ began at the same time as the movements for the initial consonant. Thus, although 'Koartikulation' was postulated as a general organisational principle of articulatory control, its experimental validation was, strictly speaking, restricted to anticipatory labial interactions between consonants and vowels in syllable-initial position.

'Steuerung' (steering, control), on the other hand, is somewhat more difficult to apply to today's understanding of coarticulation. The process was essentially limited to immediately adjacent sounds which involve the same articulator. For instance, the lip movements of vowel articulations were said to be completely controlled ('gesteuert') by the following consonant in syllables such as /am/ or /ap/. The final consonant indicates the direction of the articulatory movement during the vocalic portion. The point in time in which the 'Steuerung' starts was supposed to be dependent upon other sounds which, in turn, precede the vowel during longer utterances. In a more general sense, then, 'Steuerung' seems to apply to the deviation of an articulator in its articulatory target in one sound due to the presence of a different target in another sound.

Thus, for Menzerath and De Lacerda the structure of any utterance is a complex interweaving of simultaneous movements, i.e., 'all articulation is coarticulation'. Somewhat paradoxically, however, they summarized that the complicated combination of articulatory action exists for the sole purpose of producing speech sounds which are acoustically (and perceptually) distinct units. The articulation associated with certain sounds is not separable,[5] but the sounds themselves are.[6] In other words, the problem of segmentation was pushed from one level of speech to another (see Tillmann and Mansell (1980) for discussion).

In the same way as Menzerath and de Lacerda can be regarded as having introduced the concept of coarticulation in phonetic research, Stetson's (1951) approach to the investigation of speech can be regarded as laying the theoretical basis for the redefined notion of coproduction (see below p.xx). He was the first to apply methods and concepts from the study of skilled movement to the analysis of speech production, in work which later gained new significance in the context of Action Theory (Fowler et al., 1980) and, relatedly, Articulatory Phonology (Browman and Goldstein, 1986, 1989) and Task Dynamics (Saltzman and Munhall, 1989).

Stetson argued that it is the coordination of the articulatory movements involved in speech production which is of major importance, and this coordination is best studied by looking at different phonetic modifications, specifically those caused by changes of rate and of stress. On these grounds, he defined an allophone, rather unusually, as the variation "of the phoneme pattern due to rate vs. stress" (p.193). More generally, and prefiguring the concept of coproduction, he stated "All generalizations on the reciprocal influence of sounds, like assimilation or prevision or law of economy, must finally be referred to the skilled movements involved" (p.122).

Considering the syllable to be the fundamental unit of speech production, Stetson investigated the coordination of speech sounds primarily with respect to their function within syllable groups. A particular concern was the role and interaction of syllable final ('arresting') and syllable initial ('releasing') consonants. His data showed that if a syllable is constantly repeated with a gradually increasing speaking rate, such as /tas tas tas/, its structure will change abruptly at one moment in time, i.e. resulting in /sta sta sta/. Similar jumps in syllable composition were later taken to be a major source of evidence for the existence of coordinative structures in speech (compare Kelso et al., 1986).

Moreover, Stetson observed that articulatory movements can be present although acoustically they might be covered completely. For example, in slow productions of the nonsense syllable /ispda/ the three intervocalic consonants were kept

[5] "Artikulatorisch-konstante Laute gibt es nicht" (Menzerath and De Lacerda 1933:61).

[6] "Die Sprachlaute sind trennbar" (Menzerath and De Lacerda 1933:60).

distinct, articulatorily as well as acoustically. However, in fast productions of the same syllable the two tongue tip movements merged into one, and the lip movement for /p/, although still clearly visible in the kymogram, was overlapped in time by the integrated tongue tip movement for the alveolars, and according to Stetson there was no closure phase for the bilabial evident in the acoustic signal. The manifestation of hidden gestures of this kind was to be of central importance in Browman and Goldstein's Articulatory Phonology (1990), in particular for the discussion about the relationship between coarticulation / coproduction and casual assimilations.

2.2 The coarticulatory research paradigm: models and experiments

Since the late 1960s the experimental investigation of coarticulation has developed into a major area of research. With the detailed investigation of coarticulatory effects it became apparent that the influential interval of a speech sound varies considerably, potentially extending quite far. In particular, studies of anticipatory labial coarticulation demonstrated that the articulatory rounding gesture, associated with the vowel /u/, is initiated up to as many as four to six segments before the actual target vowel, regardless of syllable or word boundaries (Benguerel and Cowan, 1974; Sussman and Westbury, 1981). Coarticulation, therefore, appeared to be more than the pure consequence of physiological limitations and inertial effects of the speech organs, as had been proposed in some previous accounts[7].

For example, the first version of Lindblom's target undershoot model (1963) posited that a string of phonetic segments is realized by a string of commands which in running speech are issued at very short temporal intervals. As the articulators are restricted in the speed with which they can move from one articulatory target to the other, they may not always complete a given response before the command for the next articulatory configuration arrives. In other words, the articulators fail to reach ('undershoot') their targets and respond to more than one input signal simultaneously.

This approach, though, cannot fully account for more extended temporal influences. Much of the phenomenon of coarticulation, therefore, has been regarded as a more active process which occurs at some higher level. For that reason, the key attraction of coarticulation research lay in the hope that by revealing the systematic coarticulatory patterns it would be possible to discover the (universal) underlying units of the speech production process as well as the mechanisms linking them and, in this way, to clarify the mismatch between mental and physical events. Almost every articulatory subsystem was investigated in a variety of studies using a number of different techniques. However, given the large amount of variation observed, it was difficult to impose an unambiguous pattern, and an increasing number of models were developed.

This section will summarize some of the major models, but the following is neither a complete survey nor a thoroughgoing critique. Critical reviews of models of coarticulation are available in Kent and Minifie (1977), Sharf and Ohde (1981, who include an excellent comparison between data and models available prior to 1981), Fowler (1980, 1985), Kent (1983), and Farnetani (1990). Detailed experimental results referring to the different articulatory subsystems can be found in the following chapters, and the most recent theories are discussed in more detail in Chapter 2.

The publication which most clearly heralded the growing interest in coarticulation was that of Kozhevnikov and Chistovich (1965), whose major concern was the development of a general model of articulatory timing. The largest unit of speech production was assumed to be the syntagma and, within the syntagma, the basic articulatory input unit was assumed to be the articulatory syllable, which consisted of a vowel and any number of consecutive consonants that preceded it.

The concept of the articulatory syllable was derived from recordings of lip movements in Russian speakers which revealed that the lip protrusion for the vowel /u/ began simultaneously with the first consonant when either one or two consonants preceded the vowel. The same result emerged when a conventional syllable or word boundary fell between the two consonants. Thus, the temporal extent of anticipatory coarticulation was taken as an indication of the size of the articulatory syllable. Provided that the consonants did not involve contradictory movements, all the articulatory actions connected with one articulatory syllable were supposed to start at its beginning. As a corollary, this view implies that, when possible, a vowel is initiated as soon as a preceding vowel is terminated.

Subsequent studies showed, as in fact Rousselot (1897-1901; see Section 2.1 above, p.xx) before them, that coarticulatory effects occur not only within such restricted forms. In his classic spectrographic study on Swedish, Öhman (1966) showed that two vowels interact with each other across intervening stops. Vowel-to-vowel coarticulation has since been reported in both acoustic (Fowler 1981; Recasens, 1989) and articulatory studies (Carney and Moll, 1971; Kent & Moll, 1972; Kiritani et al., 1977).

Öhman developed a VCV model of coarticulation which shares with Kozhevnikov and Chistovich (1965) the assumption that vowels play some special role in the speech production process. In Öhman's account, the vowel and consonant articulations are considered to be largely independent of each other at the level of neural instructions. Vowels are produced by relatively slow diphthongal movements of the tongue body on which the articulatory gestures for the consonants are superimposed. Vowels therefore coarticulate with each other across the medial

[7] Mechano-inertial limitations certainly play some role, especially for some part of carry-over influences (see Daniloff and Hammarberg, 1973, and Fowler, 1980, for discussion).

consonant(s), while consonant-vowel effects depend on the degree of tongue involvement in the production of the consonant(s) (but see Gay, 1977, for some complicating evidence). The idea that vowels and consonants have several quite distinct properties was adapted by Perkell (1969) and, later, implemented in a larger theoretical framework by Fowler (1980, 1981).

Wickelgren (1969, 1972) put forward an alternative theory which moved coarticulation away from the speech apparatus up to a more central level of speech motor control. According to this account, language users do not have a set of phoneme-like segments as the immediate phonetic constituents of words, but an inventory of context-sensitive allophones. These stored allophones are understood to be versions of segments which are specifically defined according to the potential left and right context in which they are able to occur, i.e. each allophone is essentially an ordered triple of immediately adjacent phonemes in the phonemic spelling of the word" (1972:239). For the string *pin*, for example, the context-sensitive elements would be: ${}_{\#}p_{i}$, ${}_{p}i_{n}$, ${}_{i}n_{\#}$ (# indicates a word boundary).

Crucially, however, the theory has difficulty with coarticulatory influences that extend beyond adjacent sounds, and with effects caused by prosodic properties. Although it might be possible to cope with such problems in part by specifying all additional contexts in which a phoneme could undergo changes, this procedure would increase the number of allophones exponentially. Most criticisms of Wickelgren's theory, therefore, revolved around the high number of elements that would need to be stored (MacNeilage, 1970; Halwes and Jenkins, 1971; Kent, 1983). The approach also misses generalizations, such as the nasalization of all English vowels preceding a nasal consonant (Kent and Minifie, 1977), and at the same time is too rich in information (Fowler, 1985). In the present example it is just the nasality which is shared by the vowel but not, for instance, the place of articulation. Even more problematically, different sounds are influenced by different properties of their neighbours; oral consonants preceding a nasal do not become nasalized. Wickelgren's allophone lists conflict with fundamental assumptions about the speech production process, namely that the integration of segments is rule-governed and productive. In a sense, the classical conception of coarticulation disappears in Wickelgren's approach.

Another view which focuses on segment-specific properties is Bladon and Al-Bamerni's (1976) concept of 'coarticulation resistance' (CR). Their spectrographic investigation of the extrinsic allophones of /l/ in British received pronunciation showed that the influence of adjacent vowels on F2, and the influence of adjacent voiceless plosives on the amount of voicelessness in the laterals, decreased from clear [l] to dark [ɫ] to dark syllabic [ɫ̩]. They hypothesized that each extrinsic allophone is stored with a different numerical value of CR to which the speech production mechanism has access. Anticipatory and carry-over coarticulation is allowed to spread freely until it is inhibited by a high CR specification on some segment. The CR value was supposed to be determined by a variety of factors - universal, language-specific, context-sensitive, and, possibly, speaker-specific.

Further research indeed suggested that coarticulation phenomena seem to be sensitive to all these factors. However, to trace them back to a different underlying CR value might not constitute a very compelling phonetic explanation, as Kent and Minifie (1977) remark: "CR values seem to be little more than summary numbers that represent the contributions of many unknown, or poorly known, effects. The CR values do not themselves generate predictions" (p.120). Moreover, it is also not clear whether a single CR value suffices to describe the overall behaviour of a segment. For example, in the utterance *feel things*, the tongue body configuration of [ɫ] would presumably be resistant to adjacent coarticulatory forces (see above), but the tip free to adopt the following dental place of articulation. Perhaps a segment would end up needing a CR value for each phonetic dimension (see Nolan, 1983).

'Coarticulation resistance' is therefore now used in a more general sense to refer to the widespread observation that coarticulation is gradual and varies between segments. For example, Bladon and Nolan (1977) reported from a video-fluorographic study of English alveolars that the tongue tip position increased in variability in the order /t,d,n,l/ when the sounds occurred next to /s/ or /z/, while the fricatives themselves were not affected. This overall pattern of contextual sensitivity appears to be a rather robust effect since it has been documented in other languages, such as Italian (Farnetani, 1990), Swedish (Engstrand, 1989) and German (Kühnert et al., 1991). A completely satisfactory account of the order of alveolar variation, however, is yet to be elaborated. One plausible suggestion adduces acoustic distinctiveness, i.e. /l/ and /n/ can be produced with more variability without losing essential acoustic properties, but this hypothesis has never been strictly tested (see Hoole, 1996).

Recasens' articulatory model of lingual coarticulation (1984, 1987, 1989), which is related to Öhman's model (1966), connects the notion of coarticulatory resistance with a sound's degree of tongue dorsum elevation. On the basis of electropalatographic and acoustic studies of Catalan and Spanish VCV and VCVCV sequences, Recasens posits that the temporal extent and the strength of vowel-to-vowel coarticulation varies inversely with the degree of tongue dorsum raising required for the intervening consonants, i.e. the larger the contact between the tongue and the palate, the less the coarticulatory modifications. The production of the alveo-palatal nasal /ɲ/, for example, was shown to reduce vowel interactions more than a medial alveolar /n/ (Recasens, 1984).

The approach further suggests that sounds which block coarticulation are the least affected by other sounds while, at the same time, affecting other sounds the most (see Farnetani, 1990). For instance, Kiritani and Sawashima (1986) reported from an X-ray microbeam study that the influence of the second vowel appears later in the first vowel of a VCV token when that is a high front /i/ than when it is an /a/. Additionally, /i/ left more prominent traces on the medial consonant than /a/. Similar coarticulatory hierarchies between vowels were observed by Butcher and Weiher (1976) and Hoole et al. (1990) for German (but note Farnetani et al., 1985, for some ambiguous evidence).

A different line of thought with a long history refers to distinctive features (e.g. Jakobson et al., 1952). Feature-based models (Moll and Daniloff, 1971; Daniloff and Hammarberg, 1973; Benguerel and Cowan, 1974)[8] were initiated by the articulatory computer model developed by Henke (1966). This model starts with a segmental input in which each segment is connected to a set of phonetic features which, in turn, are transferred together to some articulatory (phonemic) goal. Only contrasting properties are specified at the level of phonological input and irrelevant value entries are left empty. For example, in English, in which nasalization is not contrastive for vowels, no vowel is assumed to have a specification for the feature [+/- nasal]. A forward-looking device below the phonological component is then assumed to scan the features of future segments and to copy them onto more immediate segments as long as their goals are not incompatible. Thus, "forward effects are due entirely to a higher level look ahead or anticipation" (Henke, 1966:47).

Support for the feature-spreading account came primarily from studies showing that anticipatory labial coarticulation started as early as the first consonant preceding a rounded vowel (see Benguerel and Cowan, 1974, for partial evidence; Lubker, 1981), and showing that velopharyngeal opening was anticipated in a string of vowels before a nasal consonant (Moll and Daniloff, 1971). However, the same subsystems were soon cited as yielding counter-evidence, leading to an ongoing discussion about apparently conflicting results (Gay, 1978; Bell-Berti and Harris, 1979, 1981, 1982; Sussmann and Westbury, 1981; Lubker, 1981; Lubker and Gay, 1982, among others).

One controversy concerned the occurrence of 'troughs'. For instance, labial movement or EMG recordings of sequences such as /usu/ show a diminution of rounding during the consonant (Gay, 1978; Engstrand, 1980; Perkell, 1986). Such data are a problem for 'look-ahead' models since there is no reason to assume a rounding specification for a non-labial consonant and hence the rounding should spread over an entire non-labial string. Explanations for 'troughs' varied considerably. While Gay (1978) interpreted his results to be in agreement with a syllable-sized production unit (Kozhevnikov and Chistovich, 1965), Engstrand (1980) suggested that acoustic and/or aerodynamic constraints on the intervocalic consonants contribute to troughs. In particular a rounding gesture might interfere with the optimal acoustic conditions for /s/, and therefore the lips exhibit some active retraction. Perkell's data (1986) pointed to a complex of factors which may underlie the reduction of rounding and were only partially consistent with Engstrand's proposals, indicating that its relative importance may vary for different subjects and different languages.

Perkell's study also suggested that alveolar consonants, previously assumed to be neutral with respect to lip rounding, may in fact have their own specific protrusion gesture. More explicit evidence came from an EMG and optoelectrical investigation by Gelfer et al. (1989) who examined the inherent lip activity for /s/ and /t/. Even in /isi/ and /iti/ some consonant related lip rounding could be registered. As Gelfer et al. point out, some of the (apparently) conflicting results concerning labial coarticulation might stem from disregarding such subtle alveolar properties (see Bell-Berti and Krakow (1991) for similar arguments concerning vowel production and velum height). Importantly, such results illustrate that any feature-spreading account depends on the *a priori* assumptions made about the feature specifications. Moreover the absence of a feature specification for a sound at an abstract level cannot be taken to imply that the relevant articulator is wholly neutral in the production of the sound. This makes the experimental investigation of coarticulatory effects a delicate matter. (Other general points of discussion of early feature-coding models have been discussed in Kent and Minifie, 1977, Sharf and Ohde, 1981, or Fowler, 1980, 1985).

Recently, Keating (1988; 1990) has proposed a refined approach to coarticulation which preserves the conception of a segmental feature input. For Keating, coarticulation can either be phonological or phonetic in character. Coarticulation at the phonological level is caused by rules of feature spreading along the lines discussed in recent developments in nonlinear phonology (Goldsmith, 1990). Phonological assimilation should result in static effects, where successive segments share a certain attribute fully.

On the other hand, phonetic coarticulation effects are typically more gradual in time and space and may affect portions of segments to varying degrees. Keating proposes that phonetic realisation involves a conversion of features or feature combinations into spatio-temporal targets. These targets, however, are not understood as fixed configurations (as in the early formulation of an invariant articulatory target model by MacNeilage, 1970). Rather, for a given articulatory or acoustic dimension every feature of a segment is associated with a range of values, called a 'window', which represents its overall contextual variability. The wider the window, the greater the permitted coarticulatory variation. Phonetic coarticulation is the result of finding the most efficient pathway through the windows of successive segments. In fact the window model of coarticulation is reminiscent of the original concept of coarticulation resistance (Bladon and Al-Bamerni, 1976, and above), and of Nolan's (1983:119-20) quantitative formalization of it, in which the notional 'costs' of alternative transitions between segments are calculated on the basis their CR values. A window is more or less the spatial expression of the degree of coarticulatory resistance of a particular phonetic dimension.

[8] Early feature-coding proposals differ somewhat in the details of their implementation and the way they treat carry-over coarticulation; however, they all share the main mechanisms which are exemplified in Henke's approach (1966).

Keating's model, especially the details of phonetic implementation patterns, still has to be evaluated in many respects, but, in principle, her concept can handle many coarticulation data more adequately than early feature-spreading approaches which were not able to account for any finer, intersegmental variations of articulatory accommodation. For example, in Keating's account individual English vowels are associated with different windows for velum height, which are relatively wide, but not as wide as possible since the velum is not allowed to lower maximally. In contrast, nasal and oral consonants, which possess either low or high velum positions, are associated with narrow windows. The contours through assumed windows of this kind, then, are compatible with the observation that the velum rises slightly for a vowel between two nasals (Kent et al., 1974), and similar assumptions about the allowed range of lip protrusion of vowels and alveolar consonants could account for the occurrence of troughs in /usu/ or /utu/ tokens.

Like its predecessors, however, Keating's account still embodies the fundamental assumption that there exists some process in which abstract (i.e. nonphysical, discrete and timeless) units are converted into the complex, continuous and time-bound realizations evident in the speech event (see Boyce et al., 1991; Fowler, 1990, 1992). This disparity between phonological descriptions and actual utterances, and hence the notion of coarticulation, is challenged by proponents of 'coproduction' theories (Fowler, 1980, 1992; Fowler et al., 1980; Fowler and Saltzman, 1993; Bell-Berti and Harris 1981, 1982; Kelso et al., 1986; Browman and Goldstein 1986, 1989; Saltzman and Munhall, 1989). There has been a long debate in the literature which focused on the philosophical differences between what have been called 'translation' or 'extrinsic timing' theories, and 'intrinsic timing' or 'coproduction' theories, and the arguments will not be repeated here (see Hammarberg 1976, 1982; Fowler 1980, 1983; Fowler et al., 1980; Nolan 1982; Parker and Walsh, 1985, among others).

Broadly speaking, in coproduction accounts the hypothesis of underlying invariance and the reality of surface variability are reconciled by redefining the primitives of the speech production process as dynamically specified units. Based on the concept of coordinative structures (see Fowler et al., 1980; Kelso et al., 1986) these underlying units are supposed to be functionally defined control structures, called 'gestures', which represent and generate particular speech-relevant goals and implicitly contain information on articulatory movement in space over time. Such gestures are, for example, the formation of a bilabial closure, or an alveolar near-closing gesture which permits frication for the production of alveolar fricatives. In the case of labial closure, for instance, the jaw, upper lip, and lower lip movements are constrained by the coordinative structure to achieve closure regardless of the phonetic context. Crucially this view suggests that, during speech production, a gesture does not change in its essential properties, but rather that it is its temporal overlap with other gestures which results in the variability observable in the vocal tract activity associated with that gesture. Thus coproduction implies that at a given point in time the influences of gestures associated with several adjacent or near-adjacent segments show their traces in the articulatory and acoustic continua. For instance, in a bilabial consonant and vowel sequence the formation of the closure will be influenced by the demands of the following vowel since both gestures share the jaw as an articulating component (for details of the transformation and the blending procedures of abstract gestural primitives into vocal tract actions, see Saltzman and Munhall, 1989, and Fowler and Saltzman, 1993). In this framework, therefore, coarticulation is rather straightforwardly the consequence of the inherent kinematic properties of the production mechanism and is considered to be merely a general co-operative process of movement coordination.

Besides the conflicting results which were reported in the literature (see references above) a major point of criticism raised against coproduction models is that they lean too heavily on the inherent kinematic properties of the production process, thus neglecting the importance of acoustic salience and perceptual distinctiveness. Lindblom (1990), for example, stresses that speech production is always a 'tug-of-war' between production-oriented factors on the one hand and output-oriented constraints on the other. Production constraints reflect the motor system's tendency towards 'low-cost' behaviour, and output constraints reflect the speaker's response to the requirements of the communicative situation. In addition, doubts have been raised over the role of coordinative structures in speech production, and the assumption that the input to the physical speech mechanism is a contrastive phonological representation (see, for example, the discussions in Nolan, 1982; Kent, 1983, 1986; Shaffer, 1984). Along these lines, Holst and Nolan (1995) and Nolan, Holst and Kühnert (1996) have claimed to show that although coproduction is capable of accounting for many of the observed forms when [s] accommodates to a following [ʃ], there are some forms which can only be accounted for if a process of phonological assimilation has taken place in the representation that is input to the speech mechanism. This view, in its separation of a more phonological and a more phonetic source of observed coarticulatory effects, is similar to the ideas of Keating discussed above.

Consistent with the coproduction view is Bell-Berti and Harris' (1979, 1981, 1982) time-locked or frame model of coarticulation. The model asserts that the component gestures of a phonetic segment begin at relatively time-invariant intervals before the phonetic target itself is achieved. Anticipatory coarticulation is therefore temporally limited and does not extend very far backward in time before the gesture becomes acoustically dominant. Thus, in contrast to feature-spreading approaches, the length of the preceding string of phones and their conflicting or non-conflicting featural specifications are irrelevant (apart from some cases in which an intervocalic consonant duration is very short). The findings in favour of this concept were again primarily derived from investigation of anticipatory labial and velar coarticulation which either showed that the EMG activity of the orbicularis oris associated with lip-rounding was initiated at a fixed interval before the onset of a rounded vowel (Bell-Berti and Harris 1979, 1982), or that the lowering of the velum started invariantly relative to the onset of nasal murmur in a nasal consonant production (Bell-Berti and Harris, 1981).

The 'hybrid model' of coarticulation, which was put forward by Perkell and Chiang (1986) using their data on lip rounding and some preliminary observations on the temporal patterning of velum lowering by Bladon and Al-Bamerni (1982), constitutes a compromise between the assumptions made by the featural look-ahead mechanism and the time-locked theory. According to this model, the anticipation of a gesture is characterized by two components. For example, in the case of anticipatory lip rounding in English there is first an initial slow phase of rounding which starts as soon as possible, specifically at the acoustic offset of a preceding unrounded vowel. Secondly, there is a more prominent rapid phase (identified by an acceleration maximum in the rounding gesture) which is supposed to be time-locked to the acoustic onset of the rounded vowel. Thus, as the intervocalic consonant string increases, the duration of the first phase increases with it while the duration of the second phase remains constant.

However, in a more recent investigation of the timing of upper lip protrusion Perkell and Matthies (1992) concluded - following the work outlined above by Gelfer et al. (1989) - that at least some protrusion effects are probably consonant-specific, and, therefore, qualified some of the original evidence which was taken as support for the hybrid model. The outcome of the study nevertheless did not support a purely time-locked coarticulation pattern, but rather suggested that coarticulatory strategies are based on competing constraints which are both kinematic and acoustic in nature, and that the balance between the constraints may vary from speaker to speaker. This raises a more general question: how does variation in coarticulation between speakers arise?

3. ACQUISITION AND VARIATION

In recent years attention has increasingly been paid to children's acquisition of coarticulatory behaviour. It is thus now possible to say something about the ontogenetic origin of coarticulation in individuals, although a far from clear general picture emerges. It may be that coarticulatory strategies are essentially idiosyncratic, with each individual free to develop a personal solution to the integration of successive segments. These alternative solutions would be expected to persist into adulthood, and evidence of coarticulatory variation in adults is briefly reviewed in the later part of this Section.

It is clear from a number of acoustic investigations (e.g., Kent and Forner, 1980) and articulatory measurements (e.g., Sharkey and Folkins, 1985) that children are much more variable in their phonetic patterns than adults. With respect to the effect of coarticulation on the variability in children's utterances two differing positions can be broadly distinguished. One account holds that children show the tendency to produce speech rather more segmentally than adults (Kent, 1983; Katz et al., 1991). This is thought to reflect an acquisition process in which the motor skill of temporal sound sequencing is acquired first, while the finer details of the temporal coordination of the articulators develop later. As a corollary, coarticulation is likely to be less prominent for young children. In contrast, an alternative approach to speech development suggests that children's productions might be characterized by more, rather than less, coarticulation (Nittrouer et al., 1989; Nittrouer and Whalen, 1989). By this view, which is consistent with the gestural approach of Browman and Goldstein (1986), children are assumed to rely to a larger extent on syllable-based speech production units and only gradually to narrow their minimal domain of articulatory organization. Thus, the spatiotemporal overlap of gestures is more prominent at early ages and diminishes through a process of differentiation.

Studies investigating the extent and degree of coarticulation in child productions have thus far yielded inconsistent results which appear to depend crucially upon the articulatory subsystem under consideration. The most divided picture arises with respect to anticipatory lingual coarticulation. For example, Nittrouer et al. (1989) presented evidence from a fricative study with eight adults and eight children at each of the ages 3, 4, 5 and 7 years that young children organize their speech over a wider temporal domain. F2 estimates and centroid frequency values showed a gradual, age-related decline of the influence of the vowels /i/ and /u/ on the preceding /s/ or /ʃ/. In an attempt to replicate this outcome with ten adults and ten 3-, 5- and 8-year olds, however, no age-dependent differences could be detected by Katz et al. (1991). On the other hand greater coarticulation for adults was observed by Kent (1983), who looked at the influence of a following consonant on a preceding vowel. Finally, Sereno and Lieberman (1987), looking at the influence on a velar stop by a subsequent vowel in /ki/ and /ka/ syllables of five adults and 14 children between the ages of 3 and 7, found consistent coarticulatory effects for adults in the form of different predominant spectral peaks in the consonant. However, their measurements varied greatly between individual children, with some of them displaying adult-like patterns while others did not show any traces of lingual coarticulation. Significantly, the differences among the child speakers did not correlate with age.

More agreement can be found in the literature with regard to anticipatory lip-rounding before /u/. The studies by Nittrouer et al. (1989) and Katz et al. (1991), as well as an investigation by Sereno et al. (1987), all indicate that labial Coarticulation in the speech of English children is roughly similar to that of adult subjects (but see Repp, 1986). Conversely, for Swedish, which has a more complex lip-rounding contrast than English, Abelin et al. (1980) reported that adults adopt a look-ahead strategy while the children's labial coarticulation appeared to be time-locked, i.e., the temporal extent of anticipation became more prominent with age. This difference between the languages is compatible with the more general observation that labial coarticulation is highly influenced by language-specific factors (Lubker and Gay, 1982). Thus, in Swedish, in which labial anticipation is constrained by several linguistic and, hence, perceptual factors, the details of possible coarticulation are refined during maturation, while this is not necessary for English.

However, studies of nasal coarticulation using English subjects, which in terms of its linguistic contrast and in terms of being a rather 'sluggish' articulation behaves in a similar way to labial coarticulation, again showed outcomes which are at variance with each other. Thompson and Hixon's (1979) nasal airflow measurements at the midpoint of the first vowel in /ini/ showed a greater proportion of anticipatory nasalization with increasing age. Flege (1988), comparing the time of velopharyngeal opening and closing during the vowel in /dVn/ and /nVd/ sequences of adults and children aged 5 and 10 years, observed that both groups of speakers nasalized most vowels with no difference in the extent and degree of nasalization.

In view of the results, it becomes obvious that it is premature to derive any general statements about the acquisitional process of coarticulation. The results testify to a general point which is best summarized in Repp's (1986) words "The various patterns of results (...) suggest that phenomena commonly lumped together under the heading of "coarticulation" may have diverse origins and hence different roles in speech development. Some forms of coarticulation are an indication of advanced speech production skills, whereas others may be a sign of articulatory immaturity, and yet others are neither because they simply cannot be avoided. Therefore, it is probably not wise to draw conclusions about a general process called coarticulation from the study of a single effect. Indeed, such a general process may not exist." (p.1634).

Speaker-specific behaviour can not only be observed during the process of speech acquisition, but to a certain extent also in the coarticulation strategies of adult speakers. Nolan (1983, 1985) argues for the view that coarticulatory strategies are potentially idiosyncratic, and presents between-speaker variation in the coarticulation of English /l/ and /r/ with following vowels.

Lubker and Gay (1982) examined upper lip movement and EMG activity of four labial muscles in speakers of Swedish and American English. In addition to language-specific differences, they observed different coarticulatory effects for speakers within one language. Among the five Swedish subjects, three seemed to use a look-ahead mechanism, i.e., they adjusted the onset of labial movement to the time available. In contrast, the onset of lip rounding for the other speakers remained essentially constant prior to the acoustic vowel onset and, hence, their behaviour is in accord with the frame model.

Perkell and Matthies (1992) examined upper lip movements using a strain-gauge cantilever system, and included minimal pairs to control for the possibility (mentioned above) of consonant-inherent protrusion effects. The subjects in this investigation were carefully preselected; only speakers without pronounced regional dialect were chosen who primarily used lip protrusion, as opposed to lip closure, in uttering words containing the vowel /u/. Nevertheless, "there was a lot of variation in the shape and timing of /u/ protrusion trajectories: on a token-to-token basis within each utterance, between utterances, and across subjects" (1992:2917). The overall outcome of the study did not support, as with most studies, the predictions of any model in a straightforward way. Rather, the authors suggested that the initiation of anticipatory lip rounding is determined by three competing constraints, reflecting elements of both the dynamic properties of coproduction and the perceptually motivated requirements of feature-spreading models. Broadly speaking, the three constraints were, (i) end the protrusion movement during the voiced part of /u/, (ii) use a preferred gesture duration, and (iii) begin the protrusion movement when permitted. Thus, variations in the degrees to which the constraints were expressed accounted for differences between the subjects (and, in addition, for interindividual variations).

These and other studies indicate that speakers seem indeed to have some freedom in coarticulatory behaviour which is beyond that attributable to anatomical differences. Thus far, however, individual differences have been little investigated, and have rarely been taken into account in the formulation of theories. On the one hand, due to methodological limitations the use of a large number of subjects is still rare in most speech production experiments. Thus, truly quantitative statements about coarticulatory effects are yet difficult to derive. A statistical case study by Forrest et al. (1990) shows clearly that care has to be taken in generalizing results from a small number of speakers.

On the other hand, the high variability found in the data makes it difficult to differentiate between effects which should be considered as being idiosyncratic and effects which simply reflect the allowed range of variation for the phenomenon. For example, while, from a classical point of view, the outcome of Lubker and Gay's study (1982) could be considered as representing two fundamentally different organisational strategies for integrating segments, the refined experiment by Perkell and Matthies (1992) rather suggests that the various forces are present at any moment in time for any speaker. That is, they are an integral part of producing and coarticulating segments and speakers only vary in the emphasis they put on the different forces or even fluctuate between them.

It may be hoped that further attention to the development of coarticulation in children, and to the variation exhibited by mature speakers, will contribute to a better overall understanding of the process of speech production.

4. CONCLUSION

This chapter has considered three senses of the 'origin' of coarticulation: the reason it exists in speech; its history as a scientific construct; and its ontogenetic development in language acquisition. It was noted in Section 1 that coarticulation presupposes the existence at some level of discrete phonological units. Most obviously, the phenomena described as coarticulation are the result of the integration of those units in the continuously flowing activity of the vocal mechanism, which is ill-suited to abrupt changes of configuration. But, additionally, it is quite possible that coarticulation is favoured for perceptual reasons: putting it crudely, the longer a phonetic property persists, the more chance it has of being spotted.

Historically, the advent of instrumental techniques in the 19th century led to a crisis in the conceptualisation of speech (Section 2). No longer was it tenable to visualise speech as the concatenation of 'letters', because physical records stubbornly resisted attempts at strict segmentation. Gradually, from perspectives as diverse as phonology, acoustics, and the study of skilled motor activity, the foundations were laid for the concept of coarticulation. At first, coarticulation was mainly an attempt to come to terms with recalcitrant instrumental data which failed to confirm researchers' preconceptions about the nature of speech, but subsequently it became recognised as a phenomenon central to the study of speech production, and in particular from the 1960s onwards became the catalyst for an impressive variety of experiments and theorising in speech motor control. Throughout, there has been a productive tension between approaches which were more dynamically or more phonologically oriented. In some of the most influential recent work, for instance in 'Articulatory Phonology', the phenomenon of coarticulation has led to a re-evaluation of such distinctions, and to the attempt to use inherently dynamic units as phonological primes.

Studies of children (Section 3) have not led to a definitive picture of whether coarticulation is greater, less, or basically the same in childhood. It may be that the emergence of clear trends are confounded by inter-individual differences, which would be compatible with the view that each language learner has to work out, within the constraints of his or her individual vocal tract, a solution to the integration of phonetic segments. This view is supported by indications that coarticulatory behaviour is quite variable even between adult speakers of the same language variety.

Coarticulation, then, bears on issues such as the control of the skilled activity of speech production, the relation between phonological systems and their realisation, and the borderline between what is shared in the social code of language and what is individual. The systematic study of coarticulation is central to the development of experimental phonetics and speech science.

REFERENCES

Abelin, A., Landberg, I. & Persson, L. (1980). *A study of anticipatory labial coarticulation in the speech of children*. Perilus 2, 2-18.

Bell, A. (1867). *Visible Speech. Universal Alphabetics or Self-Interpreting Physiological Letters for the Writing of all Languages in one Alphabet*. London: Simpkin & Marschall.

Bell-Berti, F. & Harris, K. (1979). *Anticipatory coarticulation: some implications from a study of lip rounding*. Journal of the Acoustical Society of America 65, 1268-1270.

Bell-Berti, F. & Harris, K. (1981). *A temporal model of speech production*. Phonetica 38, 9-20.

Bell-Berti, F. & Harris, K. (1982). *Temporal patterns of coarticulation: lip rounding*. Journal of the Acoustical Society of America 71, 449-454.

Bell-Berti, F. & Krakow, R. (1991). *Anticipatory velar lowering: a coproduction account*. Journal of the Acoustical Society of America 90, 112-123.

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

Bladon, A. & Al-Bamerni, A. (1976). *Coarticulation resistance in English /l/*. Journal of Phonetics 4, 137-150.

Bladon, A. & Al-Bamerni (1982). *One stage and two-stage temporal patterns of velar coarticulation*. Journal of the Acoustical Society of America 72, S104 (abstract).

Bladon, A. & Nolan, F. (1977). *A video-fluorographic investigation of tip and blade alveolars in English*. Journal of Phonetics 5, 185-193.

Boyce, S., Krakow, R. & Bell-Berti, F. (1991). *Phonological underspecification and speech motor organization*. Haskins Laboratories Status Report on Speech Research SR-l05/106, 141-152.

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

Browman, C. & Goldstein, L. (1989). *Articulatory gestures as phonological units*. Phonology 6, 201-251.

Browman, C. & Goldstein, L. (1990). *Tiers in articulatory phonology, with some implications for casual speech*. In M. Beckman and J. Kingston (eds.), Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech. Cambridge: Cambridge University Press, 341-376.

Brücke, E. (1856). *Grundzüge der Physiologie und Systematik der Sprachlaute für Linguisten und Taubstummenlehrer*. Wien: Gerold.

Butcher, A. & Weiher, E. (1976). *An electropalatographic investigation of coarticulation in VCV sequences*. Journal of Phonetics 4, 59-74.

Carney, P. & Moll, K. (1971). *A cinefluorographic investigation of fricative consonant-vowel coarticulation*. Phonetica 23, 193-202.

Daniloff, R. & Hammarberg, R. (1973). *On defining coarticulation*. Journal of Phonetics 1, 239-248.

Engstrand, 0. (1980). *Acoustic constraints or invariant input representation? An experimental study of selected articulatory movements and targets*. Report from Uppsala University Department of Linguistics 7, 67-95.

Engstrand, 0. (1989). *Towards an electropalatographic specification of consonant articulation in Swedish*. Perilus 10, 115-156.

Farnetani, E. (1990). *V-C-V lingual coarticulation and its spatiotemporal domain*. In W. Hardcastle & A. Marchal (eds.), Speech Production and Modelling. Dordrecht: Kluwer, 93-130.

Farnetani, E., Vagges, K. & Magno-Caldognetto, E. (1985). *Coarticulation in Italian /VtV/ sequences: a palatographic study*. Phonetica 42, 78-99.

Flege, J. (1988). *Anticipatory and carry-over nasal coarticulation in the speech of children and adults*. Journal of Speech and Hearing Research 31, 525-536.

Forrest, K., Weismer, G. & Adams, S. (1990). *Statistical comparison of movement amplitudes from groupings of normal geriatric speakers*. Journal of Speech and Hearing research 33, 386-389.

Fowler, C. (1980). *Coarticulation and theories of extrinsic timing*. Journal of Phonetics 8, 113-133.

Fowler, C. (1981). *Production and perception of coarticulation among stressed and unstressed vowels*. Journal of Speech and Hearing Research 24, 127-139.

Fowler, C. (1983). *Realism and unrealism: a reply*. Journal of Phonetics 11, 303-322.

Fowler, C. (1985). *Current perspective on language and speech production: a critical overview*. In R. Daniloff (ed.), Speech Science. London: Taylor and Francis, 193-278.

Fowler, C. (1990). *Some regularities in speech are not consequences of formal rules: comments on Keating's paper*. In M. Beckman and J. Kingston (eds.), Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech. Cambridge: Cambridge University Press, 476-489.

Fowler, C. (1992). *Phonological and articulatory characteristics of spoken language*. Haskins Laboratories Status Report on Speech Research SR-10911 10, 1-12.

Fowler, C., Rubin, P., Remez, R. & Turvey, M. (1980). *Implications for speech production of a general theory of action*. In B. Butterworth (ed.), Language Production I. London: Academic Press, 371-420.

Fowler, C. & Saltzman, E. (1993). *Coordination and coarticulation in speech production*. Language and Speech 36, 171-195.

Gay, T. (1977). *Articulatory movements in VCV sequences*. Journal of the Acoustical Society of America 62, 183-193.

Gay, T. (1978). *Articulatory units: segments or syllables?* In A. Bell and J. Hooper (eds.), Syllables and Segments. Amsterdam: North-Holland, 121-131.

Gelfer, C., Bell-Berti, F. & Harris, K. (1989). *Determining the extent of coarticulation: effects of experimental design*. Journal of the Acoustical Society 6, 2443-2445.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.

Halwes, T. & Jenkins, J. (1971). *Problem of serial order in behavior is not resolved by context-sensitive associative memory models*. Psychological Review 78, 122-129.

Hammarberg, R. (1976). *The metaphysics of coarticulation*. Journal of Phonetics 4, 353-363.

Hammarberg, R. (1982). *On redefining coarticulation*. Journal of Phonetics, 10, 123-137.

Hardcastle, W. (1981). *Experimental studies in lingual coarticulation*. In R. Asher & E. Henderson (eds.), Towards a history of phonetics. Edinburgh: Edinburgh University Press, 50-66.

Henke, W. (1966). *Dynamic articulatory model of speech production using computer simulation*. PhD Dissertation, MIT.

Holst T. & Nolan F. (1995) *The influence of syntactic structure on [s] to [ʃ] assimilation*. In: B. Connell and A. Arvaniti (eds), Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV. Cambridge: CUP, pp 315-333.

Hoole, P. Gfroerer, S. & Tillmann, H.-G. (1990). *Electromagnetic articulography as a tool in the study of lingual coarticulation*. Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München 28, 107-122.

Hoole, P. (1996). "*Theoretische und methodische Grundlagen der Artikulationsanalyse in der experimentellen Phonetik*", FIPKM, 34, 3-156.

Jakobson, R., Fant, G. & Halle, M. (1952). *Preliminaries to Speech Analysis: the Distinctive Features and their Correlates*. Cambridge, MA: MIT Press.

Jones, D. (1932). *Outline of English Phonetics*. Cambridge: Cambridge University Press (3rd. ed.).

Katz, W., Kripke, C. & Tallal, P. (1991). *Anticipatory coarticulation in the speech of adults and young children: acoustic, perceptual and video data*. Journal of Speech and Hearing Research 34, 1222-1232.

Keating, P. (1988). *Underspecification in phonetics*. Phonology 5, 275-292.

Keating, P. (1990). *The window model of coarticulation: articulatory evidence*. In M. Beckman and J. Kingston (eds.), Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech. Cambridge: Cambridge University Press, 451-470.

Kelso, J., Saltzman, E. & Tuller, B. (1986). *The dynamic perspective on speech production: data and theory*. Journal of Phonetics 14, 29-59.

Kent, R. (1983). *The segmental organization of speech*. In P. MacNeilage (ed.), The Production of Speech. New York: Springer, 57-89.

Kent, R. (1986). *Is a paradigm change needed?* Journal of Phonetics 14, 111-115.

Kent, R., Carney, P. & Severeid, L. (1974). *Velar movement and timing: evaluation of a model for binary control*. Journal of Speech and Hearing Research 17, 470-488.

Kent, R. & Forner, L. (1980). *Speech segment durations in sentence recitations by children and adults*. Journal of Phonetics 8, 157-168.

Kent, R. & Minifie, F. (1977). *Coarticulation in recent speech production models*. Journal of Phonetics 5, 115-133.

Kent, R. & Moll, K. (1972). *Tongue body articulation during vowel and diphthong gestures*. Folia Phoniatrica 24, 286-300.

Kiritani, S., Itoh, K., Hirose, H. & Sawashima, M. (1977). *Coordination of the consonant and vowel articulations - X-ray microbeam study on Japanese and English*. Annual Bulletin of the Research Institute of Logopedics and Phoniatrics (Tokyo University) 11, 31-37.

Kiritani, S. & Sawashima, M. (1986). *The temporal relationship between articulations of consonants and adjacent vowels*. In R. Channon and L. Shockey (eds.), In honor of Ilse Lehiste. Dordrecht: Foris.

Kozhevnikov, A. & Chistovich, L. (1965). *Speech, articulation, and perception*. Washington: National Technical Research Service, 30 (US. Department of Commerce).

Kühnert, B., Ledl, C., Hoole, P. & Tillmann, H.-G. (1991). *Tongue-jaw interactions in lingual consonants*. Perilus 12, 21-25.

Lindblom, B. (1963). *Spectrographic study of vowel reduction*. Journal of the Acoustical Society of America 35, 1773-1781.

Lindblom, B. (1990). *Explaining phonetic variation. A sketch of the H&H theory*. In A. Marchal and W. Hardcastle (eds), Speech Production and Speech Modelling. Dordrecht: Kluwer, 403-439.

Lubker, J. (1981). *Temporal aspects of speech production: anticipatory labial coarticulation*. Phonetica 38, 51-65.

Lubker, J. & Gay, T. (1982). *Anticipatory labial coarticulation: experimental, biological, and linguistic variables*. Journal of the Acoustical Society of America 71 437-448.

MacNeilage, P. (1970). *Motor control of serial ordering of speech*. Psychological Review 77, 182-196.

Menzerath, P. & De Lacerda, A. (1933). *Koartikulation, Steuerung und Lautabgrenzung*. Berlin u. Bonn: Fred. Dümmlers.

Moll, K. & Daniloff, R. (1971). *Investigation of the timing of velar movement during speech*. Journal of the Acoustical Society of America 50, 678-684.

Nittrouer, S., Studdert-Kennedy, M. & McGowan, R. (1989). *The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults*. Journal of Speech and Hearing Research 32, 120-132.

Nittrouer, S. & Whalen, D. (1989). *The perceptual effects of child-adult differences in fricative vowel coarticulation*. Journal of the Acoustical Society of America 86, 1266-1276.

Nolan, F. (1982). *The role of Action Theory in the description of speech production*. Linguistics 20, 287-308.

Nolan, F. (1983). *The phonetic bases of speaker recognition*. Cambridge: Cambridge University Press.

Nolan, F. (1985). *Idiosyncrasy in coarticulatory strategies*. Cambridge Papers in Phonetics and Experimental Linguistics 4, 1-9.

Nolan, F., Holst, T. & Kühnert B. (1996) *Modelling [s] to [ʃ] accommodation in English*. Journal of Phonetics, 24, 113-137

Öhman, S. (1966). *Coarticulation in VCV utterances: spectrographic measurements*. Journal of the Acoustical Society of America 39, 151-168.

Parker, F. & Walsh, T. (1985). *Mentalism vs physicalism: a comment on Hammarberg and Fowler*. Journal of Phonetics 13, 147-153.

Paul, H. (1898). *Prinzipien der Sprachgeschichte*. Halle: Niemeyer (3rd ed.).

Perkell, J. (1969). *Physiology of speech production: results and implications of a quantitative cineradiographic study*. Cambridge, Mass.: MIT Press.

Perkell, J. (1986). *Coarticulation strategies: preliminary implications of a detailed analysis of lower lip protrusion movements*. Speech Communication 5, 47-86.

Perkell, J. & Chiang, C.-M. (1986). *Preliminary support for a 'hybrid model' of anticipatory coarticulation*. Proceedings of the 12th International Congress of Acoustics. Toronto: Canadian Acoustical Association, A3-6.

Perkell, J. & Matthies, M. (1992). *Temporal measures of anticipatory labial coarticulation for the vowel /u/: within- and cross-subject variability*. Journal of the Acoustical Society of America 91,2911-2925.

Recasens, D. (1984). *V-to-V coarticulation in Catalan VCV sequences*. Journal of the Acoustical Society of America 76, 1624-1635.

Recasens, D. (1987). *An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences*. Journal of Phonetics 15, 299-312.

Recasens, D. (1989). *Long range coarticulation effects for tongue dorsum contact in VCVCV sequences*. Speech Communication 8, 293-307.

Repp, B. (1986). *Some observations in the development of anticipatory coarticulation*. Journal of the

Acoustical Society of America 79, 1616-1619.

Rousselot, P.-J. (1897-1901). *Principes de phonétique experimentale, I-II*. Paris: H. Welter.

Saltzman, E. & Munhall, K. (1989). *A dynamical approach to gestural patterning in speech production*. Ecological Psychology 1, 333-382.

Scripture, E. (1902). *The Elements of Experimental Phonetics*. New York: Charles Scribners Sons.

Sereno, J., Baum, S., Marean, C. & Lieberman, P. (1987). *Acoustic analyses and perceptual data on anticipatory labial coarticulation in adults and children*. Journal of the Acoustical Society of America 81,521-519.

Sereno, J. & Lieberman, P. (1987). *Developmental aspects of lingual coarticulation*. Journal of Phonetics 15, 247-257.

Shaffer, L. (1984). *Motor programming in language-production - a tutorial review*. In H. Bouma and D. Bouwhuis (eds.), Attention and Performance X: Control of Language Processes. London: Lawrence Erlbaum Associates, 17-41.

Sharf, D. & Ohde, R. (1981). *Physiologic, acoustic, and perceptual aspects of coarticulation: implications for the remediation of articulatory disorders*. In N. Lass (ed.), Speech and Language: Advances in Basic Research and Practice, Vol.5. New York: Academic Press, 153-247.

Sharkey, S. & Folkins, J. (1985). *Variability of lip and jaw movements in children and adults: implications for the development of speech motor control*. Journal of Speech and Hearing Research 28, 3-15.

Sievers, E. (1876). *Grundzüge der Lautphysiologie zur Einführung in das Studium der Lautlebre der Indogermanischen Sprachen*. Leipzig: Breitkopf and Härtel.

Stetson, R. (1951). *Motor phonetics: A study of speech movements in action*. Amsterdam: North-Holland (2nd ed.).

Sussman, H. & Westbury, J. (1981). *The effects of antagonistic gestures on temporal and amplitude parameters of anticipatory labial coarticulation*. Journal of Speech and Hearing Research 24, 16-24.

Thompson, A. & Hixon, T. (1979). *Nasal air flow during normal speech production*. Cleft Palate Journal 16, 412-420.

Tillmann, H.-G. and Mansell, P. (1980). *Phonetik*. Stuttgart: Klett-Cotta.

Tillmann, H.-G. (1994). *Phonetics, early modem, especially instrumental work*. In R. Asher (ed.), The Encyclopedia of Language and Linguistics, Vol.6. Oxford: Pergamon, 3082-3094.

Wickelgren, W. (1969). *Context-sensitive coding, associative memory, and serial order in (speech) behavior*. Psychological Review 76, 1-15.

Wickelgren, W. (1972). *Discussion paper on speech perception*. In J. Gilbert (ed.), Speech and Cortical Functioning. New York: Academic Press, 237-262.