

# 9 Coarticulation and Connected Speech Processes

---

EDDA FARNETANI AND  
DANIEL RECASENS

## 1 Speech Contextual Variability

### 1.1 Coarticulation

A fundamental and extraordinary characteristic of spoken language, of which we speakers are not even conscious, is that the movements of different articulators for the production of successive phonetic segments overlap in time and interact with one another: as a consequence, the vocal tract configuration at any point in time is influenced by more than one segment. This is what the term “coarticulation” describes. The acoustic effects of coarticulation can be observed by means of spectrographic analysis: any acoustic interval, auditorily defined as a phonetic segment, will show the influence of neighboring phones in various forms and degrees. These effects are usually not audible, which is why their descriptive and theoretical study in various languages became possible only after physiological and acoustical methods of speech analysis became available and widespread during the last 40 years.

Table 9.1 shows how coarticulation can be described in terms of: (1) the main articulators involved; (2) some of the muscles considered to be primarily responsible for the articulatory movements; (3) the movements that usually overlap in contiguous segments; (4) the major acoustic consequences of such overlap. As for lingual coarticulation, the tongue tip/blade and the tongue body can act quasi-independently as two distinct articulators, so that their activity in the production of adjacent segments may overlay in time.

Jaw movements are not included in the table since the jaw contributes both to lip and to tongue positioning and, therefore, may be considered part of the labial and lingual subsystems. Mandibular movements are analyzed especially when the goal of the experiment is to establish the role of the jaw in shaping the vocal tract and thus distinguish between active and passive tongue or lip movements, or to investigate how the jaw contributes to or compensates for coarticulatory variations.

**Table 9.1** Coarticulation: Levels of description

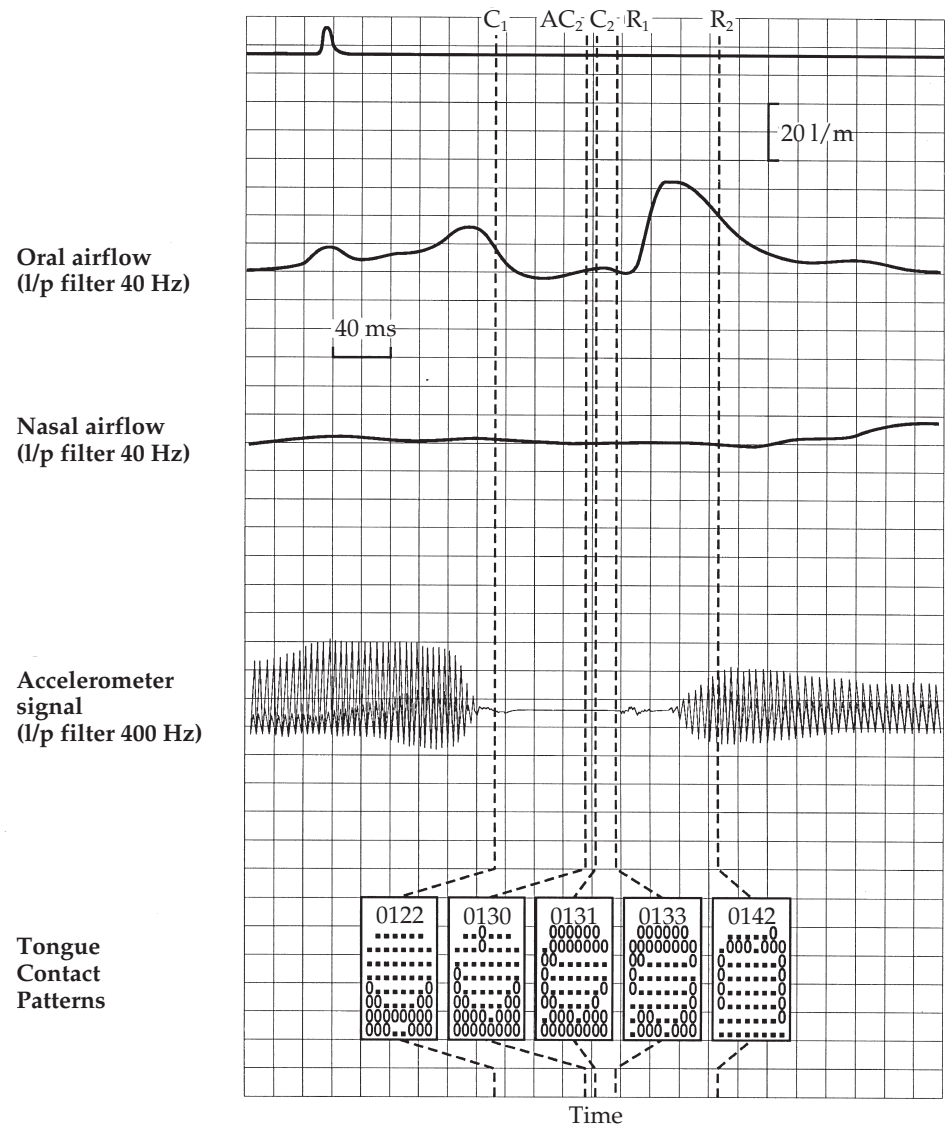
<i>Articulators</i>	<i>Muscles</i>	<i>Articulatory activity</i>	<i>Acoustic consequences</i>
LIPS	Orbicularis oris/Risorius	Lip rounding/spreading	Formant changes
TONGUE	Genioglossus, and other extrinsic and intrinsic lingual muscles	Tongue front/back and high/low displacement	Formant changes
VELUM	Levator palatini	Velum lowering	Nasal formants and changes in oral formant structure
LARYNX	Posterior cricoarytenoid/ Lateral cricoarytenoid, Interarytenoid	Vocal fold abduction/adduction	Signal (a)periodicity

Examples of coarticulation in terms of muscle activity and of articulatory movements are given in Figures 9.1 and 9.2. They show overlapping activity both at the level of commands and that of execution.

Data from Hirose and Gay (1972) illustrate coarticulation at the myomotoric level. Electromyographic activity of four muscles during the production of the sequences /əpɪb/ and /əpɪp/ indicates that the activity of orbicularis oris for the production of the first /p/ is overlapped by the activity of the genioglossus for the production of the following front vowel. Moreover the activity of the laryngeal muscles responsible for abducting and adducting the vocal folds also overlap: the onset of lateral cricoarytenoid activity (LCA, adducting) occurs when the posterior cricoarytenoid activity (PCA, abducting) is at its peak, that is, at the middle of the /p/ closure.

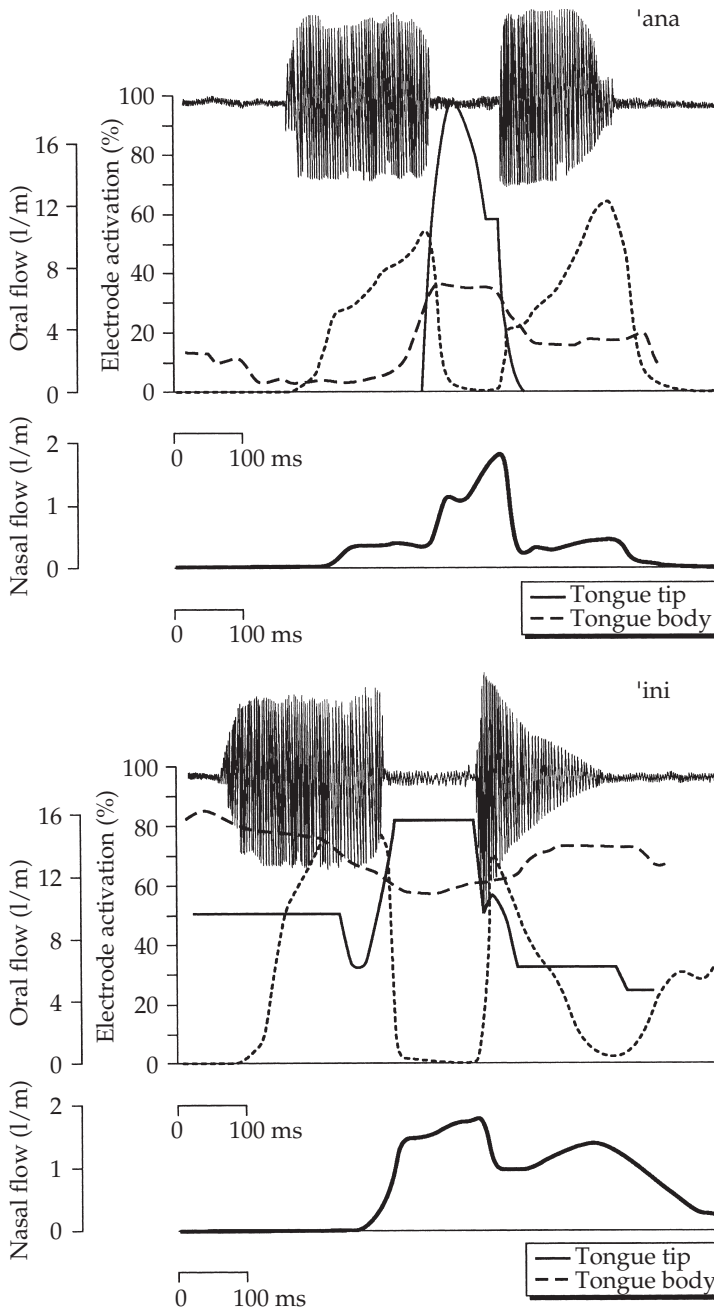
Figure 9.1 describes tongue-tip–tongue-body coarticulation in the /kl/ cluster of the English word *weakling*, analyzed with electropalatography (EPG) and synchronized with oral and nasal airflow and with the acoustic signal (from Hardcastle, 1985). The sequence of the EPG frames (sampled every 7.5 ms) describes the articulation of the /kl/ cluster: it can be seen that the tongue-body closure for the velar consonant is overlapped by the tongue-tip/blade gesture for the following /l/, detectable by a light front contact as early as frame 130. The following frames show complete overlap of /k/ and /l/ closures for about 20 ms.

Figure 9.2 is an example of velar and lingual coarticulation in the sequences /'ana/ and /'ini/ in Italian, analyzed with EPG and oral/nasal flow measurements (Farnetani, 1986). As for velar coarticulation, the nasal flow curves (continuous



**Figure 9.1** Oral and nasal flow curves, acoustic signal, and synchronized EPG activity in the production of the English word *weakling* within phrase. (From Hardcastle, 1985)

thick lines) indicate that in /'ana/ the opening of the velopharyngeal port for the production of /n/ occurs just after the acoustic onset of the initial /a/ and lasts until the end of the final /a/; in /'ini/ there is only a slight anticipation of velopharyngeal opening during the initial /i/, but after /n/ the port remains open until the end of the utterance. Thus, velar C-to-V coarticulation extends both in the



**Figure 9.2** Acoustic signal, oral and nasal flow curves, and synchronized EPG curves during /'ana/ and /'ini/ produced in isolation by an Italian subject. (From Farnetani, 1986)

anticipatory and in the carryover direction in the sequence /'ana/, while extending mostly in the carryover direction in the sequence /'ini/. The two EPG curves represent the evolution of tongue-to-palate contact over time. It can be seen that during tongue-tip/blade closure for /n/, tongue-body contact is much larger in the context of /i/ than in the context of /a/, indicating that the tongue-body configuration during the consonant is strongly affected by the vowels. These patterns describe V-to-C lingual coarticulation, that is, the effect of the vowel on the articulation of the consonant.

These examples clearly show that speaking is coarticulating gestures. The central theoretical issues in the studies of coarticulation concern its origin, function, and control. Coarticulation has been observed in all languages so far analyzed, and can be considered a universal phenomenon, even if it appears to differ among languages. Before exploring these issues, the assimilatory and connected speech processes will be discussed at some length in the next sections.

## 1.2 Assimilation

Assimilation refers to contextual variability of speech sounds, by which one or more of their phonetic properties are modified and become like those of the adjacent segment. Are assimilation and coarticulation qualitatively different processes, the former reflecting an auditory approach to phonetic analysis, and the latter an instrumental articulatory/acoustic approach? The answers are various and controversial. Standard generative phonology (Chomsky & Halle, *The Sound Pattern of English*, 1968) makes a clear-cut distinction between assimilation and coarticulation. Assimilation pertains to the domain of linguistic competence, is accounted for by phonological rules, and refers to modifications of features defined as the minimal categorical-classificatory constituents of a phoneme; hence, assimilatory processes are part of the grammar and are language-specific. Coarticulation, by contrast, results from the physical properties of the speech mechanism and is governed by universal rules; hence, it pertains to the domain of performance and cannot be part of the grammar. Chomsky and Halle include as coarticulation effects “the transition between a vowel and an adjacent consonant, the adjustments in vocal tract shape made in anticipation of subsequent motions, etc.” (1968, p. 295).

Quite often context-dependent changes involving the same articulatory structures have different acoustic and perceptual manifestations in different languages so that it is possible to distinguish what can be considered universal phonetic behavior from language-particular rules. A classic example is the difference between vowel harmony, an assimilatory process present in a limited number of languages such as Hungarian, and vowel-to-vowel coarticulation, a process attested in a number of languages and probably present in all (Fowler, 1983). In other cases, cross-language differences are not easily interpretable, and inferences on the nature of the underlying production mechanisms can be made by manipulating some of the speech parameters, for example segmental durations. In a study of vowel nasalization in Spanish and American English, Solé and Ohala (1991) were able to distinguish

phonological (language-specific) nasalization from phonetic (universal) nasalization by manipulating speech rate. They found a quite different distribution of the temporal patterns of nasalization in the two languages as a function of speaking rate: the extent of nasalization on the vowel preceding the nasal consonant was proportional to the varying vowel duration in American English, while it remained constant in Spanish. A temporal increase of nasalization as vowel duration increases in American English must be intentional, i.e. phonological; on the other hand, the short, constant extent of nasalization in Spanish must be an automatic consequence of the speech mechanism, as it reflects the minimum time necessary for the lowering gesture of the velum.

But a strict dichotomy between universal and language-specific variations fails to account for the many cross-language data showing that coarticulation differs in degree across languages. A typical example is Clumeck's study on velar coarticulation, which was found to differ in temporal extent across all the six languages analyzed (Clumeck, 1976). In this case, what criterion can be used to decide in which language the patterns are unintentional and automatic, and in which they are to be ascribed to the grammar?

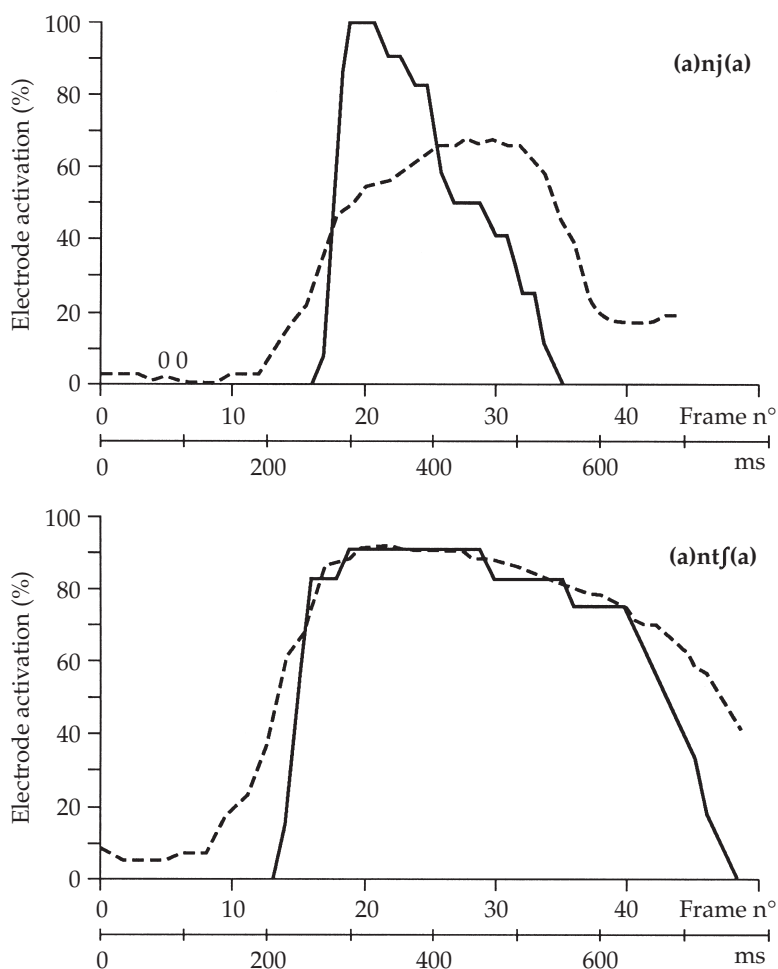
Likewise, within a given language, context-dependent variations may exhibit different articulatory patterns which can be interpreted either as the result of different underlying processes, or just as quantitative variations resulting from the same underlying mechanism. For example, Figure 9.3 shows the variations in the articulation of the phoneme /n/ as a function of the following phonetic segment, a palatal semivowel (in /anja/) and a postalveolar or alveopalatal affricate consonant (in /antʃa/). The utterances are pseudowords produced in isolation by an Italian speaker, and analyzed with EPG.

In each graph the two curves represent the evolution of the tongue-to-palate contact over time. We can see that in /anja/, as the tongue tip/blade (continuous line) achieves maximum front contact for /n/ closure, the tongue body (dashed line) moves smoothly from /a/ to /j/ suggesting overlapping activity of two distinct articulators, and two distinct goals. In /antʃa/, instead, the typical /n/ configuration has disappeared; the front contact has decreased by some percentage points, and the back contact has appreciably increased; the cluster seems to be produced with one tongue movement. These differences may indicate that two distinct processes are at work in the two utterances: anticipatory coarticulation of /j/ on /n/ in the former, and place assimilation of /n/ to /tʃ/ in the latter.

Current theories of coarticulation have controversial views on whether there are qualitative or quantitative or even no differences between assimilatory and coarticulatory processes. It will be seen that at the core of the different positions are different answers to the fundamental issues addressed above, i.e., the domain, the function and the control of coarticulation.

### 1.3 *Connected speech processes*

In speech, the phonetic form of a word is not invariable but can vary as a function of a number of linguistic, communicative, and pragmatic factors (e.g., information



**Figure 9.3** EPG curves during /anja/ and /antja/ produced by an Italian subject. Continuous lines: tongue-tip/blade contact; dashed lines: tongue-body contact. (From Farnetani, 1986)

structure, style, communicative situation). These variations are generally referred to as “alternations,” and the phonetic processes accounting for them have been termed “connected speech processes” (Jones, 1969; Gimson, 1970). According to Gimson (p. 287), these processes describe the phonetic variations characterizing continuous speech when compared to a word spoken in isolation. The author makes a detailed list of connected speech processes in English: assimilation of place, manner, and voicing; reduction of vowels to schwa in unaccented words; deletion of consonants and vowels. According to the author, the factors that contribute to modify the phonetic properties of a word are “the pressures of its sound

environment or of the accentual or rhythmic group of which it forms part," and the speed of the utterance.

Kohler (1990) proposes an explanatory account of connected speech processes in German. Basing his analysis of German on the differences between careful and casual pronunciation of the same items, the author arrives at the conclusion that the so-called connected speech processes are a global phenomenon of reduction and articulatory simplification. These processes include /r/ vocalization to [ɐ] (when not followed by a vowel), weak forms, elision, and assimilation. From an analysis of the sound categories undergoing such changes, he infers that connected processes result from articulatory constraints, such as minimization of energy, which induce a reorganization of the articulatory gestures. He also proposes a formalization of the processes in terms of sets of phonetic rules, which generate any reduced segmental pronunciation.

The two accounts, although substantially different in their perspectives, have in common the assumption that connected speech processes imply modifications of the basic units of speech, i.e., elimination and replacements of articulatory gestures, changes in articulation places, etc. Hence, the main difference between connected speech processes and the phonological assimilations described in section 1.2 is that the latter occur independently of how a word is pronounced, while the former occur in some cases (e.g., in rapid, casual speech), but are absent in others (Chomsky & Halle, 1968, p. 110).

Recent theories of coarticulation also consider connected speech processes and propose their own accounts, as will be seen below in section 2.6.

## 2 Theoretical Accounts of Coarticulation

### 2.1 *Pioneering studies: Joos' overlapping innervation theory*

That speech is a continuum, rather than an orderly sequence of distinct sounds as listeners perceive it, was pointed out long ago (Sweet, 1877, cited by Wood, 1993). Sweet saw speech sounds as points "in a stream of incessant change" and this promoted the view that coarticulatory effects result from the transitional movements conjoining different articulatory targets, and reflected acoustically in the transitions to and from targets. Menzerath and de Lacerda (1933), to whom the term "coarticulation" is attributed, showed that segments can be articulated together, not merely conjoined to each other. The pioneering acoustic analysis of American English vowels conducted by Joos (1948) revealed that vowels vary as a function of neighboring consonants not only during the transitional periods but also during their steady state. Referring to temporal evolution of the second formant, Joos observed that "the effect of each consonant extends past the middle of the vowel, so that at the middle the two effects overlap" (p. 105). In his theoretical account of the phenomenon, he contests the "glide" hypothesis, which attributes coarticulation to the inertia of vocal organs and muscles: since no shift



from one articulatory position to another can take place instantaneously, a transition intervenes between successive phones. Joos proposes instead the “overlapping innervation wave theory” (p. 109): each phonetic segment command is an invariant “wave” that “waxes and wanes smoothly”; “waves for successive phones overlap in time.”

As will be seen below, these early hypotheses on the sequential ordering of speech segments have been highly influential in the development of coarticulation theories.

## ***2.2 Coarticulation as a component of the Grammar***

The evolution of featural phonology after Chomsky and Halle (1968) is marked by a gradual appropriation of coarticulation into the domain of linguistic competence.

**2.2.1 The theory of feature spreading** Daniloff and Hammarberg (1973) and Hammarberg (1976) were the promoters of the “feature-spreading” account of coarticulation. The view that coarticulation is a pure physiological process due to mechano-inertial constraints of the speech apparatus entails a sharp dichotomy between intent and execution, and implies that articulators are unable to carry out the commands as specified. The way to overcome this dichotomy is to assume that coarticulation itself is part of the phonological component. The arguments in support of this assumption are: (1) phonology is prior to phonetics, i.e., the phonology component underlies the phonetic implementation of speech sounds; (2) phonological segments are abstract entities, and cannot be altered by the physical speech mechanism; (3) the speech mechanism can only execute higher level commands. Hence, the variations associated to coarticulation must be the input to the speech mechanism. How? Segments have inherent and derived properties. These latter result from coarticulation, which alters the properties of a segment. Phonological rules stipulate which features get modified, and the phonetic representation, which is the input of the speech mechanism, specifies the details of articulation and coarticulation.

The departure from Chomsky and Halle’s view of coarticulation was probably necessary in face of emerging data on anticipatory lip protrusion (Daniloff & Moll, 1968) and velar coarticulation (Moll & Daniloff, 1971) showing that coarticulatory movements can be initiated at least two segments before the influencing one. These findings revealed that coarticulation is not the product of inertia. Another reason (Hammarberg, 1976) was that coarticulation cannot be accounted for by universal rules, owing to interlanguage differences.

Why does coarticulation occur? The function of coarticulation is to smooth out the differences between adjacent sounds: coarticulatory modifications accommodate the segments so that the transitions between them are minimized. In Daniloff and Hammarberg’s view anticipatory coarticulation is always a deliberate process, while carryover coarticulation is in part the effect of inertia and in part a feedback assisted strategy that accommodates speech segments to each other.

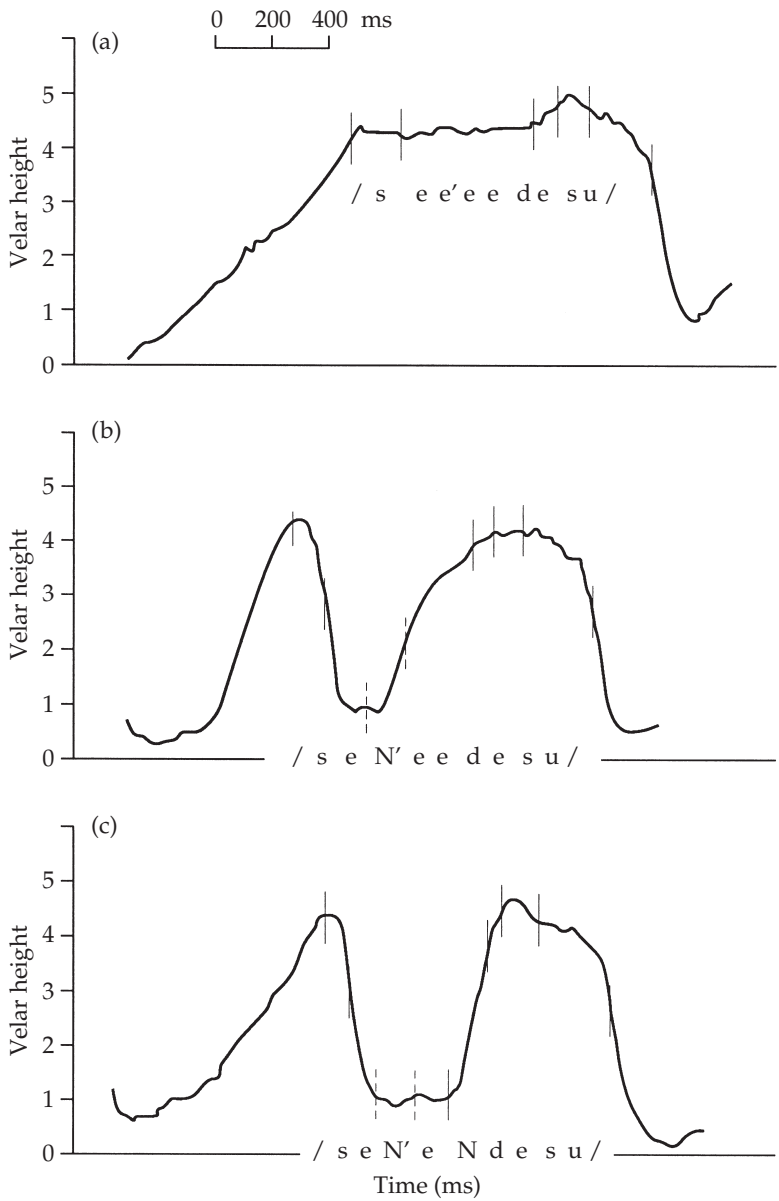
**2.2.2 Henke's articulatory model** The articulatory model of Henke (1966) best accounts for experimental data on the extent of coarticulation. It contrasts with another well-known account of coarticulation, the "articulatory syllable" model, proposed by Kozhevnikov and Chistovich (1965). This model is based on data on anticipatory labial coarticulation in Russian, where segments appeared to coarticulate within, but not across  $C_nV$  sequences. Unlike the  $C_nV$  model, Henke's model does not impose top-down boundaries on anticipatory coarticulation; instead, input segments are specified for articulatory targets in terms of binary phonological features (+ or -), unspecified features being given the value 0. Coarticulation rules assign a feature of a segment to all preceding unspecified segments by means of a look-ahead scanning mechanism. The spread of features is blocked by a specified feature: for example the feature [+nasal] will be anticipated to all preceding segments unspecified for nasality.

**2.2.3 Feature specification and coarticulation: Coarticulatory resistance** While a number of experimental results are compatible with the hypothesis of feature spreading and the look-ahead mechanism, many others contradict the spatial and/or the temporal predictions of the model. First, the model cannot explain the extensive carryover effects observed in a number of studies on V-to-V coarticulation (for example, Magen, 1989; Recasens, 1989). Other disputed aspects of the theory are:

- 1 the adequacy of the concept of specified versus unspecified features for blocking or allowing coarticulation;
- 2 the hypothesis that a look-ahead mechanism can account for the temporal extent of anticipatory coarticulation.

As for the first issue, it appears that segments specified for a contradictory feature in asymmetric VCV sequences can nonetheless be modified by coarticulation. Data on lip rounding in French (Benguerel & Cowan, 1974) and English (Sussman & Westbury, 1981) indicate that in an  $/iC_nu/$  sequence type, lip rounding for  $/u/$  can start during  $/i/$ . Also, data on lingual coarticulation show that tongue displacement towards a vowel can begin during a preceding cross-consonantal vowel even if the two vowels are specified for conflicting features, for example, F2 and tongue-dorsum lowering effects from  $/a/$  on  $/i/$  (Öhman, 1966, for Swedish and English; Butcher & Weiher, 1976, for German; Farnetani et al., 1985, for Italian; Magen, 1989 for American English). Most interestingly, these transconsonantal V-to-V effects appear to vary in degree across languages, indicating that the same vowel categories are subject to different constraints that favor or disfavor coarticulatory variations in different languages (see Manuel & Krakow, 1984, comparing Swahili, Shona and English; Manuel, 1987, comparing three Bantu languages; Choi & Keating, 1991, comparing English, Polish, Russian, and Bulgarian).

As for phonologically unspecified segments, some experimental data are compatible with the idea that they completely acquire a contextual feature. Figure 9.4



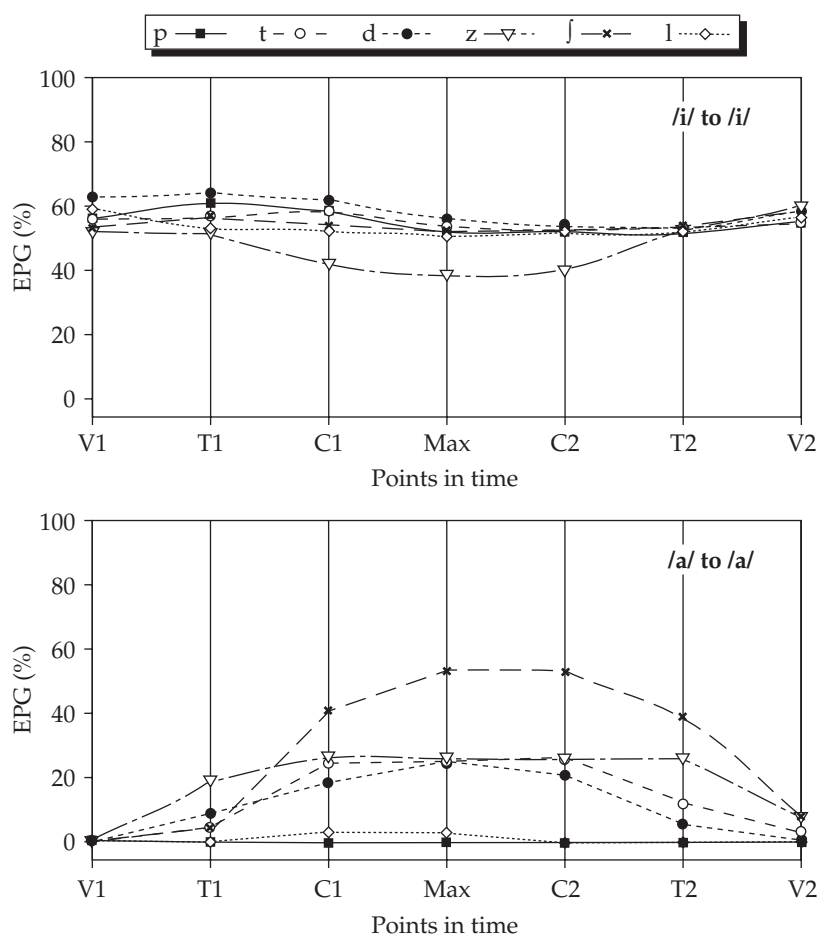
**Figure 9.4** Velar movement observed with fiberscope during Japanese utterances containing the vowel /e/ in oral and nasal contexts. (From Ushijima & Sawashima, 1972)

(from Ushijima & Sawashima, 1972) illustrates how, as predicted by the feature-spreading model, the Japanese vowel /e/ unspecified for nasality acquires the nasality feature in a symmetric context. The figure shows the amount of velum height during the vowel /e/ in Japanese, surrounded by oral consonants (panel a), by nasal consonants (panel c), and in a mixed environment (panel b). It can be seen that during /e/ the velum is as high as for the oral consonants in (a), and as low as for the nasal consonants in (c): in both cases the velum height curve runs nearly flat across the vowel. Instead, in the asymmetric example (b) the curve traces a trajectory from a high to a low position during the /e/ preceded by /s/ and the reverse trajectory during the /e/ followed by /d/. The symmetric sequences show that /e/ is completely oral in an oral context and completely nasalized in a nasal context, indicating that this vowel has no velar target of its own but acquires that of the contextual phonetic segment(s). The trajectories in the asymmetric sequences do not contradict the hypothesis that this vowel has no target for velar position, but contradict the assumption that contextual features are spread in a categorical way. Accordingly, /e/ would have to be completely nasalized from its onset when followed by a nasal, and completely oral when followed by an oral consonant, and this does not seem to occur.

Many other data indicate that phonologically unspecified segments may nonetheless exhibit some resistance to coarticulation and, therefore, are specified for articulatory targets. English data on velar movements (Bell-Berti, 1980; Bell-Berti & Krakow, 1991) show that the oral vowels are not articulatorily neutral to velar height, and have their own specific velar positions even in a non-nasal environment. As for lip position, Engstrand's data (1981) show that in Swedish /u-u/ sequences with intervocalic lingual consonants, protrusion of the upper lip relaxes during the consonants and the curve may form a "trough" between the vowels, suggesting that such consonants are not neutral to lip position. Troughs in lingual muscle activity during /ipi/ were first observed by Bell-Berti and Harris in their 1974 study (see further below for a different account of troughs).

Subsequent research on lingual consonants in Catalan, Swedish, and Italian (Recasens, 1984a, 1984c, 1987; Engstrand, 1989; Farnetani, 1990, 1991) revealed that consonants unspecified for tongue-body features coarticulate to different degrees with the surrounding vowels. As for coronals, those studies show that the amount of tongue-body coarticulation tends to decrease from alveolars to postalveolars, and from liquids (provided that /l/ is clear and the rhotic is a tap or an approximant) to stops to fricatives.

Figure 9.5 is an example of how different consonants coarticulate to different degrees with the surrounding vowels /i/ and /a/ in Italian, as measured by EPG. The trajectories represent the amount of tongue-body contact over time during the coronals /t/, /d/, /z/, /ʃ/ and clear /l/ and the bilabial /p/ in symmetric VCV sequences. The /iCi/ trajectories exhibit troughs of moderate degree for most consonants; /z/ shows the largest deviation indicating that the production of this consonant requires a lowering of the tongue body from the /i/ position. In the context of /a/, the consonants /p/ and /l/ coarticulate strongly with this vowel, as they show little or no contact, while for /t/, /d/, and /z/ the tongue body



**Figure 9.5** Tongue-body EPG contact at various points in time during symmetric (C)VCV isolated words in Italian. V1, V2 correspond to vowel mid-points; T1, T2 to offset of V1 and onset of V2 respectively; C1, C2 correspond to onset and release of consonant closures/constrictions respectively; Max refers to the point of maximum contact during the consonant. (From Farnetani, 1991)

needs to increase contact to about 20 percent. During /ʃ/, tongue-body contact reaches the same value, i.e., between 50 and 60 percent, in the two vocalic contexts, indicating that this consonant is maximally resistant to coarticulation.

The overall data on tongue-body V-to-C coarticulation indicate that no alveolar consonant fully coarticulates with the adjacent vowels, which suggests the presence of a functional and/or physical coupling between tip/blade and body. The differences in coarticulation across consonants can be accounted for by consonant-specific manner constraints: fricatives must constrain tongue dorsum position to

ensure the appropriate front constriction and the intraoral pressure required for noise production; the production of stops and several laterals imposes lesser constraints, and allows for a wider range of coarticulatory variations.

The notion of coarticulatory resistance was introduced by Bladon and Al-Bamerni (1976) in an acoustic study of V-to-C coarticulation in /l/ allophones. Their data indicated that coarticulatory variations decrease from clear to dark to syllabic /l/. These graded differences could not be accounted for by binary feature analysis, which would predict complete blocking of coarticulation in the case of dark /l/ since this consonant variety is specified as [+back]. The authors propose a numerical index of coarticulatory resistance to be attached to the feature specification of each allophone. A subsequent study on tongue-tip/blade displacement in alveolar consonants in clusters (Bladon & Nolan, 1977) confirmed the idea that feature specification alone cannot account for the observed coarticulatory behavior. The coarticulation resistance scale has been developed more recently by the degree of articulatory constraint model of coarticulation or DAC (section 2.5.4 below).

All these studies show that the assignment of contextual binary features to unspecified segments through phonological rules fails to account for the presence versus absence of coarticulation, for its graded nature, and for the linguistically relevant aspects of coarticulation associated with this graded nature, i.e., the different degree of coarticulation exhibited by the same segments across languages. Explanation of these facts must be carried out in terms of articulatory, aerodynamic-acoustic, and perceptual constraints and, therefore, requires factors outside the world of phonological features.

## 2.3 *Coarticulation as speech economy*

**2.3.1. Adaptive variability in speech** The theoretical premise at the basis of Lindblom's theory of speech variability is that the primary scope of phonetics is not to describe how linguistic forms are realized in speech, but to explain and derive the linguistic forms from "substance-based principles pertaining to the use of spoken language and its biological, sociological, and communicative aspects" (Liljencrants & Lindblom, 1972, p. 859). Accordingly, in his theory of "Adaptive Variability" and "Hyper-/Hypo-speech" (Lindblom, 1983, 1989, 1990), phonetic variation is not viewed as mere consequence of inertia in the speech mechanism, but rather as a continuous adaptation of speech production to the demands of the communicative situation. Variation arises because production strategies change as a result of the interaction between system-oriented and output-oriented motor control. Some situations will require an output with a high degree of perceptual contrast, others will require less perceptual contrast and will allow more variability. Thus, the acoustic characteristics of the same item will exhibit a wide range of variation reflected along the continuum of over- to under-articulation, or hyper- to hypo-speech.

**2.3.2 Low-cost and high-cost production behavior** What is the function of coarticulation within the hyper-hypo framework? Coarticulation, manifested as

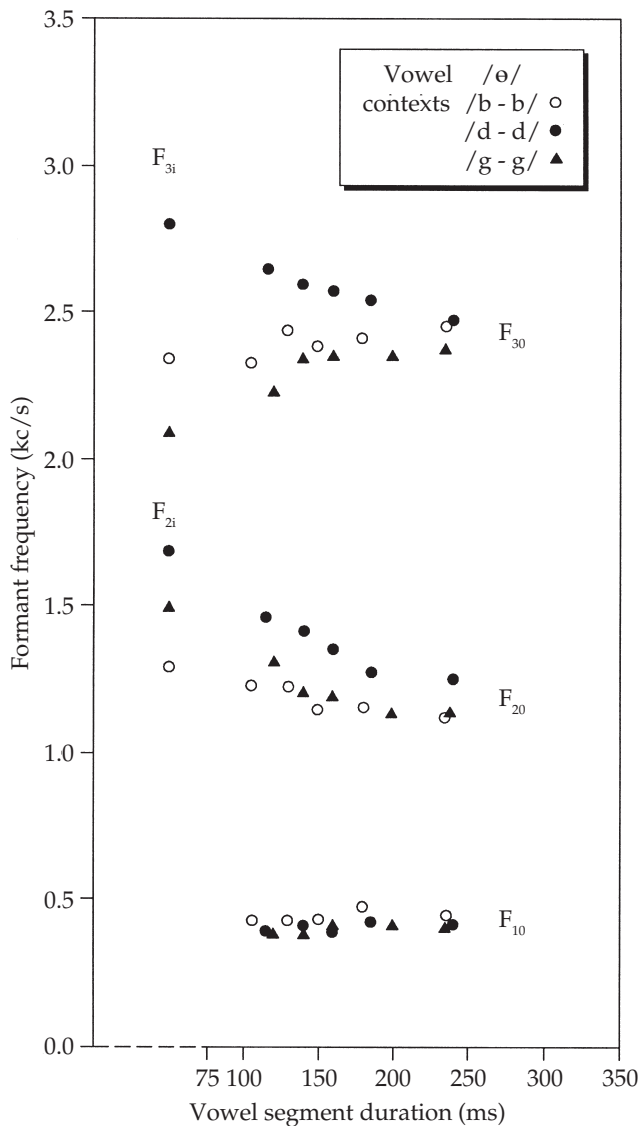
a reduced displacement and a shift of articulatory movements towards the contextual phonetic segments, is a low-cost motor behavior, an economical way of speaking. Its pervasiveness indicates that the speech motor system, like other kinds of motor behavior, is governed by the principle of economy.

In his study on vowel reduction, Lindblom (1963) introduced the notion of acoustic target, an ideal context-free spectral configuration which, in the case of vowels, is represented by the asymptotic values towards which formant frequencies aim. Lindblom's study showed that targets are quite often not realized: his data on CVC syllables indicated that the formant frequency values at vowel midpoint change monotonically with changes in vowel duration. At long vowel durations, formants tend to reach the target values; as vowel duration decreases, the formant movements are reduced and tend to shift towards the values of the adjacent consonants. This target undershoot process is shown in Figure 9.6.

Its continuous nature reveals that vowel reduction is an articulatory process, largely dependent on duration, rather than a phonological process. Indeed, the direction of the change towards the segmental context, as well as different degrees of undershoot as a function of the extent of the CV transitions (i.e., vowel reduction is minimal when the consonant-to-vowel distance is small), indicate that reduction is ruled by a coarticulatory rather than by a centralization mechanism leading towards a schwa-like configuration.

Lindblom's account of the relation between duration, target undershoot, and coarticulation was that reduction is the automatic response of the motor system to an increase in the rate of the motor commands. When successive commands on one articulator are issued at very short temporal intervals, the articulator has insufficient time to complete the response before the next signal arrives, and has to respond to different commands simultaneously, thus inducing both vowel shortening and reduced formant displacement. Subsequent research showed that the system response to high rate commands does not automatically result in reduced movements (Kuehn & Moll, 1976; Gay, 1978), and that reduction can occur also at slow rates (Nord, 1986). These studies indicated that speakers can adapt to different speaking situations and choose different production strategies to avoid or to allow reduction/coarticulation.

In the revised model of vowel undershoot (Moon & Lindblom, 1994), vowel duration is still the main factor, but variables associated with speech style, such as the rate of formant frequency change, can substantially modify the amount of formant undershoot. The model is based on an acoustic study of American English stressed vowels produced in clear speech and in citation forms, i.e., in overarticulated versus normal speech. Data on vowel duration and F2 indicate that vowels tend to be longer and less reduced in clear speech than in citation forms. A second finding is that clear speech is in most cases characterized by larger formant velocity values than citation forms. This means that the degree of context-dependent undershoot depends on speech style and tends to decrease with an increase in velocity of the articulatory movements. In this model where the speech motor mechanism is seen as a second-order mechanical system, it is proposed that undershoot is controlled by three variables reflecting the articulation



**Figure 9.6** Mean F1, F2 and F3 frequencies during the steady state of the Swedish vowel /ə/ plotted against vowel duration. The vowel is in the contexts of /b/, /d/ and /g/. As the vowel shortens, the F2 and F3 frequencies shift towards the formant values of the /bV/, /dV/ and /gV/ initial boundaries, F2i and F3i. (Reprinted with permission from B. Lindblom (1963), "Spectrographic study of vowel reduction," *Journal of the Acoustical Society of America*, 35, 1773–81. Copyright 1963, Acoustical Society of America.)



strategies available to speakers: duration, input articulatory force, and time constant of the system. An increase in input force and/or an increase in speed of the system response (i.e., a decrease in stiffness) contribute to increase the movement amplitude/velocity, and hence to decrease the amount of context-dependent undershoot. Thus, there appears to be an undershoot-compensatory reorganization of articulatory gestures in clear speech.

**2.3.3 Natural speech as a low-cost strategy** The experiment carried out by Lindblom et al. (1975) shows that a low-cost strategy, characterized by coarticulatory variations, is the norm in natural speech. The authors analyzed apical consonants in VCV utterances. Using a numerical model of apical stop production, they showed that the best match between the output of the model and spectrographic data of natural VCV utterances produced in isolation is a tongue configuration always compatible with the apical closure but characterized by a minimal displacement from the preceding vowel. In other words, the experiment shows that among a number of tongue-body shapes that facilitate tongue-tip closure, the tongue body always tends to take those requiring the least amount of movement and an adaptative behavior to the articulatory configuration of the adjacent vowels.

Lindblom's hypothesis, that the more speech style shifts towards the hypspeech pole the larger will be the amount of coarticulation, is confirmed by a number of studies on connected speech: Krull (1987, 1989), for Swedish; Duez (1991) for French; Farnetani (1991) for Italian.

**2.3.4 The Locus equation metrics as a measure of coarticulation** Locus equations were conceived by Lindblom (1963), who defined them as linear regressions of the onset of the F2 transition on the F2 target, measured at the vowel nucleus. He formulated locus equations as  $F2_{onset} = K F2_{vowel} + c$ , where the constants  $K$  and  $c$  are the slope and the intercept, respectively. Lindblom showed that the data points were aligned at about the regression line and that the slope and the intercept varied as a function of consonant place of articulation.

Krull (1989) pursued Lindblom's locus equations experiments. Most importantly, she found that the variations of slope and intercept as a function of consonant place in CV syllables were proportional to the extent of coarticulation between the vowel and the preceding consonant, such that flatter slopes indicate a more invariant locus and steeper slopes an increase in coarticulation. In other words, locus equation data are strongly related to the underlying coarticulatory behavior. Krull also found for CVC sequences that prevocalic stop consonants undergo stronger anticipatory vowel effects than postvocalic consonants undergo carryover vowel effects.

In a later study, Chennoukh et al. (1997) reported that locus equations depend on consonant place and on degree of coarticulation. Moreover, in a series of experiments conducted by Sussman and colleagues locus equations have been shown to account for degree of coarticulation in different conditions: coarticulation, obtained by locus equations, decreases in VC versus CV utterances with a stop consonant due to the greater articulatory precision in the production of prevocalic

than of postvocalic consonants (Sussman et al., 1997); slopes are identical for single and geminate open syllables and lower for closed syllables (Sussmann & Modarresi, 2003); even though there is less coarticulation for a coda stop in VC and C##V sequences than for a syllable onset stop in CV sequences, the locus equation slopes for the across-syllable and word-boundary conditions still differ as a function of place of articulation (Modarresi et al., 2004).

Lindblom et al. (2002) have proposed a novel interpretation of troughs repeatedly observed in speech. A trough is described as an apparent discontinuity of anticipatory coarticulation that takes the form of a momentary deactivation of tongue or lips movement after the first vowel in a VCV sequence. This event, according to Lindblom et al. (2002), could suggest a segment-by-segment activation pattern, as opposed to a V-to-V trajectory with an independent and superimposed consonant gesture as proposed by Öhman (1967).

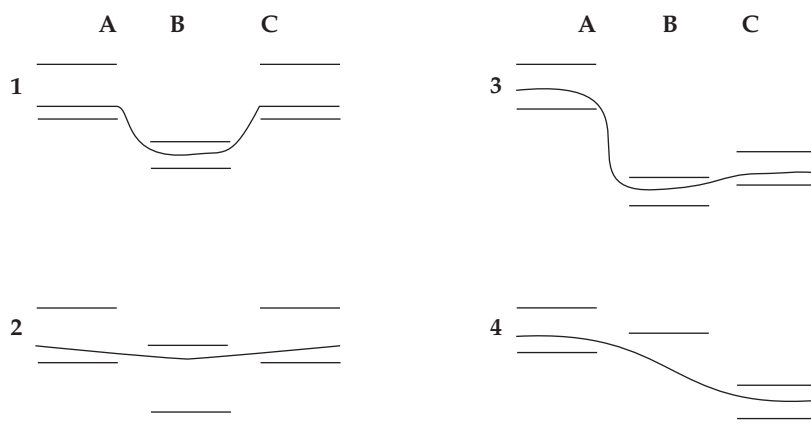
## 2.4 *The window model of coarticulation*

Keating (1985, 1988a, 1988b, 1990) formulated an articulatory model which, in the author's opinion, can account for the continuous changes in space and time observed in speech as well as for interlanguage differences in coarticulation. Keating agrees that phonological rules cannot account for the graded nature of coarticulation, but she contests the assumption that such graded variations are to be ascribed to the speech production mechanism (Keating, 1985). Her proposal is that all graded spatial and temporal contextual variations be accounted for by the phonetic rules of the grammar.

**2.4.1 The windows** Input to the window model is the phonological representation in terms of binary features. For a given articulatory or acoustic dimension, a feature value is associated with a range of values called a window. Specified features are associated with narrow windows and allow for little contextual variation; unspecified features are associated with wide windows and allow for large contextual variation. Windows are connected by interpolation functions called "paths" or contours. Paths should represent the articulatory or acoustic variations over time in a specific context (see Figure 9.7 showing some selected sequences of windows and contours).

Wide windows specify very little about a segment. On this crucial point, Boyce et al. (1991) argue that, if supposedly unspecified segments are associated in production with characteristic articulatory positions, it becomes hard to reconcile the demonstration of any kind of target with the notion of underspecification. The authors propose instead that phonologically unspecified features can influence speech production in another way: they may be associated with cross-speaker variability (as shown by their lip protrusion data during unspecified consonants) and with cross-dialectal variability.

**2.4.2 Cross-language differences** According to Keating, interlanguage differences in coarticulation may originate from phonology or from phonetics. Phonological



**Figure 9.7** Windows and paths modeling articulator movements in three-segment sequences (selected from Keating, 1988a). The effects of narrow vs. wide windows on the interpolation contours can be observed in both the symmetric (1, 2) and the asymmetric (3, 4) sequences.

differences occur when, for a given feature, phonological assimilatory rules operate in one language and not in another. Phonetic differences are due to a different phonetic interpretation of a feature left unspecified. Speech analysis will help determine which differences are phonological and which are phonetic.

In a study of nasalization in English using airflow measurements, Cohn (1993) compared nasal flow contours in nasalized vowels in English with those of nasal vowels in French and of nasalized vowels in Sundanese. Vowel nasality is phonological in French and described as the output of a phonological spreading rule in Sundanese. Cohn found that in the nasalized vowels of Sundanese the flow patterns have plateau-like shapes very similar to the French patterns; in nasalized English vowels, instead, the shapes of the contours describe smooth trajectories from the [–nasal] to the [+nasal] adjacent segments. The categorical versus gradient quality of nasalization in Sundanese versus English indicates that nasalization is the output of phonological assimilatory rules in the former language and results from phonetic interpolation rules in the latter.

Manuel (1987) disagrees with Keating's tenet that all phonetic changes have to be accounted for by grammatical rules simply because they are not universal. Referring to interlanguage differences in V-to-V coarticulation, Manuel proposes that language-particular behavior, apparently arbitrary, can itself be deduced from the interaction between universal characteristics of the motor system and language-specific phonological facts such as the inventory and distribution of vowel phonemes. Her hypothesis is that V-to-V coarticulation is regulated in each language by the requirement that the perceptual contrast among vowels be preserved, i.e., by output constraints, which can be strict in some languages and loose in others. There ought to be more coarticulatory variations in languages with smaller vowel

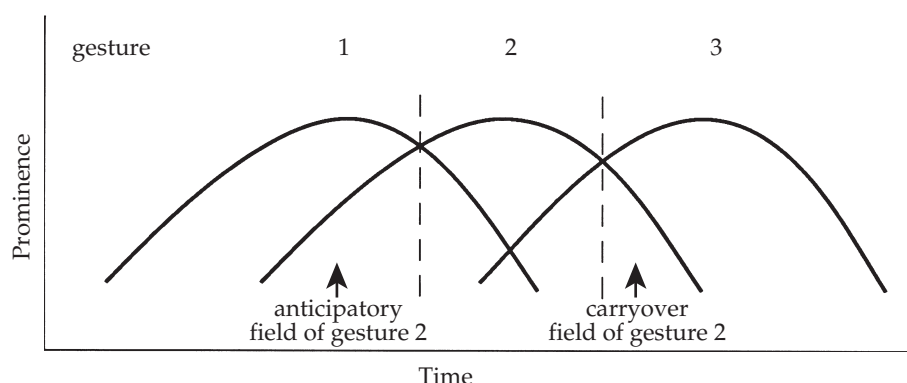
inventories, where there is less possibility of confusion, than in languages with a larger number of vowels, where coarticulation may lead to articulatory/acoustic overlap of adjacent vowel spaces. This hypothesis was tested by comparing languages with different vowel inventories (Manuel & Krakow, 1984; Manuel, 1987, 1990). Results of these studies support the output constraints hypothesis. Thus, if the output constraints of a given language are related to its inventory size and to the distribution of vowels in the articulatory/acoustic space, then no particular language-specific phonetic rules are needed since different degrees of coarticulation across languages can be predictable to some extent.

## 2.5 *Coarticulation as coproduction*

The coproduction theory has been elaborated through collaborative work of psychologists and linguists, starting from Fowler (1977, 1980, 1985), Fowler et al. (1980) and Bell-Berti and Harris (1981). In conjunction with the new theory, Kelso et al. (1986), Saltzman and Munhall (1989) and Saltzman (1991) have developed a computational model, the task-dynamic model, whose aim is to account for the kinematics of articulators in speech. Input to the model are the phonetic gestures, the dynamically defined units of gestural phonology, proposed as an alternative to features by Browman and Goldstein (1986, 1989, 1990a, 1990b, 1992).

The present account centers on four topics: the nature of phonological units, coarticulation resistance, anticipatory labial and velar coarticulation, and the DAC model of coarticulation.

**2.5.1 The dynamic nature of phonological units** The central point of Fowler's criticism of feature-based theories (Fowler, 1977, 1980) is the dichotomy between the abstract, discrete, and timeless units posited at the level of language knowledge, and the physical, continuous, and context-dependent articulatory movements at the level of performance. In other words, she contests the assumption that what speakers know about the phonological categories of their language is substantially different from the units they use when they speak. According to Fowler, all current accounts of speech production need a translation process between the abstract and the physical domain: the speech plan supplies the spatial targets to be reached and a central clock specifies when the articulators have to move to the targets ("The articulator movements are excluded from the domain of the plan except as it is implied by the different successive articulatory positions"; Fowler, 1977, p. 99). An alternative proposal that overcomes the dichotomy between linguistic and production units and gets rid of a time program separated from the plan, is to modify the phonological units of the plan. The plan must specify the act to be executed, not only describe "an abstract summary of its significance" (Fowler et al., 1980, p. 381). The production units, the articulatory gestures, must be planned actions serially ordered, specified dynamically and context-free. It is their specification in terms of dynamic parameters (such as force and stiffness) that automatically determines the kinematics of speech movements. Gestures have their own intrinsic temporal structure, which allows them to overlap in time when executed, and



**Figure 9.8** Representation of a sequence of three overlapping gestures. (From Fowler & Saltzman, 1993)

the degree of gestural overlap is controlled at the plan level. So gestures are not altered by adjacent gestures but coproduced with them. Figure 9.8 taken from Fowler and Saltzman (1993) illustrates the coproduction of articulatory gestures.

The activation of a gesture increases and decreases smoothly in time, and so does its influence on the vocal tract shape. In the figure, the vertical lines delimit a temporal interval (possibly corresponding to an acoustic segment) during which gesture 2 is maximally prominent, i.e., it has maximal influence on the vocal tract shape, while the overlapping gestures 1 and 3 have a weaker influence. The influence of gesture 2 is clearly less before and after this interval during its initiation and relaxation period, respectively.

The view of gestures as intervals of activation gradually waxing and waning in time, echoes the early insight by Joos (1948) who proposed the “innervation wave theory” to account for coarticulation (section 2.1 above).

**2.5.2 Coarticulation resistance** Articulatory gestures are implemented by coordinative structures, i.e., by transient functional dependencies among the articulators that contribute to a gesture. These constraints are established to ensure invariance of the phonetic goal. For instance, upper lip, lower lip, and jaw are functionally linked in the production of bilabial closures, so that one will automatically compensate for a decreased contribution of another due to perturbation or coarticulatory variations (see Löfqvist, this volume).

How are coarticulatory variations accounted for within the gestural framework? According to Fowler and Saltzman (1993), variations induced by coproduction depend on the degree to which the gestures share articulators, i.e., on the degree of spatial overlap. When subsequent gestures share only one articulator, such as the jaw in /VbV/ sequences, the effects of gestural interference will be irrelevant, and temporal overlap between vocalic and consonantal gestures will take place

with minimal spatial perturbations. The highest degree of spatial perturbation occurs when two overlapping gestures share the articulators directly involved in the production of gestural goals, and impose competing demands on them. Browman and Goldstein (1989) and Saltzman and Munhall (1989) propose that the phasing of gestures may be context-free and that the output of a gestural conflict may be simply a blend of the influence of the overlapping gestures. According to Fowler and Saltzman (1993), the outcome of gestural blending depends on the degree of "blending strength" associated with the overlapping gestures: "stronger" gestures tend to suppress the influence of "weaker" gestures, while the blending of gestures of similar strength will result in an averaging of the two influences. In agreement with experimental findings (Bladon & Nolan, 1977; Recasens, 1984b; Farnetani & Recasens, 1993; Fowler & Brancazio, 2000), Fowler and Saltzman's account of coarticulatory resistance implies that gestures with a high degree of blending strength resist interference from other gestures, and at the same time themselves induce strong coarticulatory effects. On this account, the highest degree of blending strength appears to be associated with consonants requiring extreme constrictions and/or placing strong constraints on articulator movements, while a moderate degree of blending strength appears to be associated with vowels. A compatible proposal may be found in Lindblom (1983), according to which coarticulatory adaptability, maximal in vowels and minimal in lingual fricatives, varies as a function of the phonotactically based sonority categories.

The coproduction account of coordination and coarticulation also implies that speakers do not need to perform a continuous feedforward control of the acoustic output and consequent articulatory adjustments. Likewise, cross-language differences do not result from online control of the output. Languages may differ in degree of coarticulation in relation to their inventories, but these differences are consequences of the different gestural set-up, i.e., the parameters that specify the dynamics of gestures and their overlap, which are learned by speakers of different languages during speech development.

**2.5.3 Anticipatory extent of labial and velar coarticulation** According to the coproduction theory, articulatory gestures have their own intrinsic duration. Hence, the temporal extent of anticipatory coarticulation must be constant for a given gesture. Compatibly, Bell-Berti and Harris (1979, 1981, 1982), proposed the "frame" or time-locked model of anticipatory coarticulation on the basis of experimental data on lip rounding and velar lowering. The model states that the onset of an articulator movement is independent of the preceding phone string length and occurs at a fixed time before the acoustic onset of the segment with which it is associated.

Findings reported in other studies, however, are more consistent with the look-ahead model (section 2.2.2) and reveal that the onset of anticipatory lip rounding or anticipatory velar lowering is not fixed but extends as a function of the number of neutral segments preceding the influencing segment (see Daniloff & Moll, 1968 and Sussman & Westbury, 1981 for lip rounding coarticulation, and Moll & Daniloff, 1971 for velar coarticulation). Yet, other results on velar coarticulation in Japanese

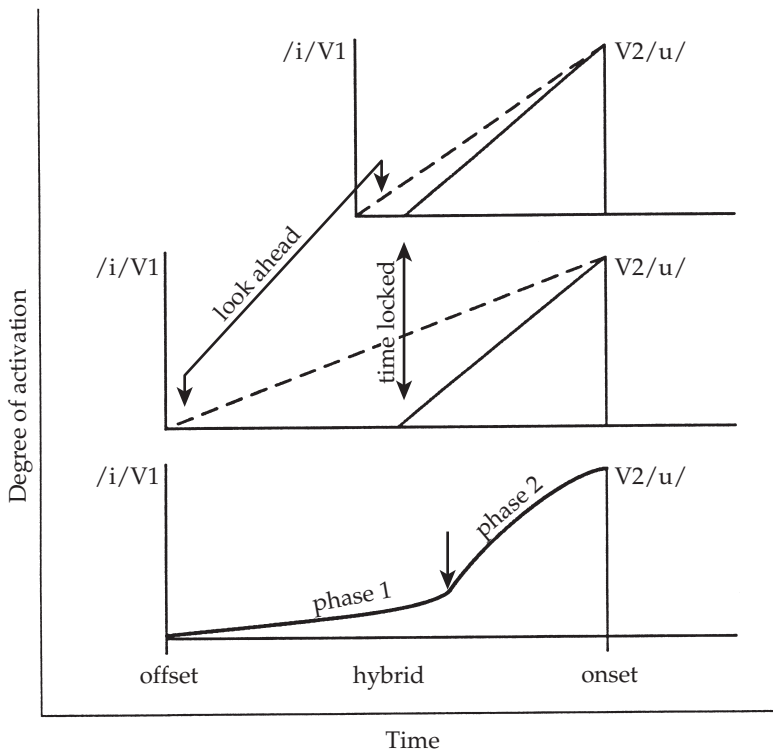


(Ushijima & Sawashima, 1972; Ushijima & Hirose, 1974) and in French (Benguerel et al., 1977a, 1977b) indicate that velar lowering for a nasal consonant does not start earlier in sequences of three than of two oral vowels preceding the nasal.

An important finding of Benguerel et al. (1977a, 1977b), apparently disregarded in the literature, was the distinction between velar lowering associated with the oral segments preceding the nasal, and a subsequent more rapid velar lowering for the nasal which causes the opening of the velar port. Bladon and Al-Bamerni (1982) reported similar findings on velar coarticulation in CV<sub>n</sub>N sequences in English: speakers seem to use two production strategies, either a single velar opening gesture, or a two-stage gesture whose onset is aligned with the first oral vowel and whose higher velocity stage is coordinated with the nasal consonant. Perkell and Cohen (1986) and Perkell (1990) were the first to observe two-stage patterns in lip rounding movements, which converge with Bladon and Al-Bamerni's observations: in /iC<sub>n</sub>u/ utterances there was a gradual onset of lip protrusion linked to the offset of /i/, followed by an increase in velocity during the consonants and an additional protrusion motion closely linked with /u/ and quite invariant. The authors interpreted the composite movements of the two-stage patterns as a mixed coarticulation strategy, and proposed a third model of anticipatory coarticulation, the hybrid model. According to this model, the early onset of the protrusion movement would reflect a look-ahead strategy, while the rapid increase in protrusion at a fixed interval before the rounded vowel would reflect a time-locked strategy. Figure 9.9 compares the three models of anticipatory coarticulation. Perkell's data on three English subjects indicated that two of the three subjects used the two-stage pattern and, therefore, were consistent with the hybrid model (Perkell, 1990).

Boyce et al. (1990) argue that many of the conflicting results on the extent of anticipatory coarticulation stem from the assumption that phonologically unspecified segments are also articulatorily neutral (see section 2.2.3): a number of studies have attributed the onset of lip rounding or velar lowering to anticipatory coarticulation without testing first whether or not the phonologically neutral contextual phonetic segments had specific lip or velar target positions. Data on velar lowering for vowels in nasal and oral contexts reported by Bell-Berti and Krakow (1991) show that the early onset of velar lowering in two-stage patterns is associated with the characteristic velar positions of the oral vowels, while the second stage is associated with the production of the nasal consonant; therefore, the two-stage patterns of interest do not reflect a mixture of coarticulation strategies but simply a vocalic gesture followed by a consonantal gesture. Moreover, the study shows that the patterns of velar movement are not random but depend on speech rate and on the number of vowels in the string, e.g., the two-movement pattern prevails in longer versus shorter utterances.

In a subsequent study on four American speakers, Perkell and Matthies (1992) tested whether the onset of the initial phase of lip protrusion in /iC<sub>n</sub>u/ utterances is related to consonant-specific protrusion targets as proposed by Boyce et al. (1990), and whether the second phase starting at the maximum acceleration event is indeed stable as predicted by the hybrid and coproduction models, or else is



**Figure 9.9** Schematic representation of the three models of anticipatory lip-rounding coarticulation proposed by Perkell (see text for description).

itself affected by the number of consonants. In agreement with Boyce et al. (1990), the movement patterns in the control  $/iC_n i/$  utterances showed consonant-related protrusion gestures (especially for  $/s/$ ), while those in the  $/iC_n u/$  utterances exhibited earlier labial activity when the first consonant in the string is  $/s/$ . Both data confirm that the consonant contributes to the onset of lip movement. The analysis of the second-phase movement, i.e., of the  $/u/$ -related component of lip protrusion, revealed that the interval between the acceleration peak and the onset of  $/u/$  tended to vary as a function of consonant duration for three subjects (although the correlations were very low, with  $R^2$  ranging from 0.057 to 0.35). According to the authors, the timing and the kinematics of this gesture reflect the simultaneous expression of competing constraints, that of using the same kinematics (as predicted by the time-locked model), and that of starting the protrusion gesture for  $/u/$  when it is permitted by the relaxation of the retraction gesture for  $/i/$  (as predicted by the look-ahead model). The variability between and within subjects would reflect the degree to which such constraints are implemented. Also, according to data for French (Abry & Lallouache, 1995), the lip-protrusion



movement measured from acceleration maximum to protrusion maximum varies in duration as a function of the consonant interval; however, in disagreement with the look-ahead model, its duration does not decrease from the /iC<sub>1</sub>y/ utterance type to the utterance /iy/; in other words, lip protrusion can expand in time but cannot be compressed.

The possibility that the slow onset of the lip-protrusion movement occurring around the offset of /i/ may reflect a passive movement due to the relaxation of the /i/ retraction gesture rather than an active look-ahead mechanism has not yet been explored. Sussman and Westbury (1981) did observe for /iC<sub>n</sub>u/ sequences that lip protrusion started before the onset of orbicularis oris activity, and suggested that this movement might be the passive result of the cessation of risorius activity and simply reflects a return of the lips to the neutral position.

As it might be expected, cross-language studies indicate that anticipatory coarticulation varies in timing and amplitude across languages. Clumeck (1976) observed that the timing and amplitude of velar lowering varies across the six languages analyzed; Lubker and Gay (1982) showed that anticipatory lip protrusion is much more extensive in Swedish than in English; an investigation on lip rounding in English and Turkish conducted by Boyce (1990) indicated that, while the English patterns were consistent with the coproduction model, the plateau-like protrusion curves of Turkish rendered lip-rounding coarticulation a phonological process.

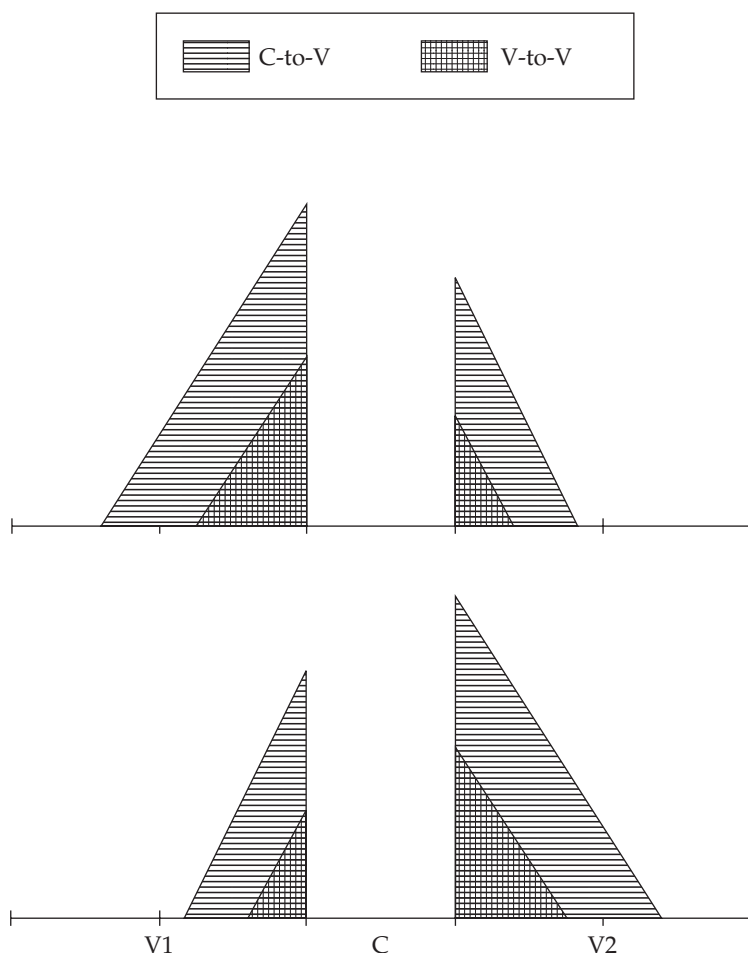
**2.5.4 The DAC model of coarticulation** The degree of articulatory constraint or DAC model of coarticulation has been proposed by Recasens, mostly with data for the Catalan language, in order to deal with the complexity of lingual coarticulatory effects in speech (Recasens et al., 1997; Recasens, 2002). It claims that the size, temporal extent, and direction of lingual coarticulation are conditioned by the severity of the requirements imposed on the tongue for the production of vowels and consonants. Vowels and consonants are assigned specific DAC values. Thus, front vowels are more constrained than low and back rounded vowels in line with the biomechanical properties involved in displacing the tongue dorsum upwards and frontwards, and the least constrained vowel is schwa since it has no clear articulatory target. Differences in degree of constraint are available for consonants as well: the DAC value is highest for consonants requiring much articulatory precision in the performance of frication for lingual fricatives, trilling for the alveolar trill, and tongue predorsum lowering and tongue postdorsum retraction for dark /ɫ/; labials are least constrained since the tongue body does not intervene essentially in their production.

An increase in degree of articulatory constraint causes an increase in coarticulatory resistance and coarticulatory dominance, i.e., in the strength of the coarticulatory effects from and onto the adjacent segments, respectively. Thus, in comparison to the less constrained alveolar /n/, the alveopalatal /ɲ/ is less coarticulation-sensitive to tongue-dorsum lowering effects from /a/ and exerts more prominent tongue-dorsum raising effects on this vowel. On the other hand, alveopalatals are subject to depalatalization by more constrained consonants

involving tongue-dorsum lowering and retraction (lingual fricatives, the alveolar trill, dark /l/), while palatalization of the latter consonants by alveolopalatals is less prone to occur (Recasens & Pallarès, 2001).

A relevant aspect of the DAC model concerns patterns of coarticulatory direction. It shows that vowels and consonants usually favor a particular coarticulatory direction over another, i.e., anticipation or carryover, based on the displacement and temporal characteristics of the lingual gestures involved. Thus, dark /l/ favors anticipation (the tongue lowers and backs in anticipation of the tongue-tip raising gesture for this consonant), while alveolopalatals such as /ɲ/ may favor the carryover component over the anticipatory component (which is in line with the articulatory configuration for this consonant being /n/-like at closure onset and /j/-like at closure release). When the complex coarticulatory interactions in VCV sequences are taken into consideration, vowel effects turn out to be affected by consonantal effects at the temporal site where the two coarticulation types conflict with each other. Therefore, the prominence of the vowel-dependent anticipatory effects decreases with an increase in the degree of consonant-dependent carryover coarticulation, while the strength of the vowel-dependent carryover effects varies inversely with the salience of the consonant-dependent anticipatory component. Consequently, dark /l/ allows more anticipatory than carryover tongue-dorsum raising effects from /i/ (as shown in the top graph of Figure 9.10), and the alveolopalatal /ɲ/ more carryover than anticipatory tongue-dorsum lowering effects from /a/ (as shown in the bottom graph of the figure). Trends in vowel-dependent coarticulatory direction may also be predicted for VCV sequences with consonants showing less clear patterns of C-to-V coarticulatory direction provided that sufficient attention is paid to their manner requirements and tongue-body configuration characteristics.

Data on the temporal extent of coarticulation have led to proposals by the DAC model about the role of planning and mechanical factors in speech production. In contrast with previous accounts (see section 2.5.3), they show that vowel anticipatory effects in tongue contact and displacement in VCV or longer VCVCV sequences are not planned to start invariably at the same moment in time. Instead, they begin earlier when the immediately preceding consonant and/or the trans-consonantal vowel are relatively unconstrained than when they are highly constrained, e.g., when V1 is /a/ and the consonant is labial or alveolar than when V1 is /i/ and the consonant is alveolopalatal. Among highly constrained consonants, as pointed out above, those that exert less C-to-V carryover allow more vowel anticipation than those that exert more prominent C-to-V carryover effects, both during the consonant and the preceding vowel. This is not to say, however, that anticipatory effects do not involve any planning at all: anticipatory effects turn out to be more fixed than carryover effects, though influenced to some extent by the articulatory requirements for the contextual segments. Carryover effects, on the other hand, are more variable because they are conditioned by inertia and by the biomechanical requirements for the contextual segments, and may be particularly long if resulting from articulatory overshoot (e.g., in the sequence /iɲV/ where /i/ and /ɲ/ reinforce each other).



**Figure 9.10** Schematic representation of C-to-V and V-to-V effects in VCV sequences with dark /ɪ/ (top) and alveolopalatal /ɲ/ (bottom). (From D. Recasens and M. D. Pallarès (1998) "An electropalatographic and acoustic study of temporal coarticulation for Catalan dark /ɪ/ and German clear /ɪ/," *Phonetica*, 55, 53–79. Reproduced with permission from S. Karger AG, Basel.)

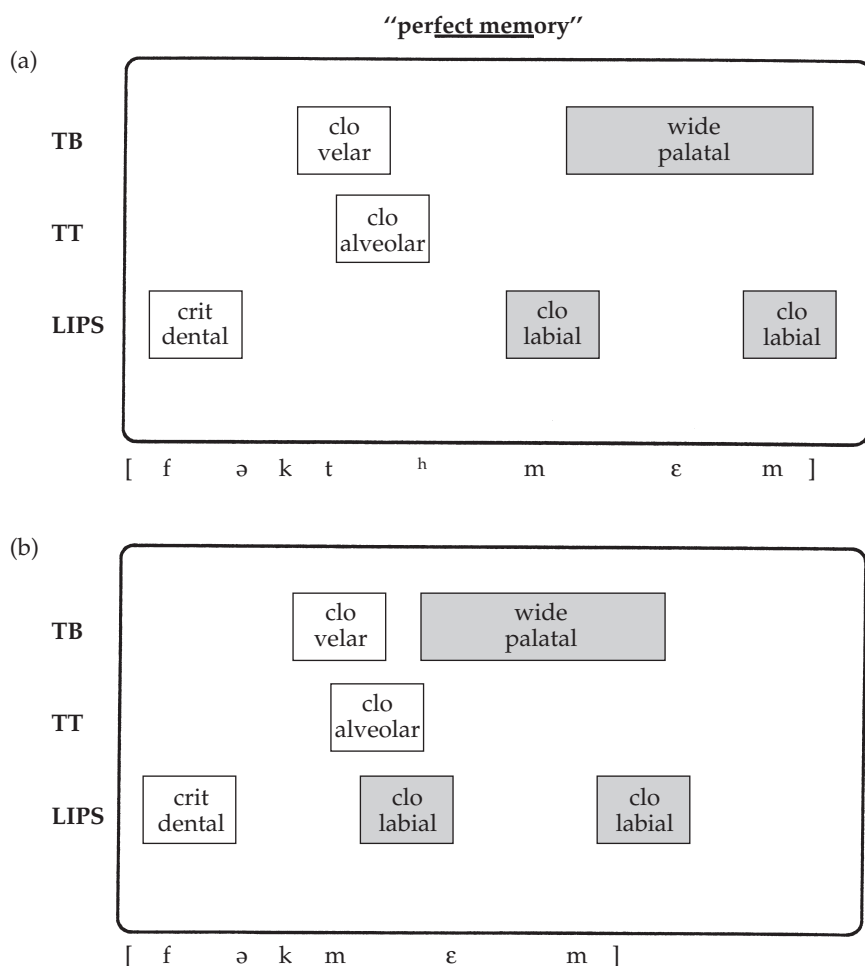
The DAC model has confirmed Fowler and Saltzman's notion that gestures specified for competing demands prevail over or are overridden by other gestures depending on gestural strength (see section 2.5.2). Indeed, differences in DAC value between the two consonants in a consonant cluster may account for changes in place of articulation in CC sequences composed of dentals, alveolars, and alveolopalatals produced with a tongue-front articulator (Recasens, 2006). Three scenarios may take place here: regressive assimilation in unconstrained + constrained consonant sequences, by which C1 adapts completely to the C2 place of

articulation during its entire duration; blending in unconstrained + unconstrained consonant sequences through the addition of the closure extent for the two consonants; two separate place of articulation targets for C1 and C2 and possible C1-to-C2 carryover coarticulation effects in constrained + unconstrained sequences. Clear /l/ and dark /l/ do not undergo some or any of these three processes, which reveals that laterality may also contribute significantly to an increase in articulatory constraint. The model may also explain syllable-position-dependent differences in the same consonant cluster structures just referred to (Recasens, 2004): when not subject to the coarticulatory influence of other highly constrained consonants, more unconstrained consonants show less closure or constriction fronting and less dorsopalatal contact syllable-finally than syllable-initially; highly constrained consonants, on the other hand, do not exhibit such syllable-position-dependent articulatory effects.

## 2.6 Connected speech processes

**2.6.1 Articulatory model** According to Browman and Goldstein (1990b, 1992), gestural phonology provides an explanatory and unifying account of apparently unrelated speech processes (coarticulation, allophonic variations, alternations) requiring a number of separate phonological rules in featural phonology. Here the phonological structure of an utterance is modeled as a set of overlapping gestures specified on different tiers (see Figure 9.11 for the vocal tract variables). Gradient variations in overlap, or quantitative variations in gestural parameters, can account for a large number of allophonic variations as a function of stress and position, as well as for the alternations observed in connected speech. Connected speech processes such as assimilations, deletions, and reductions or weakenings can be accounted for by an increase in gestural overlap and a decrease in gestural amplitude. In casual rapid speech, subsequent consonantal gestures can so far overlap as to hide each other when they occur on different tiers, or to completely blend their characteristics when they occur on the same tier. Hiding gives rise to perceived deletions and/or assimilations, while blending gives rise to perceived assimilations. For example, the deletion of /t/ in a rapid execution of the utterance “perfect memory” is only apparent; X-ray trajectories reveal the presence of the /t/ gesture, overlapped by the following /m/ gesture. Figure 9.11 shows a schematic gestural representation of part of the utterance “perfect memory” spoken (a) in isolation and (b) within a fluent phrase.

In the figure the extent of each box represents the duration (or activation interval) of a gesture. It can be seen that within each word articulatory gestures always overlap, but in version (b) the labial closure for the initial /m/ of word 2 overlaps and hides the alveolar gesture for the final /t/ of word 1. According to the authors, hidden gestures may be extremely reduced in magnitude or completely deleted. As pointed out by Browman and Goldstein (1990b, p. 366): “Even deletion, however, can be seen as an extreme reduction, and thus as an endpoint in a continuum of gestural reduction, leaving the underlying representation unchanged.”



**Figure 9.11** Representation of the phonological structure of the utterance “perfect memory” produced in isolation (a) and in continuous speech (b). Vocal tract variables are from top: tongue body, tongue tip, and lips. (From Browman & Goldstein, 1989)

An example of within-tiers blending is the palatalization of /s/ followed by a palatal in the utterance “this shop”: the articulatory analysis should show a smooth transition between the first and the second consonant, not the substitution of the first by the second. Finally, in CVCV utterances with unstressed schwa as first vowel, the gestures on the consonantal tier can overlap the schwa gesture so far as to completely hide it, giving the impression of deletion of the unstressed syllable.

**2.6.2 Experimental data** Consistently with both gestural theory and Lindblom’s hyper-/hypo-speech account, an increase in coarticulation and reduction in rapid,

fluent speech has been found to hold in a number of acoustic and articulatory studies (see section 2.3.2).

The proposition that position-dependent allophonic variations proceed continuously rather than categorically is supported by experimental data on contrast neutralization. This approach differs from generative phonology which accounts for contrast neutralization through rules that delete the feature(s) responsible for the contrast. An acoustic-perceptual experiment on vowel contrast neutralization in devoiced syllables in Japanese shows that contrast is not completely neutralized (Beckman & Shoji, 1984): listeners are able to recover the underlying vowels /i/ and /u/, possibly from coarticulatory information present in the preceding consonant. In an acoustic-perceptual study on neutralization of the voicing contrast in word-final obstruents in German, Port and O'Dell (1985) found that voicing is not completely neutralized and that listeners are able to distinguish the voiced from the voiceless consonants with better-than-chance accuracy.

Also the majority of English data on alveolar-velar place assimilation in connected speech reported so far is consistent with the proposition that the nature of the segmental adaptive changes is gradient. EPG studies on VC-CV sequences, where C1 is an alveolar stop (Kerswill & Wright, 1989; Wright & Kerswill, 1989; Nolan, 1992), show an intermediate stage between absence of assimilation and complete assimilation, which the authors refer to as residual alveolar gesture. It is also shown that the occurrences of partial and complete assimilations increase from careful/slow speech to normal/fast speech. Most interestingly, the rate of correct identification of C1 decreases, as expected, from unassimilated to assimilated alveolars, but never falls to zero, suggesting that also in the cases of apparently complete assimilation where lingual alveolar contact is absent, listeners can make use of some residual cues to the place distinction. The data are in agreement with the hypothesis that in English the assimilation of alveolars to velars is a continuous process. This is confirmed by recent research on alveolar nasal + velar stop clusters (Hardcastle, 1994).

Other data (Barry, 1991; Nolan et al., 1993) challenge some of the assumptions of gestural phonology. The cross-language study by Barry (1991) on English and Russian alveolar-velar clusters confirms that assimilation in English is a graded process. In Russian, instead, assimilation never occurs when C1 is an oral stop; when C1 is a nasal, assimilation may be continuous or categorical depending on syllabic structure. Data on /s#f/ sequences reported by Nolan et al. (1993) do show intermediate articulations between two-gesture and one-gesture patterns, as predicted by gestural phonology. Accordingly, the one-gesture or static patterns should reflect complete spatio-temporal overlap, i.e. they should show a blending of the /s/ and /f/ influences and a duration comparable to that of a single consonant. Contrary to this hypothesis, preliminary results indicate that the static patterns have the spatial characteristic of a typical /f/, and are 16 percent longer than an initial /f/. Recent EPG and durational data on Italian clusters with C1 = /n/ followed by an oral C1 differing in place and manner of articulation suggest that both categorical and continuous processes may coexist in a language, the occurrence of the ones or the others depending on cluster type and individual

speech style. Moreover, the finding that in Italian the alveolar–velar assimilation in /nk/ clusters is always categorical indicates, in agreement with Barry (1991), that the assimilatory process for the same cluster type may differ qualitatively across languages (Farnetani & Busà, 1994). Also, as pointed out by the DAC model (section 2.5.4), different predictions appear to be needed for sequences of consonants produced with close and distant primary tongue articulators; in particular, complete place adaptation appears to be an efficient strategy for the implementation of two consecutive consonants produced with the same or a close articulator and differing in manner of articulation requirements.

### 3 Summary

This excursus on the problem of contextual variability shows, on one hand, the incredible complexity of the speech production mechanism, which renders the task of understanding its underlying control principles so difficult. It shows, on the other hand, the enormous theoretical and experimental ongoing progress, as reflected in continuously evolving and improving models, and in increasingly rigorous and sophisticated research methodologies.

We started with the questions of the origin, function, and control of coarticulation. At the moment there is no single answer to these questions. For generative phonology, assimilations, connected speech processes, and coarticulation are different steps linking the domain of competence with that of performance, with no bottom-up influences from the physical to the cognitive structure of the language. For both the theory of “adaptive variability” and the theory of gestural phonology the origin of coarticulation lies in speech (in its plasticity and adaptability for the former, in its intrinsic organization in time for the latter). Both theories assume that the nature of speech production itself is at the root of linguistic morpho-phonological rules, which are viewed as adaptations of language to speech processes, sometimes eventuating in historical sound changes. However, there is a discrepancy between the two theories on the primacy of production vs. perception in the control of speech variability. Gestural phonology considers acoustics/perception as the effect of speech production, whilst Lindblom’s theory of “adaptive variability” sees acoustics/perception as the final goal of production, hence perception itself shapes production.

Two general control principles for speech variability have been repeatedly advocated: economy (by Lindblom and by Keating) and output constraints (advocated by Lindblom for the preservation of perceptual contrast across styles within a language and extended by Manuel to account for interlanguage differences in coarticulation).

If we confront the various articulatory models with experimental data, it seems that the overall results on coarticulation resistance are more consistent with the gestural model than with other models, although certain patterns of coarticulation resistance could be better explained if aerodynamic/acoustic constraints, in addition to articulatory constraints, were taken into account (Sussman & Westbury, 1981;



Engstrand, 1983). The challenging hypothesis of gestural phonology that connected speech processes are not substantially different from coarticulation processes (i.e., are continuous and do not imply qualitative changes in the categorical underlying units) is supported by a large number of experimental results. However, recent data, based on both spatial and temporal parameters, indicate that assimilation can also be a rule-governed categorical process.

As for anticipatory coarticulation, no model in its present version can account for the diverse results within and across languages: the review shows that articulatory structures and languages differ both quantitatively and qualitatively in the way they implement this process. Lingual coarticulation appears to be subject to a more restricted set of mechanisms than labial and velar coarticulation which calls for models (such as the DAC model) relying on detailed information about the articulatory constraints involved in the production of specific vowel and consonant types. Languages differ considerably in the anticipatory coarticulation strategies for lips and velum. Thus, English and Swedish seem to differ quantitatively in lip-rounding anticipation (Lubker & Gay, 1982), while the plateau-patterns observed in some languages (Boyce, 1990; Cohn, 1993) suggest that the process is phonological in some languages and phonetic in others. Most intriguing in the data on anticipatory coarticulation are the discrepancies among the results for the same language, as those on vowel nasalization in American English (cf. Moll & Daniloff, 1971 vs. Bell-Berti, 1980, vs. Solé & Ohala, 1991). Such discrepancies might be due to different experimental techniques, or the different speech material may itself have conditioned the speaking style or rate and hence the coarticulatory patterns. The discrepancies might also reveal actual regional variants, suggesting ongoing phonetic changes, yet to be fully explored.

## NOTE

We thank Michael Studdert-Kennedy for his extensive comments on the form and substance of the present work, and Bjorn Lindblom and Bill Hardcastle for their valuable suggestions. This research was supported in part by projects ESPRIT/ ACCOR, WG 7098 from the European Union, and HUM2006-03743 from the Spanish Ministry of Education.

## REFERENCES

- |   |  |
|---|--|
| <p>Abry, C. &amp; Lallouache, M. T. (1995) Le MEM: un modèle d'anticipation paramétrable par locuteur. Données sur l'arrondissement en français. <i>Bulletin du Laboratoire de la Communication Parlée</i>, Centre National de la Recherche Scientifique, Grenoble, 3, 85–99.</p> | <p>Barry, M. (1991) Temporal modelling of gestures in articulatory assimilation. <i>Proceedings of the 12th International Congress of Phonetic Sciences</i> (Aix-en-Provence), 4, 14–17.</p> <p>Beckman, M. &amp; Shoji, A. (1984) Spectral and perceptual evidence for CV</p> |
|---|--|



- coarticulation in devoiced /si/ and /syu/ in Japanese. *Phonetica*, 41, 61–71.
- Bell-Berti, F. (1980) Velopharyngeal function: A spatial-temporal model. In N. J. Lass (ed.), *Speech and Language: Advances in Basic Research and Practice* (pp. 291–316). New York: Academic Press.
- Bell-Berti, F. & Harris, K. (1974) More on the motor organization of speech gestures. *Haskins Laboratories Status Report on Speech Research*, 37/38, 73–7.
- Bell-Berti, F. & Harris, K. (1979) Anticipatory coarticulations: Some implications from a study of lip rounding. *Journal of the Acoustical Society of America*, 65, 1268–70.
- 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, 449–54.
- Bell-Berti, F. & Krakow, R. (1991) Anticipatory velar lowering: A coproduction account. *Journal of the Acoustical Society of America*, 90, 112–23.
- Benguerel, A. P. & Cowan, H. (1974) Coarticulation of upper lip protrusion in French. *Phonetica*, 30, 41–55.
- Benguerel, A. P., Hirose, H., Sawashima, M., & Ushijima, T. (1977a) Velar coarticulation in French: A fiberoptic study. *Journal of Phonetics*, 5, 149–58.
- Benguerel, A. P., Hirose, H., Sawashima, M., & Ushijima, T. (1977b) Velar coarticulation in French: An electromyographic study. *Journal of Phonetics*, 5, 159–67.
- Bladon, A. & Al-Bamerni, A. (1976) Coarticulation resistance in English /l/. *Journal of Phonetics*, 4, 137–50.
- Bladon, A. & Al-Bamerni, A. (1982) One-stage and two-stage temporal patterns of velar coarticulation. *Journal of the Acoustical Society of America*, 72, S104.
- Bladon, A. & Nolan, F. (1977) A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics*, 5, 185–93.
- Boyce, S. (1990) Coarticulatory organization for lip rounding in Turkish and English. *Journal of the Acoustical Society of America*, 88, 2584–95.
- Boyce, S., Krakow, R., & Bell-Berti, F. (1991) Phonological underspecification and speech motor organization. *Phonology*, 8, 219–36.
- Boyce, S., Krakow, R., Bell-Berti, F., & Gelfer, C. (1990) Converging sources of evidence for dissecting articulatory movements into core gestures. *Journal of Phonetics*, 18, 173–88.
- Browman, C. P. & Goldstein, L. M. (1986) Towards an articulatory phonology. In C. Ewan & J. Anderson (eds.), *Phonology Yearbook 3* (pp. 219–52). Cambridge: Cambridge University Press.
- Browman, C. P. & Goldstein, L. M. (1989) Articulatory gestures as phonological units. *Phonology*, 6, 201–51.
- Browman, C. P. & Goldstein, L. M. (1990a) Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, 18, 299–320.
- Browman, C. P. & Goldstein, L. M. (1990b) Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech* (pp. 341–76). Cambridge: Cambridge University Press.
- Browman, C. P. & Goldstein, L. M. (1992) Articulatory phonology: An overview. *Phonetica*, 49, 155–80.
- Butcher, A. & Weiher, E. (1976) An electropalatographic investigation of coarticulation in VCV sequences. *Journal of Phonetics*, 4, 59–74.
- Chennoukh, S., Carré, R., & Lindblom, B. (1997) Locus equation in the light of articulatory modelling. *Journal of the Acoustical Society of America*, 102, 2380–8.
- Choi, J. D. & Keating, P. (1991) Vowel-to-vowel coarticulation in three Slavic languages. *UCLA Working Papers in Phonetics*, 78, 78–86.

- Chomsky, N. & Halle, M. (1968) *The Sound Pattern of English*. New York: Harper and Row.
- Clumeck, H. (1976) Patterns of soft palate movements in six languages. *Journal of Phonetics*, 4, 337–51.
- Cohn, A. C. (1993) Nasalisation in English: Phonology or phonetics. *Phonology*, 10, 43–81.
- Daniloff, R. & Hammarberg, R. (1973) On defining coarticulation. *Journal of Phonetics*, 1, 239–48.
- Daniloff, R. & Moll, K. (1968) Coarticulation of lip rounding. *Journal of Speech and Hearing Research*, 11, 707–21.
- Duez, D. (1991) Some evidence of second formant locus-nucleus patterns in spontaneous speech in French. *PERILUS XII*, University of Stockholm, 109–26.
- Engstrand, O. (1981) Acoustic constraints or invariant input representation? An experimental study of selected articulatory movements and targets. *Reports from Uppsala University Department of Linguistics*, 7, 67–95.
- Engstrand, O. (1983) Articulatory coordination in selected VCV utterances: A means-end view. *Reports from Uppsala University Department of Linguistics*, 10, 1–145.
- Engstrand, O. (1989) Towards an electropalatographic specification of consonant articulation in Swedish. *PERILUS X*, University of Stockholm, 115–56.
- Farnetani, E. (1986) A pilot study of the articulation of /n/ in Italian using electro-palatography and airflow measurements. *15e Journées d'Études sur la Parole*. GALF 23–6.
- Farnetani, E. (1990) V-C-V lingual coarticulation and its spatiotemporal domain. In W. J. Hardcastle & A. Marchal (eds.), *Speech Production and Speech Modelling* (pp. 93–130). Dordrecht: Kluwer.
- Farnetani, E. (1991) Coarticulation and reduction in consonants: Comparing isolated words and continuous speech. *PERILUS XIV*, University of Stockholm, 11–15.
- Farnetani, E. & Busà, M. G. (1994) Italian clusters in continuous speech. *Proceedings of the 1994 International Conference on Spoken Language Processing*, Yokohama, 1, 359–62.
- Farnetani, E. & Recasens, D. (1993) Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian. *Language and Speech*, 36, 279–302.
- Farnetani, E., Vaggies, K., & Magno-Caldognetto, E. (1985) Coarticulation in Italian /VtV/ sequences: A palatographic study. *Phonetica*, 42, 78–99.
- Fowler, C. A. (1977) *Timing Control in Speech Production*. Bloomington: Indiana University Linguistics Club.
- Fowler, C. A. (1980) Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8, 113–33.
- Fowler, C. A. (1983) Realism and unrealism: A reply. *Journal of Phonetics*, 11, 303–22.
- Fowler, C. A. (1985) Current perspectives on language and speech production: A critical overview. In R. G. Daniloff (ed.), *Speech Science* (pp. 193–278). London: Taylor and Francis.
- Fowler, C. A. & Brancazio, L. (2000) Coarticulation resistance of American English consonants and its effects on transconsonantal vowel-to-vowel coarticulation. *Language and Speech*, 43, 1–41.
- Fowler, C. A., Rubin, P., Remez, R., & Turvey, M. (1980) Implications for speech production of a general theory of action. In B. Butterworth (ed.), *Language Production I: Speech and Talk* (pp. 373–420). London: Academic Press.
- Fowler, C. A. & Saltzman, E. (1993) Coordination and coarticulation in speech production. *Language and Speech*, 36, 171–95.
- Gay, T. (1978) Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America*, 63, 223–30.

- Gimson, A. C. (1970) *An Introduction to the Pronunciation of English*, 2nd edn. London: Edward Arnold.
- Hammarberg, R. (1976) The metaphysics of coarticulation. *Journal of Phonetics*, 4, 353–63.
- Hardcastle, W. (1985) Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication*, 4, 247–63.
- Hardcastle, W. (1994) EPG and acoustic study of some connected speech processes. *Proceedings of the 1994 International Conference on Spoken Language Processing*, Yokohama, 2, 515–18.
- Henke, W. L. (1966) Dynamic articulatory model of speech production using computer simulation. Doctoral Dissertation, MIT.
- Hirose, H. & Gay, T. (1972) The activity of the intrinsic laryngeal muscles in voicing control: An electro-myographic study. *Phonetica*, 25, 140–64.
- Jones, D. (1969) *An Outline of English Phonetics*, 9th edn. Cambridge: Cambridge University Press.
- Joos, M. (1948) Acoustic phonetics. *Language Monographs*, 23 (Supplement to *Language*, 24).
- Keating, P. A. (1985) Universal phonetics and the organization of grammars. In V. Fromkin (ed.), *Phonetic Linguistics: Essays in Honor of Peter Ladefoged* (pp. 115–32). Orlando: Academic Press.
- Keating, P. A. (1988a) Underspecification in phonetics. *Phonology*, 5, 275–92.
- Keating, P. A. (1988b) The window model of coarticulation: Articulatory evidence. *UCLA Working Papers in Phonetics*, 69, 3–29.
- Keating, P. A. (1990) Phonetic representations in a generative grammar. *Journal of Phonetics*, 18, 321–34.
- Kelso, J. A. S., Saltzman, E., & Tuller, B. (1986) The dynamical perspective on speech production: data and theory. *Journal of Phonetics*, 14, 29–59.
- Kerswill, P. & Wright, S. (1989) On the limits of auditory transcription: A sociophonetic approach. *York Papers in Linguistics*, 14, 35–59.
- Kohler, K. J. (1990) Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. J. Hardcastle & A. Marchal (eds.), *Speech Production and Speech Modelling* (pp. 69–92). Dordrecht: Kluwer.
- Kozhevnikov, V. A. & Chistovich, L. A. (1965) *Speech: Articulation and Perception* (trans. US Department of Commerce, Clearing House for Federal Scientific and Technical Information), No. 30, 543. Washington, DC: Joint Publications Research Service.
- Krull, D. (1987) Second formant locus patterns as a measure of consonant–vowel coarticulation. *PERILUS V*, University of Stockholm, 43–61.
- Krull, D. (1989) Second formant locus pattern and consonant–vowel coarticulation in spontaneous speech. Phonetic experimental research at the Institute of Linguistics, *PERILUS X*, University of Stockholm, 87–108.
- Kuehn, D. P. & Moll, K. L. (1976) A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, 4, 303–20.
- Liljencrants, J. & Lindblom, B. (1972) Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 48, 839–62.
- Lindblom, B., (1963) Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35, 1773–81.
- Lindblom, B. (1983) Economy of speech gestures. In P. F. MacNeilage (ed.), *The Production of Speech* (pp. 217–45). New York: Springer Verlag.
- Lindblom, B. (1989) Phonetic invariance and the adaptive nature of speech. In B. A. G. Elsendoorn & H. Bouma (eds.), *Working Models of Human Perception* (pp. 139–73). London: Academic Press.
- Lindblom, B. (1990) Explaining phonetic variation: A sketch of the H & H theory.

- In W. J. Hardcastle & A. Marchal (eds.), *Speech Production and Speech Modelling* (pp. 403–40). Dordrecht: Kluwer.
- Lindblom, B., Pauli, S., & Sundberg, J. (1975) Modeling coarticulation in apical stops. In G. Fant (ed.), *Speech Communication*, vol. 2 (pp. 87–94). Uppsala: Almqvist and Wiksell.
- Lindblom, B., Sussman, H., Modarresi, G., & Burlingame, E. (2002) The trough effect: Implications for speech motor programming. *Phonetica*, 59, 245–62.
- Lubker, J. F. & Gay, T. (1982) Anticipatory labial coarticulation: Experimental, biological and linguistic variables. *Journal of the Acoustical Society of America*, 71, 437–48.
- Magen, H. S. (1989) An acoustic study of vowel-to-vowel coarticulation in English. Doctoral dissertation, Yale University, New Haven, CT.
- Manuel, S. (1987) Acoustic and perceptual consequences of vowel-to-vowel coarticulation in three Bantu languages. Doctoral dissertation, Yale University, New Haven, CT.
- Manuel, S. (1990) The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America*, 88, 1286–98.
- Manuel, S. & Krakow, R. (1984) Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Report on Speech Research*, 77/78, 69–78.
- Menzerath, P. & Lacerda, A. de (1933) *Koartikulation Steuerung und Lautabgrenzung*. Bonn: Ferdinand Dummlers Verlag.
- Moll, K. & Daniloff, R. (1971) Investigation of the timing of velar movements during speech. *Journal of the Acoustical Society of America*, 50, 678–84.
- Modarresi, G., Sussman, H., Lindblom, B., & Burlingame, E. (2004) Stop place coding: An acoustic study of CV, VC#, and C#V sequences. *Phonetica*, 61, 2–21.
- Moon, S. J. & Lindblom, B. (1994) Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America*, 96, 40–55.
- Nolan, F. (1992) The descriptive role of segments: Evidence from assimilation. In G. J. Docherty & D. R. Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 261–80). Cambridge: Cambridge University Press.
- Nolan, F., Holst, T., & Kühnert, B. (1993) Modelling [s] to [ʃ] assimilation. *Journal of Phonetics*, 24, 113–38.
- Nord, L. (1986) Acoustic studies of vowel reduction in Swedish. *Quarterly Progress and Status Report*, Department of Speech Communication, KTH, Stockholm, 4, 19–36.
- Öhman, S. E. G. (1966) Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 39, 151–68.
- Öhman, S. E. G. (1967) Numerical model of coarticulation. *Journal of the Acoustical Society of America*, 41, 310–20.
- Perkell, J. S. (1990) Testing theories of speech production: Implications of some detailed analyses of variable articulatory data. In W. J. Hardcastle & A. Marchal (eds.), *Speech Production and Speech Modelling* (pp. 263–88). Dordrecht: Kluwer.
- Perkell, J. S. & Cohen, C. (1986) Preliminary support for a “hybrid model” of anticipatory coarticulation. *Proceedings of the 12th International Congress of Acoustics*, Toronto, A3-6.
- Perkell, J. S. & Matthies, M. L. (1992) Temporal measures of anticipatory labial coarticulation for the vowel /u/: Within- and cross-subject variability. *Journal of the Acoustical Society of America*, 91, 2911–25.
- Port, R. & O’Dell, M. (1985) Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455–71.
- Recasens, D. (1984a) V-to-C coarticulation in Catalan VCV sequences: an articulatory and acoustical study. *Journal of Phonetics*, 12, 61–73.

- Recasens, D. (1984b) Vowel-to-vowel coarticulation in Catalan VCV sequences. *Journal of the Acoustical Society of America*, 76, 1624–35.
- Recasens, D. (1984c) Timing and coarticulation: Alveolo-palatals and sequences of alveolar + [j] in Catalan. *Phonetica*, 41, 125–39.
- Recasens, D. (1987) An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish V-C-V 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.
- Recasens, D. (2002) An EMA study of VCV coarticulatory direction. *Journal of the Acoustical Society of America*, 111, 2828–41.
- Recasens, D. (2004) The effect of syllable position in consonant reduction (evidence from Catalan consonant clusters). *Journal of Phonetics*, 32, 435–53.
- Recasens, D. (2006) Integrating coarticulation, blending and assimilation into a model of articulatory constraints. In L. Goldstein, D. Whalen, & C. Best (eds.), *Laboratory Phonology 8* (pp. 611–34). Berlin/New York: Mouton de Gruyter.
- Recasens, D. & Pallarès, M. D. (1998) An electropalatographic and acoustic study of temporal coarticulation for Catalan dark /l/ and German clear /l/. *Phonetica*, 55, 53–79.
- Recasens, D. & Pallarès, M. D. (2001) Coarticulation, blending and assimilation in Catalan consonant clusters. *Journal of Phonetics*, 29, 273–301.
- Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997) A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, 102, 544–61.
- Saltzman, E. (1991) The task dynamic model in speech production. In H. F. M. Peters, W. Hulstijn, & C. W. Starkweather (eds.), *Speech Motor Control and Stuttering* (pp. 37–52). Amsterdam: Excerpta Medica.
- Saltzman, E. & Munhall, K. (1989) A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1, 333–82.
- Solé, M. J. & Ohala, J. (1991) Differentiating between phonetic and phonological processes: The case of nasalization. *Proceedings of the 12th International Congress of Phonetic Sciences* (Aix-en-Provence), 2, 110–13.
- Sussman, H., Bessel, N., Dalston, E., & Majors, T. (1997) An investigation of stop place articulation as a function of syllable position: A locus equation perspective. *Journal of the Acoustical Society of America*, 101, 2826–38.
- Sussman, H. & Modarresi, G. (2003) The stability of locus equation encoding of stop place. *Proceedings of the 15th International Congress of Phonetic Sciences* (Barcelona), 2, 1931–34.
- Sussman, H. M. & Westbury, J. R. (1981) The effects of antagonistic gestures on temporal and amplitude parameters of anticipatory labial coarticulation. *Journal of Speech and Hearing Research*, 24, 16–24.
- Sweet, H. (1877) *Handbook of Phonetics*. Oxford: Clarendon.
- Ushijima, T. & Hirose, H. (1974) Electromyographic study of the velum during speech. *Journal of Phonetics*, 2, 315–26.
- Ushijima, T. & Sawashima, M. (1972) Fiberscopic observation of velar movements during speech. *Annual Bulletin of Research Institute of Logopedics and Phoniatics*, University of Tokyo, 6, 25–38.
- Wood, S. (1993) Crosslinguistic cineradiographic studies of the temporal coordination of speech gestures. *Working Papers*, Lund University, Department of Linguistics, 40, 251–63.
- Wright, S. & Kerswill, P. (1989) Electropalatography in the analysis of connected speech processes. *Clinical Linguistics and Phonetics*, 3, 49–57.