Linguistic Disorders and Pathologies

An International Handbook

Edited by
Gerhard Blanken · Jürgen Dittmann
Hannelore Grimm · John C. Marshall
Claus-W. Wallesch

Offprint

Walter de Gruyter · Berlin · New York
1993

Stemberger, J. P. & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition, 14*, 17–26.

Tannen, D. (1982a). Oral and literate strategies in spoken and written narratives. *Language, 58*, 1–21.

Tannen, D. (1982b). The oral/literate continuum in discourse. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy. 1–16*. Norwood, NJ: Ablex Publishing Corporation.

Tannen, D. (Ed.) (1984). *Coherence in Spoken and Written Discourse*. Norwood: Ablex Publishing Corporation.

Tannen, D. (1987). Repetition in conversation: toward a poetics of talk. *Language, 63*, 574–605.

Wunderlich, D. (1976). *Studien zur Sprechakttheorie*. Frankfurt a. M.: Suhrkamp.

Claus Heeschen, Nijmegen (The Netherlands)

3. Phonological and Articulatory Characteristics of Spoken Language

1.   Introduction

Speaking may be our most impressive motor skill. We speak rapidly, and production of each word involves intricate sequencing and temporal interleaving of gestures for the component, ordered consonants and vowels of the word. The problem of understanding speech production at this level is that of understanding how speakers accomplish the feat of fluent consonant and vowel production. Solving that problem involves solving another one, however. It is to understand what speaking is essentially. That is, it is to understand how a series of complicated actions of a vocal tract can serve to convey a message composed of rulefully-patterned symbols to members of a language community. In fact, the kind of solution an investigator seeks to the problem of understanding how vocal-tract actions are executed depends on how the investigator looks at the relation between vocal-tract action and the linguistic message itself.

Phonology is traditionally seen as the discipline that concerns itself with the building blocks of linguistic messages. It is the study of the structure of sound inventories of languages and of the participation of sounds in rules or processes. Phonetics, in contrast, concerns speech sounds as produced and perceived. Two extreme positions on the relationship between phonological messages and phonetic realizations are represented in the literature. One holds that the primary home

for linguistic symbols, including phonological ones, is the human mind, itself housed in the human brain. The second holds that their primary home is the human vocal tract.

Consider the first position and the conceptualization of speech production to which it leads. For at least two reasons, the vocal tract is rejected as a natural home for phonological segments of the language. A philosophical reason is that phonemes are not the kinds of things that can occur or exist outside the mind. They are ideas or concepts without real-world actualization. Articulatory gestures or their acoustic consequences can serve as cues to phonological segments, but they cannot be phonological segments.

"[Segments] are abstractions. They are the end result of complex perceptual and cognitive processes in the listener's brain" (Repp 1981, 1462).

"Phonological representation is concerned with speakers' implicit knowledge, that is with information in the mind ... [Phonetic] representation ... is not cognitive because it concerns events in the world rather than events in the mind" (Pierrehumbert 1990).

A practical reason why phonological segments cannot occur in the vocal tract is that linguistic symbols have other properties, aside from being covert kinds of things, that preclude the vocal tract from representing them veridically or even analogically. In particular, a central and important fact about language is that its messages are composed of discrete symbols. Phonological segments are discrete in the sense that they do not overlap and blend. Moreover, until recently, they have been represented in linguistic theories as if they were composed of lists of coextensive (and by implication, cotemporal) features (cf.

Chomsky/Halle 1968). The features themselves described static postures of the vocal tract or their acoustic consequences; accordingly, the feature lists of a word described a succession of vocal-tract or acoustic snapshots. The vocal-tract actions that somehow convey a message to a listener have none of those properties. Actions associable with a given consonant or vowel do overlap and do appear to blend with actions of neighbors. Actions identifiable with the component features of a consonant or vowel are not cotemporal. Finally, fundamental units of articulation appear to be actions, not postures; accordingly, time is intrinsic to speech, rather than extrinsic as it is to the linguistic message. One interpretation of these mismatches is that they reflect the mismatch between the ideal of linguistic competence and the degraded physical reality of linguistic vocal performance; the latter necessarily is a considerable distortion of the former due to the limitations of mechanico-inertial systems. This way of looking at speech production promotes development of a kind of theory of the 'how' of speech production that have been termed 'translation theories' (Fowler/Rubin/Remez/Turvey 1980). The mismatch between the character of the planned message, presumably a sequence of linguistic symbols, and of its physical, phonetic, realization requires a translation over stages of processing out of the ideal, mental, domain of the plan into the real, physical-nonmental, domain of a vocal tract.

The other extreme perspective on the nature of speaking is that consonants and vowels are actions of the vocal tract that have linguistic, including phonological, significance in a language community. They are, certainly, psychological actions that require knowledge about them to be performed. However, the knowledge is not a superior 'ideal' that the actions cannot implement; rather, the knowledge is about the actions, derived from perceptual and articulatory experience with them. From this perspective, the mismatch between linguistic segments and articulation described above is apparent rather than real. It is the product of three kinds of error: 1. a mistaken ascription of primacy to linguistic knowledge (competence) over linguistic activity (performance); 2. an incorrect characterization of phonological segments in linguistic theory; 3. an incorrect characterization of the vocal tract actions of speech production. As to the first 'error', the

argument is that we treat language differently from other human creations when we decide that its components exist only in the mind. Other human creations include, for example, automobiles, baseball games and musical pieces. Automobiles definitely exist in the world and so do baseball games and musical pieces when they are played. What is in the mind of those who know about automobiles, baseball and a musical piece, is only what they know about those things; it is not the things themselves. If linguistic concepts are like these other concepts, they are knowledge about real-world objects or events; the events have a psychological nature — in this case, they are actions of the vocal tract, identified as phonological segments. If the phonology in the mind of a language user is what the user knows about the actions that implement a linguistic message, then there need be no mismatch between knowledge and action. If a phonological theory ascribes properties to phonological segments as known that are impossible to realize in vocal-tract action, then the first hypothesis should be that the theory is wrong, not that vocal-tract action distorts components of linguistic competence. If descriptions of vocal-tract actions include properties, such as coarticulatory blending, that would distort the phonological message, then the first hypothesis should be that the descriptions are wrong. From this perspective, an important aim is to work on development of a phonology that does not ascribe properties to phonological segments that are unproduceable as vocal-tract action (cf. Browman/Goldstein 1986; Browman/Goldstein 1989). A second aim is to find a perspective on vocal-tract action from which macroscopic order is evident that conforms to the phonological structure of spoken utterances (e. g., Fowler/Rubin/Remez/Turvey 1980; Fowler 1990; Saltzman 1986; Saltzman/Munhall 1989).

This theoretical perspective promotes a theory of speech production different from a translation theory as outlined earlier. Speech production does not involve a translation out of an ideal, mental domain into a physical, nonmental, domain. Rather, the plan for a sequence of phonological segments, physically instantiated in the brain, replicates itself in a new physical medium, the moving vocal tract. A speech plan, in some way, brings about vocal-tract actions having linguistic significance.

In the remainder of this chapter, I pursue the different outlooks on a central aspect of

speech production, coarticulation, that these different theoretical perspectives promote. I then consider the implications of our understanding of coarticulation for understanding another central aspect of speech production: the coordinated actions of the vocal tract that constitute token phonological segments.

2. Two Perspectives on Coarticulation

All sources of evidence regarding speech production, whether they are acoustic or articulatory, provide the same general picture of context-sensitivity in speech production. An acoustic signal displayed spectrographically or as a waveform, for example, can be divided into phonological-segment sized regions (e. g. Klatt 1973) by identifying acoustic properties that are more strongly associated with one particular segment of an utterance than with others. For example a stop burst can be assigned to a stop in a stop-vowel utterance, and the following voiced formants can be assigned to the vowel. Even so, however, the display has not thereby been partitioned into phonological segments or even into their acoustic consequences. This is so, in part, because there may be no obvious place to locate a boundary separating the acoustic consequences of one phoneme from those of another. For example, the voiceless formant transitions following a voiceless stop consonant in a consonant-vowel sequence belong with the stop, because they are voiceless, but they also belong with the vowel, because formants are characteristic of vowels and other sonorants (see, e. g. Peterson/Lehiste 1960). Indeed, generally, there are no boundaries between segments so that a partition leaves all and only the acoustic information for one segment on one side of the boundary and all and only that for another segment on the other side (cf. Fant/Lindblom 1961). Moreover, the overlap is not only in a potential boundary region. Spectral analysis of the signal well within a domain associated with a particular phonetic segment — well within the frication region for a fricative or within the steady-state formants, if any, of a vowel, for example — is likely to reveal influences of context. (I will use the term 'domain' to refer to the temporal region in which the features of a segment dominate in articulation or in the acoustic signal. The domain does not include the whole articulatory extent of a segment or the whole region in which it influences the acoustic signal, but only the

region in which it is dominant; see also Löfqvist 1990.)

Examination of the articulatory behaviors that give rise to acoustic speech signals reveals a compatible picture. Articulatory movements can be found that are identifiable with one of the phonetic segments in an utterance — movement toward bilabial closure in a bV sequence, for example. In addition, boundaries can be located around that movement. In the example, a boundary may be located where closing is first detectable and another at the point of release of the closure. Once again, however, the boundaries are not boundaries between phonological segments or their articulatory consequences so that all and only movements associated with /b/ occur within the boundaries and movements associated with other segments fall outside the boundaries. During the closing and closure gestures, the tongue body will be conforming itself to the requirements of the following vowel (e. g. Öhman 1966), and once again, the movements within the boundaries are context-sensitive. For example, the jaw moves to a higher point of maximum closing for /b/ followed by a high than a low vowel (Keating/Lindblom/Lubker, cited in Keating 1990).

Sources of context sensitivity are bidirectional. Effects of earlier segments in the string extending beyond their domains of prominence are termed 'left-to-right', 'perseverative' or 'carryover' effects. Effects of later segments are called 'right-to-left' or 'anticipatory.' Estimates of the coarticulatory field — that is, the interval of time or the number of segmental domains affected by a segment in either direction — vary considerably, but may be quite large. For example, Magen (1989) reports anticipatory effects of V_3 on $V_1 C \partial C V_3$ sequences in English. While some part of the carryover influences can bee ascribed to inertial properties of the vocal tract and to its inability instantaneously to adopt a characteristic posture for one phonological segment without exhibiting transitional movements between the postures, anticipatory coarticulation cannot have that explanation, and carryover effects are sometimes more extensive than can be realistically ascribed to these mechanical factors (Daniloff/ Hammarberg 1973). These considerations have suggested to many investigators that coarticulation is planned. Generally accounts of coarticulation diverge along the theoretical lines distinguished in the introduction.

2.1. Coarticulation as Assimilation by Feature Spreading

In a translation theory, coarticulation serves an important function of, indeed, translating a planned symbol string into a form more compatible with the capabilities of vocal-tract action. (The role of phonetic rules generally, according to Keating (1990), is to make the linguistic representation "more physical.")

One example of a theory in which coarticulation serves that function is that of Daniloff/Hammarberg (1973). Daniloff/Hammarberg described the phonological segments that serve as "input" in a plan to speak as "canonical forms" — that is. "invariant, ideal, uncoarticulated forms" — the phonological types of a linguistic theory. These forms undergo "articulatory encoding" to tailor them to the vocal tract. The encoding processes include application of context-sensitive rules of feature spreading. An example they provide of such a rule is one that spreads a rounding feature from a vowel to a preceding /ʃ/: ʃ → ʃʷ / __ [+ round. ÷ V]. By this rule, the /ʃ/ in *shoe*, for example. acquires the feature [÷ round] from its context, a following rounded vowel. Generally (following Henke 1966), rules cause a feature to spread in an anticipatory direction to any phonetic segment that is 'unspecified' for that feature. Feature values in phonological theory generally are binary, and a segment may be 'specified' for a feature by having either a '+' or a '−' value of that feature. (Accordingly, a rounded vowel is [+ round] while an unrounded vowel is [− round].) To count as an instance of a segment specified for some feature value, a token occurrence of the segment must have the appropriate feature value; changing the value may change one segment into another and hence, in a sequence of phonemes, may change one word into another. These feature values thereby serve a 'contrastive' function in the language. At least hypothetically, the contrastive feature values cannot be changed by a feature-spreading coarticulatory rule. However, some features are irrelevant to the identification of some segments. For example, in English, rounding is not contrastive for consonants; accordingly, making a consonant rounded does not change it from one consonant of English to another. Consonants are said to be 'unspecified' for rounding, and they are subject to coarticulatory rules of feature spreading.

Evidence compatible with the feature-spreading theory includes findings (or, perhaps, interpretations of findings; see 2.2) that lip rounding anticipates a rounded vowel across any number of preceding consonants (e. g. Daniloff/Moll 1968; Benguerel/Cowan 1974) and that nasality anticipates a nasal consonant across any number of vowels uninterrupted by oral consonants (Moll/Daniloff 1971).

The simple characterization of coarticulation fails in several ways. One is that the coarticulatory field very often does not respect boundaries drawn between segments. That is, the hypothesis of feature spreading as the sole source of coarticulation predicts that the spread feature should be uniformly present throughout the production of the segment — at least to the same extent that other features of the segment are present, but that is generally not the case (e. g., Benguerel/Cowan 1974; Krakow 1989). Second, the magnitude of effects of ostensibly spread features is gradient rather than categorial. For example, Manuel/Krakow (1984) found that a following (front, high) vowel /e/, raises and fronts following (low. back) vowel /a/, but (front, high) /i/ raises it even more. Likewise. Marchal (1988) reported graded effects of one stop consonant on another in /kt/ sequences that suggested varying degrees of coarticulatory overlap between them. A third problem is that coarticulatory influences may affect realizations of specified features. In Marchal's findings, just cited, coarticulatory influences occur between stops specified for different places of articulation. A final problem relates to the idea of underspecification. The problem here is that segments considered to be unspecified for a feature involving some articulator — say, rounding and the lips (in English, consonants) or nasality and the velum (in English, vowels) — are not wholly neutral with respect to the demands they make on the articulator. Some consonants are associated with rounding movements of the lips (for example, /l/, /r/ and /s/ and /ʃ/ (Bell-Berti/Harris 1982; Delattre/Freeman 1968; Leidner 1973). Compatibly, vowels, ostensibly unspecified for nasality are associated with characteristic postures of the velum (Bell-Berti 1980; Moll 1962). Despite their not being wholly unspecified in terms of articulatory control, they are subject to coarticulatory influences from specified neighbors and they coarticulate with neighbors. For example, the different velum heights associated with vowels of different heights both influence velum height for neighboring consonants and

they are recipients of coarticulatory influences from nasal consonants (Bell-Berti 1980). Accordingly, in contrast to the feature-spreading account of coarticulation, coarticulatory influences occur in the absence of any linguistic features to spread.

Recently, Keating (1988, 1990) has proposed an alternative account of specification and its role in coarticulation that preserves the idea of coarticulation as a participant in a translation from the mental to the physical domain of talking. She proposes that coarticulation includes processes at two levels at least, one phonological and one phonetic. At the phonological level, coarticulation is assimilatory feature spreading. Since Keating's focus has been on phonetic coarticulation, she simply alludes to this type of coarticulation without providing an example. However, a possible example is provided by Daniloff/Hammarberg (1973). They point out that in the word *width*, there is, apparently, a spreading of the interdental place of articulation of /θ/ to /d/ (which, by the way, is specified for a different place of articulation; however, in this case, the feature change does not yield a different phoneme of English). As for phonetic coarticulation, Keating proposes a "targets and connections" model. In the model, phonetic segments are associated with characteristic targets, and segments are sequenced by interpolating between successive targets. A novel aspect of Keating's idea of targets, however (but see Manuel 1987, for a similar idea), is that the targets are regions ("windows"), rather than fixed postures. Windows differ in their widths, and a target's instantiation within its window will depend on its neighbors in that the speaker will generally select the most efficient path from segment to segment that passes through each target region. The idea of target windows replaces the idea of underspecification as categorial with a gradient version. A segment with the narrowest possible window for some feature is "specified" for that feature value; one with the widest possible window for a feature is unspecified. However, most segments have intermediate target window sizes for their component features. Vowels have wider windows for velum height than do nasal and oral consonants, but the window is not as wide as possible. Accordingly, a vowel's window region does affect the articulatory path through the target window of neighboring segments.

This model handles the data of coarticulation considerably better than does the feature spreading model of Daniloff/Hammarberg (1973); yet it preserves the idea of coarticulation as among the processes that make the planned utterance 'more physical.' The targets and connections model is not obviously consistent with all of the data, however. In particular, one finding that the model does not seem to handle well is the ubiquity of coarticulatory fields that extend beyond immediate neighbors. The targets and connections idea explains how contiguous segments can be produced smoothly, but it does not readily predict strong coarticulatory influences of a segment C on A in an ABC sequence. Two other problems emerge below. They are that some coarticulation is difficult to characterize as anything other than overlap (for example, findings by Marchal 1988, cited above). A second is that a segment's 'aggressiveness' (here, having a narrow window) in its own domain appears always to be associated with a compatible degree of aggressiveness outside of its domain, frequently beyond any transitional region between target regions.

2.2. Coproduction Theories

A "coproduction" theory (Fowler 1977) explains coarticulation as the overlapping production of — to a first approximation — invariant sequences of consonants and vowels. The context sensitivity apparent in the acoustic signal and in articulation is not 'deep' context sensitivity in the sense that consonants or vowels have undergone assimilatory change (as in a feature spreading theory). Rather it is a more peripheral blending of consonants and vowels that are unchanged with respect to their essential, specified, properties.

Öhman's (1966; 1967) theory provides a seminal example of such a theory (but see also, however, (Kozhevnikov/Chistovich 1965). In a spectrographic analysis of $V_1 C V_2$ disyllables, Öhman noticed many instances in which the closing transitions into the consonant depended not only on V_1, but also on V_2. Likewise, transitions following consonant release depended on both vowels. X-ray tracings (see also Öhman 1967) showed clear evidence that the tongue body conformation during C closure was different in the context of different flanking vowels. Öhman (1966, 166) suggested that the stop gestures were "superimposed" on a diphthongal vowel-to-

vowel gesture of the tongue body and that the "tongue is able to make a distorted vowel gesture, while it is executing the stop consonant". More speculatively, he proposed three neuromuscular systems for controlling the tongue. The systems, though distinct, would use overlapping muscles. One system, the apical system, is used to produce dental, alveolar and retroflex consonants; the dorsal system produces palatal and velar consonants, and the tongue body system produces vowels. During speech production, a consonant and vowel system may be controlling the tongue in overlapping time frames and the result is "a complex summation (neural, muscular and probably mechanical also) of the responses to each of the components of the instruction" (1966, 166). Öhman's observations have been replicated many times. For example, Perkell (1969) noticed that the /k/ constriction during /hõkɛ/ consisted of a sliding movement of the tongue dorsum toward the front location for /ɛ/. Compatible evidence of vowel-to consonant anticipatory and carryover coarticulation and sometimes vowel-to-vowel coarticulation in V C Vs is provided by Barry/Kuenzel (1975), Butcher/Weiher (1976) and Carney/Moll (1971).

These findings are not captured naturally in a feature spreading account of coarticulation. The main reason is that they reveal the dynamic nature of changing articulatory parameter values during speech. Consider Perkell's finding just described. There is no change in a feature value for /k/'s place of articulation that would yield a sliding place value. The outcome is explained more naturally as a growing influence of /ɛ/'s articulatory demands during /k/.

Öhman (1967) developed a quantitative model of vowel-consonant-vowel coarticulation that did a satisfactory job of predicting the changing vocal-tract shapes (as indexed by X-ray tracings) during V C V production. Notably it includes a parameter value, k, and other parameters labeled q, to implement consonant and vowel production respectively over time. To implement the temporal articulatory domain of a consonant or vowel, the associated parameter increases over time and then decreases. That is, to generate coarticulatory influences of the vowel on the consonant, for example, the vowel's influence on the vocal tract gradually waxes and then wanes. Elsewhere, we have described this waxing and waning of a segment's implementation over time as a "prominence curve"

(Fowler/Smith 1986); see Löfqvist's (1990) similar idea of "dominance").

In light of this evidence favoring coproduction, let us reconsider the data considered most supportive of feature spreading theory, evidence that lip rounding anticipates across consonant strings unspecified for rounding and that velum lowering anticipates across vowel strings unspecified for nasality. Difficulties with the idea of underspecification have already been cited. More than that, however, work by Bell-Berti and her colleagues show quite convincingly that the error of accepting underspecification has led to considerable overestimation of anticipation of velum lowering and lip rounding (see also Boyce/Krakow/Bell-Berti/Gelfer 1990).

2.2.1. Anticipatory Lowering of the Velum for Nasal Consonants

Consider the literature on nasalization first. Researchers typically examined $C V_n N$ strings (where Ns are nasal consonants and the subscript on the vowel signifies that different numbers of vowels intervened between C and N). Velar lowering following C was taken as evidence for onset of anticipatory nasalization from N (Moll/Daniloff 1971). However, Bell-Berti (1980) points out that vowels are associated with lower velum heights than are oral consonants; accordingly the initial drop of the velum will be due at least to the vowel: it may or may not reflect an influence of N as well. That can be determined only by comparing $C V_n N$ sequences with corresponding $C V_n C$ sequences. Such a comparison indeed shows a lowering of the velum at the onset of a vowel string in $C V_n C$ utterances that, of course, must be ascribed to the vowel rather than to coarticulatory effects of a nasal consonant (Bell-Berti/Krakow, submitted). When effects of the vowel are eliminated from velum movements in $C V_n N$ utterances, findings are no longer consistent with feature spreading theories. Rather, they suggest an invariant onset of velum lowering relative to onset of nasal murmur in nasal consonant production. Bell-Berti/Harris (1981) interpret the findings as favoring a particular version of a coproduction theory, that they call "frame theory" in which the temporally-staggered onsets of component gestures of a phonetic segment are staggered in a time-invariant way.

The findings by Bell-Berti and her colleagues also help to explain an otherwise complicating finding by Bladon/Al-Bamerni (1982). Bladon/Al-Bamerni had found evi-

dence for two patterns of anticipatory coarticulation of velum lowering — a one-step pattern of lowering, timed consistently with predictions of feature-spreading theory (that is, beginning at the onset of the first vowel in a string) and a two step pattern, the first step beginning at the onset of the first vowel and the second, as frame theory predicts, an invariant interval before the oral closing gesture for the nasal consonant. Bladon/Al-Bamerni were unable to find anything systematically different in the contexts in which each pattern was observed; therefore, they suggested that selection among the strategies was unsystematic. An alternative interpretation, however, is that sometimes the vocalic velum lowering movement (always beginning near vowel onset) overlaps completely with the lowering gesture for the nasal consonant, whereas at other times, it precedes velum lowering for the consonant. Bell-Berti/Krakow (submitted; see also Boyce/Krakow/Bell-Berti/Gelfer 1990) found increasing evidence of two- or multi-stage velum lowering as vocalic segments were added before the nasal consonant. Likewise, of their three talkers, one produced the target words at a considerably faster rate than the others and that subject showed a one-stage lowering pattern for all but the longest vowel segments. Finally, one talker who produced the words at two rates showed two- or multi-stage lowering only at the slower rate.

Overall, the findings on anticipatory velum lowering — originally considered to provide strong evidence in favor of a feature spreading theory of coarticulation — do not; rather, they provide better support for the view that coarticulation is coproduction. Notice, too, that Keating's targets and connections account must at least be modified to fit the data. In particular, the model does not predict that target windows for successive segments will overlap; however, the data just described shows convincingly that they do. That is, this model too must admit the possibility of coproduction. Coarticulation is not wholly finding the most efficient pathway from one target window to another; sometimes windows overlap.

2.2.2. Lip Rounding

The literature on lip rounding, like that on nasalization, has failed to support the feature spreading account. Generally, it supports frame theory. As Kent/Minifie (1977) pointed out, contradictory evidence was available

even in one study commonly cited as supporting feature spreading, namely that of Benguerel/Cowan (1974). In their findings more than half the time, rounding spread not only through a preceding consonant string, but beyond it into a preconsonantal unrounded vowel. Bell-Berti/Harris (1979) obtained similar results for both of their speakers. The study by Bell-Berti/Harris (1979) and a later one (Bell-Berti/Harris 1982) showed a generally invariant relation between onset of EMG (orbicularis oris muscle of the lips) for a rounded vowel and measured acoustic onset of the rounded vowel over a variable number of prevocalic consonants.

The research by Bell-Berti/Harris tested for and found lip EMG activity for /l/, one of the consonants in the strings they used as stimuli. As noted earlier, other investigators have found rounding for other consonants. These consonantal influences on lip configuration are likely to have contaminated estimates of onset of lip rounding in the earlier research in the same way that the vocalic influences on velum height contaminated estimates of onset of velum lowering for nasal consonants. These contaminating influences can only be identified by examining control utterances that lack the specified segment (that is, V Cₙ V utterances in which both vowels are unrounded), and investigators have not done that generally. However, using appropriate control utterances, Boyce (1988) has shown that overlapping consonantal and vocalic lip movements approximately add so that effects of consonants on the lips in a utterance such as /kuktluk/ can be eliminated by subtracting the movement trace from /kiktlik/ from it. Whereas Boyce did not then test for the invariance of EMG onset relative to acoustic onset of the rounded vowel that Bell-Berti/Harris had reported earlier, she did find a clear intervocalic trough in lip movement activity and bimodal peaks of EMG activity in utterances with two rounded vowels. The pair of findings suggests that during the consonantal string /ktl/, rounding from the first vowel wanes while that for the second vowel increases. Hence there are two distinct rounding gestures that wax and wane in the consonantal string — just as Öhman's account of vowel-consonant production proposed. There is not a spreading of a rounding feature from vowel to consonant. Compatibly, Gelfer/Bell-Berti/Harris (1989) superimposed graphs of lip EMG activity (orbicularis oris) for utterances such as /ist # tu/ and

/ist ≠ ti/ having varying numbers of intervocalic consonants and final /u/ or /i/. By eliminating the activity common to both utterances, and hence due to the consonant string, they were able to identify the onset time of EMG activity associated with the rounded vowel itself. Onset times bore a near-invariant relation to release of the occlusion of the final consonant in strings of two or more consonants.

2.2.3. Lingual Coarticulation

The literature on coarticulation involving the tongue supports and augments the idea of coarticulation as gestural overlap. Öhman's model suggests that demands on the articulators made by a segment increase gradually over time and decrease gradually. The serial ordering of segments in articulation is maintained not by preserving discreteness of segment production along the time axis, but, rather perhaps, by maintaining a serial ordering of their times of maximum control in the vocal tract. In addition, however, segments differ one from the other in the strengths of demands they place on different articulators (or on different articulatory systems: see below under 'coordination' and cf. Keating's idea of windows discussed above). The differences in strength have an observable consequence that is described differently (e. g. Farnetani 1990) depending on where it is observed. If discrete domains are identified for segments in an utterance by drawing boundaries at points where coarticulating segments shift in their relative dominance in the vocal tract, then one can say that in their own domain, segments that make strong demands on an articulator 'resist' coarticulatory influences from neighbors (Bladon/Al-Bamerni 1976); in the domains of near neighbors, they exert a strong coarticulatory influence. From the perspective of a coproduction theory, resistance to coarticulation and a strong coarticulatory influence covary because they are really the same thing — namely a segment's exerting a relatively strong influence on articulators throughout its temporal domain.

Recasens (1984; 1985; 1987; 1989) has conducted much of the work that has uncovered variation in coarticulation resistance in movements involving the tongue dorsum. In general, resistance to coarticulation of a consonant or vowel is associated with the amount of tongue dorsum-palatal contact associated with production of the segment (see also Farnetani 1990). Compatibly, using acoustic and electropalatographic measures, Recasens (1984; 1987) found a decrease in vowel-to-vowel coarticulation in V C V sequences in which C is produced with considerable contact between the tongue dorsum (an important articulator in vowel production) and the palate. For example, there is less V-to-C coarticulation across palatal /j/ than across dental /n/. Compatibly, the vowel /i/, which requires a constriction in the palatal region resists consonant-vowel coarticulatory influences moreso than do other vowels (Recasens 1985), and it resists vowel-to-vowel coarticulatory overlap as well (Recasens 1987; 1989). In addition, as noted earlier, segments such as /i/ that are resistant to coarticulation in their own coarticulatory domains themselves exert strong coarticulatory influences on neighbors (see Tables II-VI in Recasens 1987; see also Butcher/Weiher 1976; Farnetani/Vagges/ Magno-Caldognetto 1985).

It may be tempting to conclude from this research that production of consonants and vowels is context sensitive after all in that coarticulatory anticipation of V2 in a V C V sequence must be delayed and reduced if V1 is /i/ as compared to /a/ or if C is /j/ as compared to /n/. However, possibly, the planned segment can be invariant, while its surface manifestations vary according to its neighbor's patterns of coarticulation resistance. Consider, by analogy, the different surface consequences of an invariant squeezing action of the hand depending on whether the hand is empty, or else holding a sponge or a rock. The outcome at the surface is different both in the extent to which the hand (metaphorically, the segment being produced) closes and in the extent that it deforms the sponge (a little coarticulation resistance) and the rock (a lot of resistance). Perhaps by the same token, an invariant plan for a segment can have different surface consequences if coarticulation resistance is implemented as a real physical variable in the vocal tract. There is one striking outcome reported by Recasens (1984) that suggests exactly that. He reported instances both of anticipatory and of carryover coarticulation in which coarticulatory effects were discontinuous. That is, vowel-to-vowel effects were observed in V C V sequences even though, in consonants with considerable tongue dorsum/palatal contact, vowel-to-consonant coarticulation was absent. It is unlikely that talkers plan to begin production of V2 in V1, to stop production

of V2 during C, and to recommence its production after C. An analogous plan for carryover coarticulation is even less likely.

2.3. Some Tentative Conclusions about Coarticulation

The findings just reviewed suggest the following summary. Each consonant or vowel of the language is implemented by one or more vocal-tract actions. Actions are of two varieties: gestures (Browman/Goldstein 1986) that are linguistically significant (and contrastive) and other, noncontrastive, ones that may occur because they are easier to produce than to suppress. Gestures for a segment may be timed or phased invariantly one with respect to another as frame theory proposes. Each vocal tract gesture has a prominence pattern of increasing then decreasing articulatory strength, where prominence refers to the extent to which the gesture exerts an influence on the character of movements in the vocal tract. Vocal tract actions differ one from the other in relative strength so that, for example, demands of /j/ or /i/ on the tongue dorsum-palate relation exceed those of /n/ and /a/. The extent to which a segment-specific action influences what is happening in the vocal tract at any point in time reflects the strength of that action and its strength relative to that of other ongoing actions affecting the same vocal-tract structures. 'Strength' appears to be implemented in such a way that its effects arise at the articulatory surface, not in differential planning for a segment depending on its context. The account is incomplete in a variety of ways, lacking detail in important areas, including a specification of how strength variations are realized. It is also too simple in some respects. In particular, patterns of relative timing of gestures for a segment are not invariant — they may vary over position in a syllable as Krakow (1989) has shown for the relative timing of velum lowering and lip closing actions for syllable-initial and -final /m/. They are likely to vary over stress and rate manipulations as well. In short, the state-of-the art in coarticulation research leaves investigators still with many problems to tackle.

3. Coordination

From the perspective of a coarticulating segment encroaching on the domain of a second segment, the second segment applies restrictions on where and to what extent encroach-

ment can occur ('coarticulation resistance'). Accordingly, coarticulation by the same segment in the same (anticipatory, carryover) direction will be differentially manifested depending on the nature and strength of the restrictions applied in its coarticulatory field. Looked at from the perspective of the influenced segment, however, the restrictions are the segment's own identity; they are actions or postures the achievement of which counts as production of that segment. Somehow realization of the segment correspondingly prohibits contradictory actions. Here I examine implementation of those restrictions in speech production.

The vocal tract includes large numbers of muscles and structures that the muscles move or deform. Relative to the catalogue of movements that could occur were contractions of all possible combinations of vocal tract muscles used and contractions of all possible magnitudes, the movements that do occur in speech are limited in number and in kind. They are constrained, of course, to structure the air so that listeners can hear them. But more than that, they are low-dimensional movements — movements with order that spans groups of muscles and groups of vocal-tract structures. They are, indeed, coordinated actions.

Coordination achieves several things. Most importantly, structures of the vocal tract work together to achieve some end. For example, in production of /b/, the jaw and lips work together to achieve bilabial closure. The couplings among structures also preclude actions that violate the couplings; thereby they prohibit coarticulatory influences that would prevent the goal of the coordinative linkages. They do not completely eliminate variability or flexibility, however. For example, bilabial closure is realized with a variety of contributions by the jaw and lips. When /b/ is coarticulated with an open vowel, the jaw is lower during closure, and hence the lips do more of the closing work, than when /b/ coarticulates with /i/. Research using a perturbation procedure (e. g. Abbs/Gracco 1984; Kelso/Tuller/Vatikiotis-Bateson/Fowler 1984; Shaiman 1989) helps to expose couplings across structures of the vocal tract. In one of these experiments, Kelso/Tuller/Vatikiotis-Bateson/Fowler 1984) asked talkers to produce *It's a _ again*, with /baeb/ or /baez/ serving as target syllable. On a low proportion of trials, randomly selected, during the closing gesture for the second /b/ in /baeb/ or for the

/z/ in /baez/, the talker's jaw was unexpectedly braked, preventing its normal contribution to closure for the consonantal constriction. On perturbed relative to unperturbed trials, within 20 – 30 ms of the perturbation in /baeb/, the orbicularis oris muscle of the up-- per lip showed extra activation and by achievement of closure, the lip had moved farther down than on unperturbed trials. If the jaw was braked during closing for /z/, extra activation was observed in the genio-glossus muscle of the tongue allowing the tongue to compensate for the unusually low position of the jaw. The upper lip did not show the same extra downward movement on /z/-perturbed trials that it showed on /b/-perturbed trials. Other research (Shaiman 1989) shows that when an articulator of the vocal tract is perturbed that is not involved in a consonantal closing gesture, closing on perturbed and unperturbed trials is alike. In short, the responses to perturbation are adaptive and they reveal a coupling among selective articulators of the vocal tract that jointly achieve some phonetic gestural end. Coupled structures and their neuromuscular underpinnings are know as 'synergies' or 'coordinative structures.' Whereas Löfqvist (1990) suggests that there are no dynamic perturbations in speech analogous to a jaw pull, perhaps there are. Coarticulatory encroachments from low vowels can perturb a talker's jaw, pulling it down during closure for /b/. Possibly, then, the couplings serve two functions; they bring about the coordinated action that constitutes a linguistic gesture of the vocal tract, and they permit only those coarticulatory encroachments that will not prevent the gesture from being realized.

The short-latency responses to the perturbations suggest that the couplings are low-level. That is, they are not cognitive couplings, but, rather neuromuscular ones. This may help to rationalize findings by Recasens summarized earlier of discontinuities in coarticulatory influences. Whereas it would be surprising for speakers to plan for V-to-V coarticulatory influences, yet plan for no V-to-C influences in a V C V sequence, the finding of discontinuities in coarticulation is less surprising if segments are planned to have an invariant coarticulatory field that then gets differentially suppressed by other synergies active in the vocal tract.

Following Browman/Goldstein (1986; 1989), we may call the vocal tract actions of a synergy a 'phonetic gesture' or, more simply,

a 'gesture.' Phonetic gestures are, then, linguistically significant actions of the vocal tract. In the research using the perturbation technique just described, perturbations disrupted movements by one articulator among two or more that participated in a phonetic gesture. That is, perturbations and compensations were interarticulatory, but intragestural. However, some phonetic segments are defined by more than one gesture, and the timing or phasing between or among gestures may also be crucial to the identity of the segment. For example, the timing of an oral constriction gesture and a glottal devoicing gesture determines whether a consonant is preaspirated or aspirated (see, e. g., Löfqvist 1980; Löfqvist/Yoshioka 1984). Presumably, then, intrasegmental gestures must be coupled and one should see evidence of the coupling in perturbation experiments. To date, there is little evidence on the topic. However, Munhall/Löfqvist/Kelso (1988) have perturbed the lower lip during closing from a vowel to a /p/. The perturbation delayed achievement of closure, thereby lengthening the vowel. However, onset of glottal opening for /p/ was also delayed, giving rise to a perceptually adequate aspirated /p/. (Even so, there was disruption of the coordinative relation between the gestures such that the voice-onset times on perturbed trials were unusually long.)

Another index, perhaps, of a coupling relation between the gestures of a segment is provided by tests for invariant relative timing (as summarized in Löfqvist 1990). Coupling between gestures of a segment should give rise to invariance of relative timing between the gestures so that, as the segment is produced at various rates or with different levels of stress, temporal intervals between gesture onsets scale proportionately to changes in other intervals produced by the coupled actions. (The idea is that if the gestures are products of a common synergy, and rate changes are achieved by changes in a parameter that is common to the synergy, all temporal intervals produced by the gestures will scale proportionately.) Löfqvist (1990) applied a test for proportionality of intervals proposed by Gentner (1987) to several sets of data including measures of intrasegmental-intergestural intervals and intersegmental-intergestural intervals. Whereas 90% of tests for proportional changes in intervals over variation in rate and stress were rejected in tests of the latter intervals, just 33% were rejected in tests of the former intervals.

Löfqvist does not consider this particularly strong support for the proportional-durational test of coupling between gestures of a segment, because the reason why 67% of tests failed to reject the hypothesis of proportional durations for intrasegmental-intergestural intervals was not that intervals were relatively invariant, but rather because they were extremely noisy (see his Figures 11−15). Even so, his data do reveal marked differences in the temporal relations among gesture belonging to the same and to different phonological segments, with the latter relations showing systematic departures from the proportional-duration hypothesis and the former showing only unsystematic departures.

4. Speech Dynamics

There is a new development in the study of speech production that I will describe only briefly. It is as yet relatively untried; however, it promises to have a marked influence on research in the field. Although speech production is remarkable as a motor activity, it is not wholly unique. Some common issues arise in investigations of a variety of intentional motor skills. More fundamentally, however, some theorists suggest that intentional actions in general (Kugler/Turvey 1987; Kugler/Kelso/Turvey 1980) and speech production in particular (Saltzman 1986; Saltzman/Kelso 1987; Saltzman/Munhall 1989; Kelso/Tuller 1984) constitute a special instance of 'self-organization' in physical systems. Accordingly, they may be best understood by embedding their investigation in the larger context of the study of self-organizing physical systems. Complex physical systems that are open to the flow of energy from the environment, whether they are living systems or not, develop macroscopic, low dimensional patterned and stable activities that can be modeled as attractors of just a few sorts. Most simply, a physical system can be modeled as a 'point attractor' if, when perturbed, it tends to return to the same final target — much as the vocal tract does if it is perturbed during bilabial closure (e. g. Saltzman/Kelso 1987).

Saltzman and colleagues have shown that many central features of speech production, including adaptive responses to perturbations and consequences of coarticulatory overlap (see Saltzman/Munhall 1989), can be modeled if phonetic gestures are modeled as dynamical systems. On the other side, Tuller/Kelso (1990) have shown that speech production

exhibits some of the central characteristic features of dynamical systems (cf. art. 37). Finally, Browman/Goldstein (1986) have developed an "articulatory phonology" whose primitive units, phonetic gestures, are defined by dynamical parameters of the vocal-tract point attractors of Saltzman's articulatory ("task-dynamic") model. Possibly, embedding the investigation of speech production in the context of studies of complex open physical systems generally will help to deepen our understanding of synergies and their achievement of low-dimensional, coordinated actions. In turn, understanding of these physical systems may literally add substance to the linguist's concepts of phonological segments and their features.

5. References

Abbs, J. & Gracco, V. (1984). Control of complex gestures: Orofacial muscle responses to load perturbations of the lip during speech. *Journal of Neurophysiology, 51,* 705−723.

Barry, W. & Kuenzel, H. (1975). Coarticulatory airflow characteristics of intervocalic voiceless plosives. *Journal of Phonetics, 3,* 263−282.

Bell-Berti, F. (1980). Velopharyngeal function: A spatio-temporal model. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice. 291−316.* New York: Academic Press.

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

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

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

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

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

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

Bladon, A. & Al-Bamerni, A. (1982). One-stage and two-stage temporal patterns of coarticulation. *Journal of the Acoustical Society of America, 72,* 104.

Boyce, S. (1988). *The influence of phonological structure on articulatory organization in Turkish and English: Vowel harmony and coarticulation.* PhD dissertation, Yale University.

Boyce. S., Krakow, R., Bell-Berti, F., & Gelfer, C. (1990). Converging sources of evidence for dissecting articulatory movements into gestures. *Journal of Phonetics. 18,* 173—188.

Browman, C. & Goldstein, L. (1986). Towards an articulatory phonology. In C. Ewan & J. Anderson (Eds.), *Phonology yearbook 3. 219—252.* Cambridge: Cambridge University Press.

Browman, C. & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6,* 201—252.

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

Carney, P. & Moll, K. (1971). A cinefluorographic investigation of fricative consonant-vowel coarticulation. *Phonetica, 23,* 192—202.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English.* New York: Harper and Row.

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

Daniloff, R. & Moll, K. (1968). Coarticulation of lip rounding. *Journal of Speech and Hearing Research, 11,* 707—721.

Delattre, P. & Freeman, D. (1968). A dialect study of American rs by X-ray motion picture. *Linguistics. 44,* 29—68.

Fant. G. & Lindblom, B. (1961). Studies of minimal speech and sound units. *Speech Transmission Laboratory: Quarterly Progress Report. 2,* 1—21.

Farnetani. E. (1990). V-C-V lingual coarticulation and its spatiotemporal domain. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling. 93—130.* The Netherlands: Kluwer.

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

Fowler, C., Rubin, P., Remez. R., & Turvey, M. (1980). Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Langage production I: Speech and talk. 373—420.* London: Academic Press.

Fowler, C. & Smith, M. (1986). Speech perception as "vector analysis": An approach to the problems of segmentation and invariance. In J. Perkell & D. Klatt (Eds.), *Invariance and variability of speech processes. 123—136.* Hillsdale, N. J.: Lawrence Erlbaum Associates.

Fowler, C. A. (1977). *Timing control in speech production.* Bloomington, Indiana: Indiana University Linguistics Club.

Fowler, C. A. (1990). Comments on the contributions by Pierrehumbert and Nearey. *Journal of Phonetics, 18,* 425—434.

Gelfer, C., Bell-Berti, F., & Harris, K. (1989). Determining the extent of coarticulation: Effects of experimental design. *Journal of the Acoustical Society of America, 86,* 2443—2445.

Gentner, D. (1987). Timing of skilled motor performance: Tests of the proportional duration model. *Psychological Review, 94,* 255—276.

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

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

Keating, P. (1990). The window model of coarticulation: articulatory evidence. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I. 451—470.* Cambridge: Cambridge University Press.

Kelso, J. A. S. & Tuller, B. (1984). Converging evidence in support of common dynamic principles for speech and movement control. *American Journal of Physiology, 246,* R982—R935.

Kelso, J. A. S., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C. A. (1984). Functionally-specific articulatory cooperation following jaw perturbation during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance. 10,* 812—832.

Kent, R. & Minifie, F. (1977). Coarticulation in recent speech production models. *Journal of Phonetics, 5,* 115—144.

Klatt, D. (1973). Vowel-lengthening is syntactically determined. *Journal of Phonetics, 3,* 129—140.

Kozhevnikov, V. A. & Chistovich. L. A. (1965). *Speech: Articulation and Perception* (30, 543). Washington D. C.: Joint Publications Research Service.

Krakow, R. (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures.* PhD, Yale University.

Kugler, P., Kelso, J. A. S., & Turvey, M. (1980). On the concept of coordinative structures as dissipative structures, I. Theoretical lines of convergence. In G. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior. 3—47.* Amsterdam: North Holland.

Kugler, P. & Turvey, M. (1987). *Information. natural law and the self-assembly of rhythmic movement.* Hillsdale, N. J.: Lawrence Erlbaum Associates.

Leidner, D. (1973). *An electromyographic and acoustic study of American English liquids.* PhD, University of Connecticut.

Löfqvist, A. (1980). Interarticulator programming in stop production. *Journal of Phonetics, 8,* 475—490.

Löfqvist, A. (1990). Speech as audible gestures. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling. 289—332.* The Netherlands: Kluwer.

Löfqvist, A. & Yoshioka, H. (1984). Intrasegmental timing: Laryngeal-oral coordination in vowel-consonant production. *Speech Communication, 3,* 279—289.

Magen, H. S. (1989). *An acoustic study of vowel-to-vowel coarticulation in English.* PhD dissertation, Yale University.

Manuel, S. (1987). *Acoustical and perceptual consequences of vowel-to-vowel coarticulation in three Bantu languages.* PhD dissertation, Yale University.

Manuel, S. & Krakow, R. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Reports on Speech Research, SR-77/78,* 69—87.

Marchal, A. (1988). Coproduction: Evidence from EPG data. *Speech Communication, 7,* 287—295.

Moll, K. (1962). Velopharyngeal closure in vowels. *Journal of Speech and Hearing Research, 36,* 30—77.

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

Munhall, K., Löfqvist, A., & Kelso, J. A. S. (1988). Lip-larynx coordination in speech: A perturbation study. *Journal of the Acoustical Society of America, 80,* S109.

Öhman, S. (1966). Coarticulation in V C V utterances: Spectrographic measurements. *Journal of the Acoustical Society of America, 39,* 151—168.

Öhman, S. (1967). Peripheral motor commands in labial articulation. *Speech Transmission Laboratory: Quarterly Progress Report, 4,* 30—63.

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

Peterson, G. & Lehiste, I. (1960). Duration of syllable nuclei. *Journal of the Acoustical Society of America, 32,* 693—703.

Pierrehumbert, J. (1990). Phonological and phonetic representations. *Journal of Phonetics, 18,* 315—394.

Recasens, D. (1984). V-to-C coarticulation in Catalan V C V sequences: An articulatory and acoustical study. *Journal of Phonetics, 12,* 61—73.

Recasens, D. (1985). Coarticulatory patterns and degrees of coarticulation resistance in Catalan C V sequences. *Language and Speech, 28,* 97—114.

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

Recasens, D. (1989). Long range coarticulatory effects for tongue dorsum contact in V C V C V sequences. *Speech Communication, 8,* 293—301.

Repp, B. (1981). On levels of description is speech research. *Journal of the Acoustical Society of America, 69,* 1462—1464.

Saltzman, E. (1986). Task dynamic coordination of the speech articulators. In H. Heuer & C. Fromm (Eds.), *Generation and modeling of action patterns. 129—144.* New York: Springer-Verlag.

Saltzman, E. & Kelso, J. A. S. (1987). Skilled action: A task-dynamic approach. *Psychological Review, 94,* 84—106.

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

Shaiman, S. (1989). Kinematic and electromyographic respones to perturbation of the jaw. *Journal of the Acoustical Society of America, 86,* 78—88.

Tuller, B. & Kelso, J. A. S. (1990). Phase transitions in speech production and their perceptual consequences. In M. Jeannerod (Ed.), *Attention and Performance VXII. 429—452.* Hillsdale, N. J.: Lawrence Erlbaum Associates.

Carol A. Fowler, Hanover, New Hampshire (USA)

4. Characteristics of Hesitation Phenomena

1.   The Hesitant Nature of Speech

Normal spontaneous speech is not as smooth flowing as it might seem. Rather, it is characterized by hesitations which listeners filter out. These hesitations can take the form of silent pauses, filled pauses, such as *um* and *ah*, false starts, repetitions, or syllable lengthening. Excluding silent pauses that are so short that they could simply be a normal part of the articulation of certain sounds, it would be quite normal, for example, to find that in spontaneous speech a speaker produced a silent pause before 65% of the sentences uttered and a filled pause before 25%. Further, it would be quite normal for over 20% of clauses within the sentences to be preceded by a silent pause and for over 4% to be preceded by a filled pause. Pausing before