

Phonetica 1988; 45: 140–155

Some Notes on Syllable Structure in Articulatory Phonology

Catherine P. Browman, Louis Goldstein

Haskins Laboratories and Department of Linguistics, Yale University, New Haven, Conn., USA

Abstract. Two approaches to seeking stable patterns in the gestural organization of speech are examined: local organization (individual gestures coordinated with other individual gestures) and global organization (gestures forming larger conglomerates). Articulatory evidence from American English words with a variety of initial consonants and clusters shows that syllable-initial consonants form a global organization (indexed by a metric we term the C-center) that is coordinated with the syllable's vowel gesture. For syllable-final consonants, however, the evidence suggests that a local organization is employed: The first postvocalic consonant gesture is coordinated with the vowel gesture. Implications of these different styles of organization for the perceptual and phonological structure of speech are discussed.

Introduction

By observing vocal tract activity during the act of speaking (for example in a cine-fluorographic film), various stable (almost repeatable) patterns may be seen, such as the formation and release of constrictions made with the lips or the blade of the tongue. Such characteristic patterns of movement, or *gestures*, can form the basis for a phonological description of speech [Browman and Goldstein, 1986; see McCarthy, this volume, for a review of other recent proposals that incorporate articulatory structure into the phonology]. The stability of these patterns has been seen both in their

repeatability in different contexts [Macchi, this volume; Sussman et al., 1973] and in their response to externally applied perturbations [Abbs et al., 1984; Kelso et al., 1984].

Using the task dynamics model [Saltzman, 1986; Saltzman and Kelso, 1983], it is possible to characterize individual gestures in terms of coordinated movement patterns of sets of articulators. However, while it is clear that gestures must themselves be coordinated in larger scale stable patterns, it is not yet clear what these patterns are. Nor is it clear to what phonological units (if any) such larger scale stability might correspond: consonant clusters, onsets and

rimes, syllables, words, phrases, etc. While Browman and Goldstein [1986] have described the phonological structure of a lexical item as a 'constellation' of gestures, that is, a stable organization among gestures, this lexically based organization is modified for words in a phrase [Browman and Goldstein, in press]. Other attempts to find stabilities have yielded conflicting results for variations in stress and rate of speech. On the one hand, some results suggest that the onset of a consonant gesture is timed invariantly relative to a particular phase of the ongoing vowel gesture's production [Tuller et al., 1982; Tuller and Kelso, 1984; Kelso and Tuller, 1985]. On the other hand, more recent work has indicated that such timing is not, in fact, absolutely invariant [Nitrouer et al., 1986]. Thus, the question of intergestural coordination is still very much unresolved.

In conjunction with our colleagues Elliot Saltzman and Philip Rubin at Haskins Laboratories, we are currently developing a gesture-based computational model of speech production [Browman et al., 1986; Saltzman et al., 1987]. In a preliminary attempt to expand the rule set of this model to handle intergestural coordination for American English, two different organizations of the set of oral gestures associated with an utterance were postulated: an articulatory and a functional organization [Browman and Goldstein, in press]. In the articulatory organization, oral gestures are organized into tiers that correspond to the anatomical structures of lips, tongue tip, and tongue body. [These tiers correspond to the features in 'articulator theory', as described by McCarthy, this volume]. In the functional organization, the oral gestures are projected onto two functional tiers, the

consonant and vowel tiers. Intergestural coordination is expressed with respect to this functional division between consonantal and vocalic gestures, with individual vocalic and consonantal gestures coordinated with each other. While the assumption of the importance of functional (i.e., consonant and vowel) organization has a fair amount of support in both the phonological and phonetic literature [Öhman, 1966; McCarthy, 1981; Clements and Keyser, 1983; Fowler, 1983; Keating, 1985], two specific aspects of the preliminary formulation of the rules warranted further investigation: the lack of any effect of syllable affiliation on gestural coordination, and the lack of any intermediate levels of organization. In the current paper, therefore, we explore both the role of syllable affiliation and the possibility of a type of coordination other than individual gesture to individual gesture.

Articulatory Analyses: C-Centers

In order to explore the effect of syllable affiliation, we compared gestural patterns of utterances with different presumed word affiliations for consonantal gestures, for example, 'pea splots' versus 'piece plots'. In particular, we examined articulatory trajectories for the utterances shown in table I (one token per utterance). The relevant portion of these utterances is the sequence $V_1C(C)V_2$, where V_1-V_2 is either [i-a] or [a-i] and consonants include [s], [p], and [l], and the combinations [sp], [sl], [pl], and [spl]. (Phonetic transcriptions are broad everywhere except in figure 1, i.e., aspiration not indicated.) In addition, a word boundary occurs either before or after the

Table I. List of utterances

Frame sentence: 'Miss _____ it.'	
[pa 'pidz]	[pap 'its]
[pa 'sidz]	[pas 'its]
[pa 'lidz]	[pas 'lidz]
[pa 'spidz]	[pas 'pidz]
[pa 'plidz]	[pap 'lidz]
[pa 'splinz]	[pas 'plidz]
[pi 'pats]	[pip 'adz]
[pi 'sats]	[pis 'adz]
[pi 'lats]	[pis 'lats]
[pi 'spats]	[pis 'pats]
[pi 'plats]	[pip 'lats]
[pi 'splats]	[pis 'plats]

Word boundaries were indicated in the orthographic materials from which the speaker read.

first consonant of the string. Thus, there are pairs of utterances differing only in the presumed syllable affiliation of the first consonant in the consonant sequence.

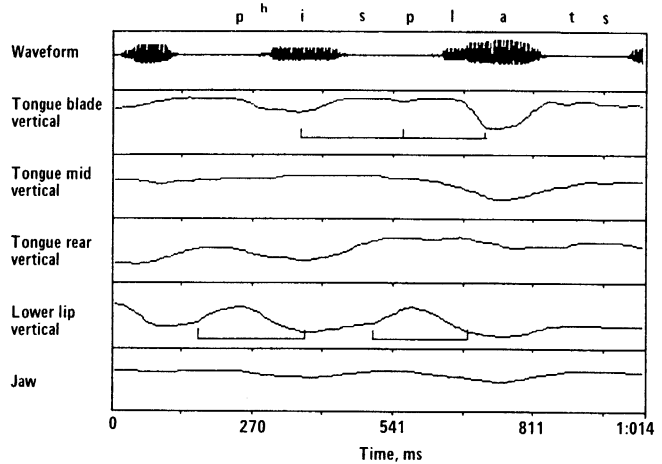
Table I also has sets of utterances with the same syllable affiliation, but with different numbers of consonants in the sequence – for example, [pi 'sats], [pi 'spats], [pi 'splats]. We used this variation in the data to explore the possibility of a different, more global, measure of gestural coordination. In earlier analyses of intergestural coordination, we had examined stabilities in the phasing of a single gesture with respect to other single gestures. This approach is, by its nature, a *local* description of gestural relations. Such a local description, however, while useful as a first step, probably oversimplifies actual gestural relations. Individual pairwise relations among gestures may not be maintained – observation of articulatory data suggests that a whole complex of gestures is rearranged as some linguistic variable changes. It would be useful, then,

to develop a measure for a more *global* description of stable gestural relations. The global metric we propose in this paper (discussed below) is that of the **C-center** (consonant center).

The data examined were taken from the Tokyo X-ray microbeam data base, processed at AT&T Bell Labs [Miller and Fujimura, 1982]. (A subset of the data presented here is discussed in Browman and Goldstein [in press].) The microbeam system [Fujimura et al., 1973; Kiritani et al., 1977] tracked the vertical and horizontal displacements of small lead pellets placed on the surface of the tongue, lips, and teeth. For example, figure 1 shows an acoustic waveform and articulatory movements for the utterance 'pea splots' ([pi 'splats]) as spoken by a speaker of American English. The vertical displacement of pellets on the blade, mid, and rear of the tongue and on the lower lip and lower teeth (i. e. jaw) is shown (after correcting for head movement using reference pellets placed on the nose and the upper incisor). The two bilabial closure gestures (for the [p]'s) can be seen in the movement of the lip. The alveolar fricative gesture (for [s]) and the alveolar lateral gesture (for [l]) can be seen in the movement of the tongue blade.

While we initially intended to examine the relationship between the achievement of target for the initial consonantal gestures (in words like 'splot') and the achievement of target for the vocalic gestures, a reinspection of both the horizontal and vertical movements of all the tongue pellets convinced us that an attempt to identify the vocalic movement with a single dimension of a single pellet was too gross an approximation for the present analysis. This is because the gestures associated with [s] and [l] influ-

Fig. 1. Utterance 'pea splots' produced by a speaker of American English, taken from the X-ray microbeam data base. The top channel shows the acoustic waveform and the lower five channels show the vertical position of pellets attached to the blade, mid and rear portions of the tongue, the lower lip and the jaw, over a period of approximately 1 s.



ence the same tongue pellets that could be expected to reflect the formation of constrictions for [i] and [a]. Therefore, instead of attempting to decompose the consonant and vowel effects on these pellets, we examined the relationship between the intervocalic consonantal gesture(s) and the transvocalic consonantal gestures. For example, in the utterance [pi 'splats], we related the cluster [spl] to the final [t] in [splats] (for the second syllable) and to the word-initial [p] in [pi] (for the first syllable). Thus, we restricted our analyses to the vertical movements of the pellets on the blade (for [s], [l], and [t]) and lower lip (for [p]).

We first examined the relationship of the intervocalic consonantal gestures to the second syllable in the case where the gestures were all word-initial. That is, in the example [pi 'splats], we examined the relationship between the gestures associated with [spl] and [t] in [splats]. Figure 2 shows the acoustic waveform and vertical displacements of the blade and lower lip pellets for individual tokens of utterances of the form [pi# 'C(C)(C)ats]. The tokens are lined up

at a point on the movement curve of the blade pellet that corresponds to the attainment of the target, and also to the acoustic closure, for the [t] of [ats], as indicated by the long line on the right extending the length of the figure. This lineup point will hereafter be called the 'anchor point'. Within each display, the filled portions lie under the plateaus for the peak displacement of the consonantal gestures. The edges of the plateaus were derived automatically by an algorithm that found the peak displacement of each trace, and then defined the plateau as the region of the curve within four X-ray units of displacement (about 1.3 mm) from the peak.

Looking down the displays in figure 2, note that the left edges of the plateaus for the initial (leftmost) consonant gesture in each sequence do not line up. That is, they do not bear a constant relationship to the anchor point. Likewise, the right edges of the plateaus for the rightmost consonants in the intervocalic sequences do not line up. None of the points associated with the individual consonant gestures line up satisfac-

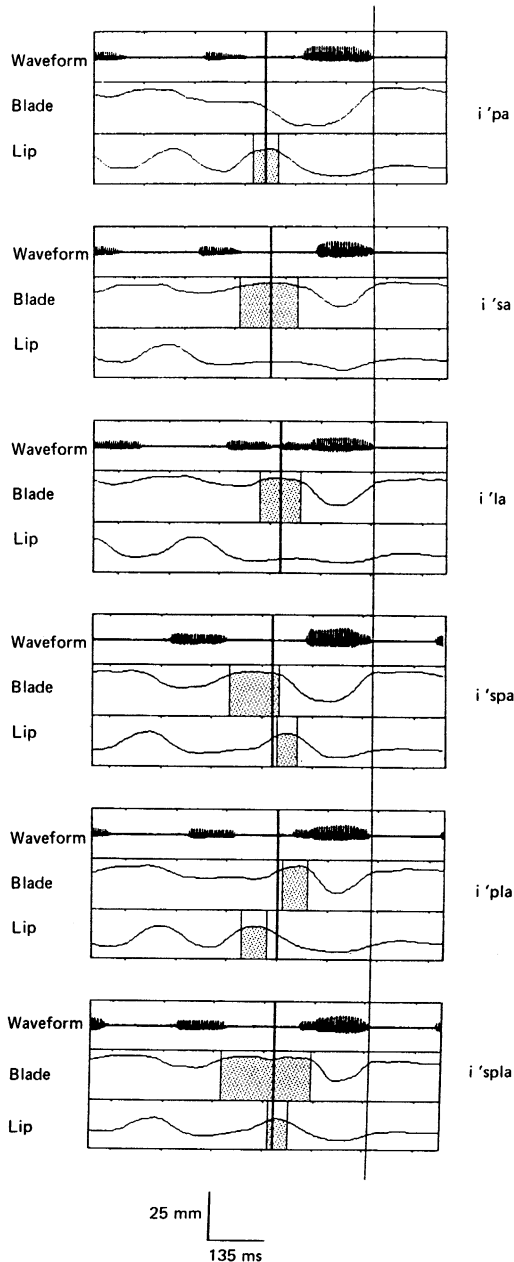


Fig. 2. Acoustic waveform and vertical displacement of pellets attached to the blade of the tongue and the lower lip for utterances of the form /pi#'(C)(C)ats/, with the C-centers for the consonant sequences marked in each display. The utterances are aligned with respect to the anchor point for the second word, the line extending the length of the figure on the right.

torily. However, the dark lines approximately in the middle of the filled portions, and extending the length of each token display, do line up fairly well. These lines represent what we call the *C-center* (consonant center) of each sequence of consonantal gestures. It is a simple global property of this sequence that was derived as follows: for every consonantal gesture, the (temporal) midpoint between the left and right edges of the plateau was computed. The C-center of a sequence is the mean of all the midpoints of the gestures in that sequence.

The C-center seems to be more tightly related than any individual consonant gesture to the anchor point. To provide a quantitative measure for the stability of the C-center, we computed its standard deviation across the utterances. That is, we measured the temporal distance from the anchor point to the C-center and then calculated the standard deviation of this distance across the six utterances. For the C-center, the value is 15.8 ms. Since the pellet movement records in fact represent discrete measures, where a frame equals 6.76 ms, the standard deviation for the C-center is equivalent to 2.4 frames. This value is considerably less than either the standard deviation for the distance from the left edge of the first consonant of the sequence to the anchor point – 37.7 ms – or the standard deviation for the distance from the right edge of the last consonant to the anchor point – 33.6 ms. Thus, the global property of the C-center, applied to syllable-initial sequences,

seems to provide the most stable measure of the relation between those sequences and the first final consonant in the same syllable (i.e., the anchor point).

Figure 3 shows the comparable utterances for the reverse vowel pattern: [pa# 'C(C)(C)idz] ([spl] has different final consonants: [nz]). Once again, the same pattern holds. The left and right edges of the consonant sequences line up rather poorly (standard deviations are 30.1 and 34.2 ms, respectively), while the C-center lines up better (SD = 17.8 ms). While we have only looked at a single token of each type, the fact that the [a-i] utterance and the [i-a] utterances behave similarly suggests that, as we look at more tokens, the C-center should continue to be a stable metric.

Thus, at least word-initially, a sequence of consonant gestures seems to form an aggregate such that its relation to the anchor point of the syllable with which it is affiliated is better (more stably) measured by the single global metric of the C-center than by metrics based on the boundaries of individual consonantal gestures. Does the C-center capture interword relations as well as these intrasyllabic relations? That is, does the C-center provide as stable a measure of the relationship between the same word-initial clusters and the preceding word, with which they are not affiliated? The answer appears to be 'no'.

Figure 4 shows the same utterances as figure 2 ([pi# 'C(C)(C)ats]), this time lined up at a new anchor point, the midpoint of

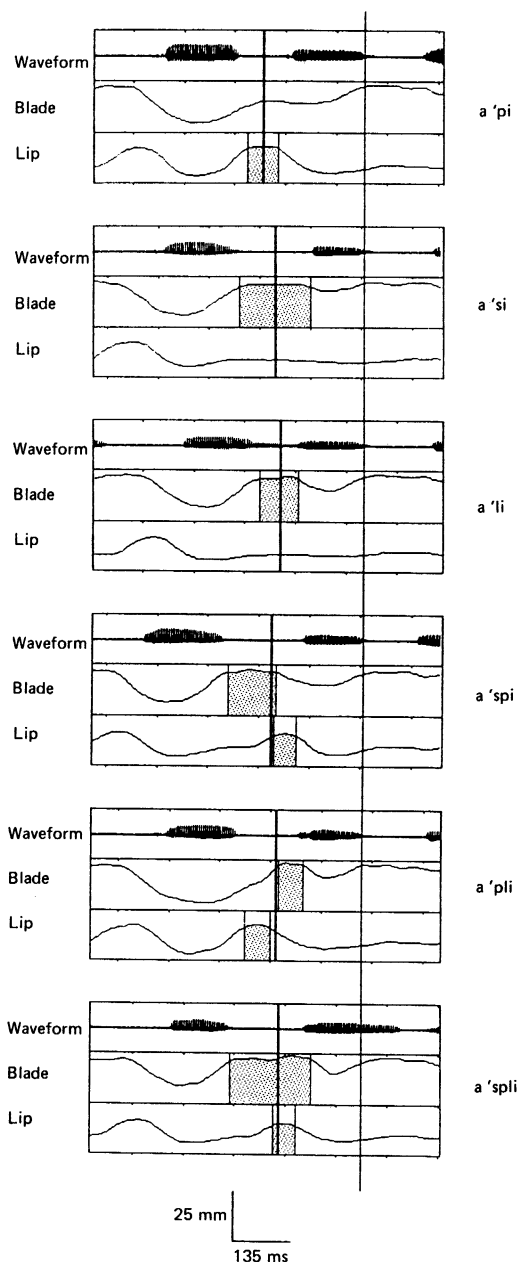


Fig. 3. Acoustic waveform and vertical displacement of pellets attached to the blade of the tongue and the lower lip for utterances of the form /pa# 'C(C)(C)idz/ (exception: /pa# 'splinz/), with the C-centers for the consonant sequences marked in each display. The utterances are aligned with respect to the anchor point for the second word, the line extending the length of the figure on the right.

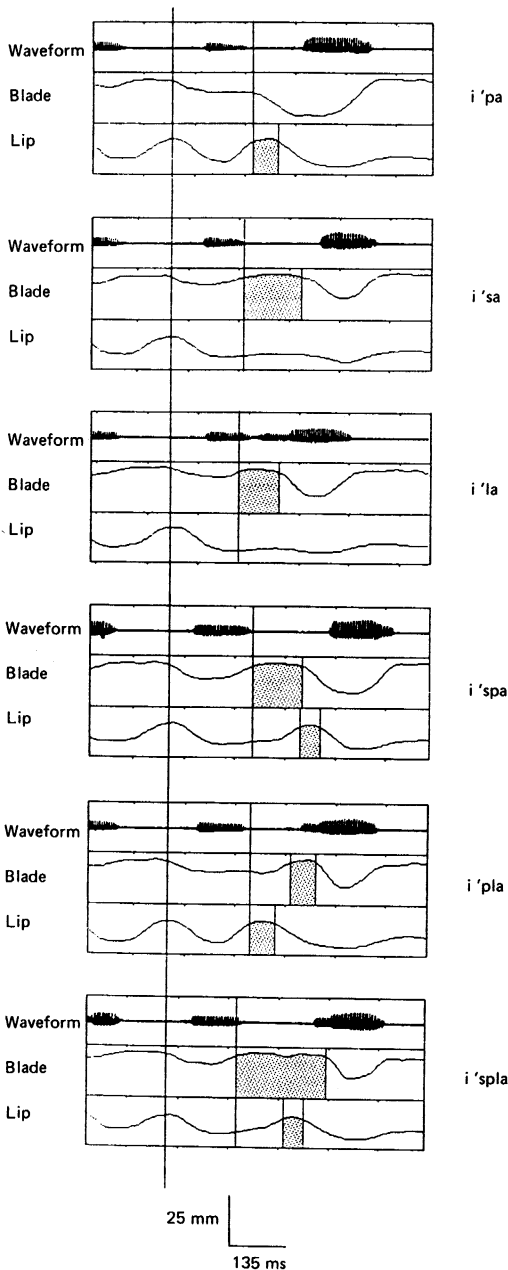


Fig. 4. Acoustic waveform and vertical displacement of pellets attached to the blade of the tongue and the lower lip for utterances of the form /pi# 'C(C)(C)ats/, with the beginning of the plateau of peak displacement for the leftmost consonantal gesture of the consonant sequences marked in each display. The utterances are aligned with respect to the anchor point for the first word, the line extending the length of the figure on the left.

the bilabial closure gesture associated with the word-initial consonant [p] of the first word [pi] (indicated by the line extending the length of the figure on the left).

The midpoint was chosen as the anchor point to be consistent with the finding that initial consonants are measured in terms of C-centers; the midpoint for the bilabial gesture associated with [p] is effectively the C-center for this word-initial single consonant. Since the word-initial consonant is a [p] in all the utterances, the choice of a center rather than boundary lineup point should not affect the results except for a constant offset.

Here it is the relation between the anchor point and the left edge of the first intervocalic consonant (defined as in fig. 2) that is more stable (SD = 18.5 ms). The C-center, with a SD = 43.4 ms, is much more variable. A similar pattern can be seen in figure 5 for [pa# 'C(C)(C)idz] (corresponding to fig. 3). Again, the left edge of the intervocalic consonant sequence lines up better (SD = 17.9 ms) than the C-center (SD = 28.0 ms). Note that the greater variability of the C-centers is not just a consequence of their greater distance from the anchor point, since the C-centers in figures 2 and 3 are farther than the right edges from the anchor point, but there the C-centers show less variability rather than more. Thus, a local rather than global property – the achievement of target of the leftmost consonant – is apparently the most stable measure of the relation between a word-initial consonant sequence and the initial consonant of a preceding word (the anchor point).

To what extent does this relationship depend on the individual consonants' word (or syllable) affiliation? To answer this, we looked at the relationship between the consonant sequence and the preceding word in cases where the first word is a closed syllable rather than an open syllable, i.e., the first consonant in the sequence is affiliated with the preceding word. That is, rather than looking at the sequences such as [pi 'splats] that were explored in figures 2–5, we looked at sequences such as [pis 'plats]. Based on the results so far, there are two possible predictions for the relationship of the [s] to the anchor point [p] in [pis]. If syllable or word affiliation is the organizing principle, then we would expect the most stable metric for the final [s] to be the C-center, since we found that the C-center appeared to be the best metric for affiliated consonants (fig. 2, 3). If, however, sequential order, i.e. prevocalic or postvocalic, is the organizing principle, then we would expect the left edge of the [s] to provide the most stable metric, since that is what we found to be the relationship between the anchor point and following consonants (fig. 4, 5).

Figures 6 and 7 display the movement curves for [piC#'(C)(C)ats] and [paC#'(C)(C)idz], respectively, with the anchor point defined in the same way as for figures 4 and 5. For these utterances, the left edges of the consonants line up somewhat better (SD = 12.1 and 11.6 ms) than the C-centers (SD = 15.3 and 17.8 ms) of

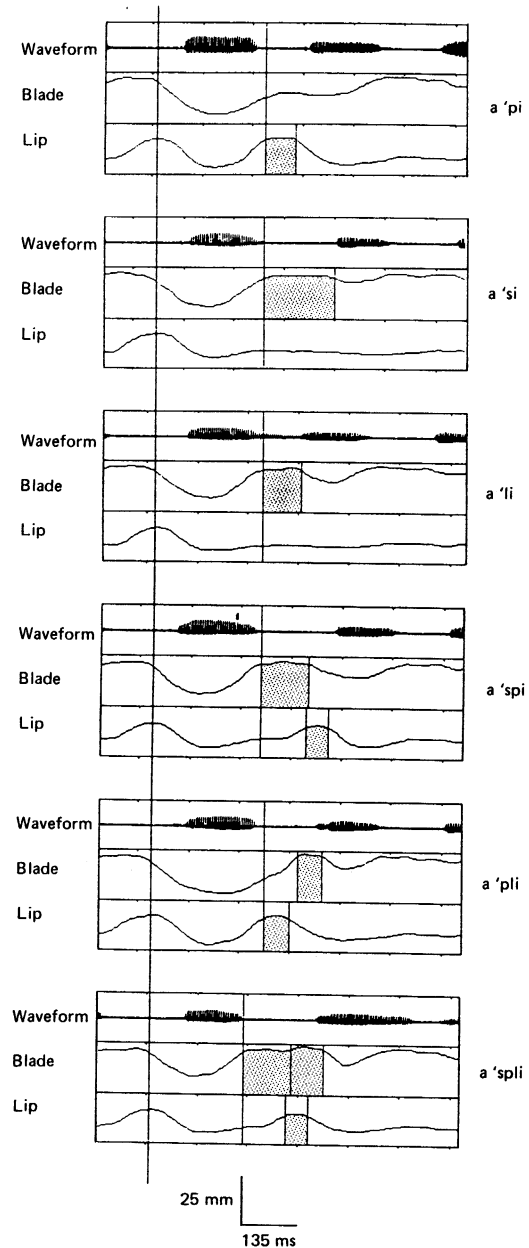


Fig. 5. Acoustic waveform and vertical displacement of pellets attached to the blade of the tongue and the lower lip for utterances of the form /pa#'(C)(C)idz/ (exception: /pa#'splinz/), with the beginning of the plateau of peak displacement for the leftmost consonantal gesture of the consonant sequences marked in each display. The utterances are aligned with respect to the anchor point for the first word, the line extending the length of the figure on the left.

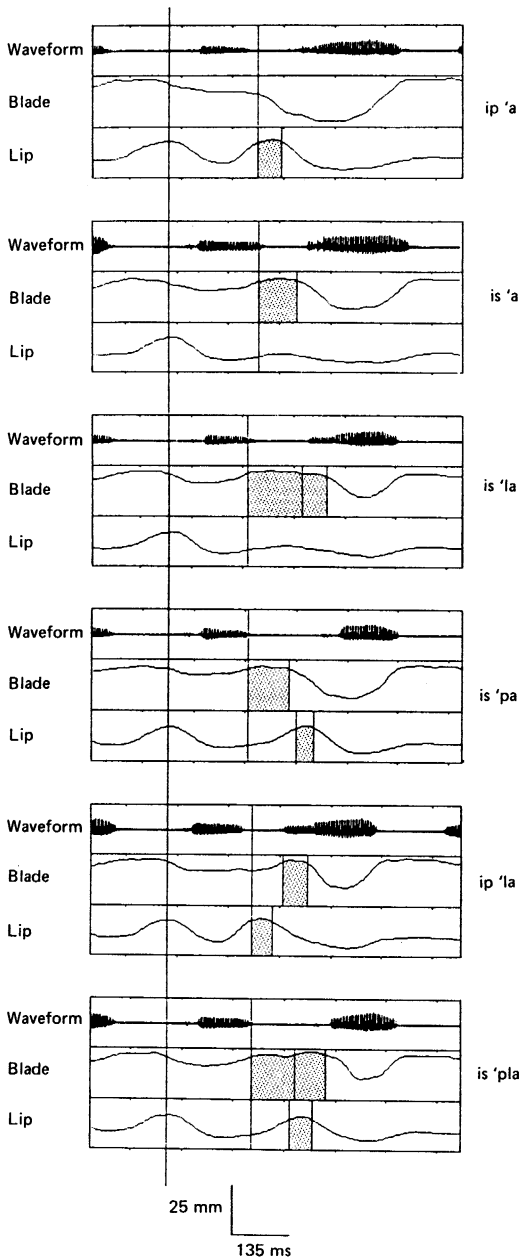


Fig. 6. Acoustic waveform and vertical displacement of pellets attached to the blade of the tongue and the lower lip for utterances /pip# 'adz/, /pis# 'adz/, and /piC# 'C(C)ats/, where C=/s/, /p/ or /l/, with the beginning of the plateau of peak displacement for the leftmost consonantal gesture of the consonant sequences marked in each display. The utterances are aligned with respect to the anchor point for the first word, the line extending the length of the figure on the left.

the single final consonants. (The C-centers of the entire sequence, i.e., the syllable-final consonant plus all following syllable-initial consonants, are even more variable; SD = 25.9 and 26.9 ms.) Note that the differences are small between the two metrics. This is to be expected since two points in the same consonantal gesture are being measured. The only source of difference in this case is the duration differences ([s] is longer than [p]), which would be large enough to result in a noticeable shift in the left edges, if it were the C-center of this consonant that were aligned with the anchor point. Such shifts between the left edges of [s] and [p] can be seen in the syllable-initial cases in figures 2 and 3, where the C-center is the most stable metric, but not in the syllable-final cases in figures 6 and 7. Thus, it appears that postvocalic consonants are organized on the basis of their sequential relation to the vowel rather than on the basis of their syllable affiliation. That is, postvocalic consonants are apparently organized with respect to their left edges (achievement of target) rather than their C-centers, both within and between words – they use what we have termed a local metric.

There is, however, some effect of consonant affiliation on the organization of postvocalic consonants. The organization of consonants ending closed syllables (the within-word consonants of fig. 6 and 7) differs in two ways from that of syllable-initial consonants following open syllables (the

between-word consonants of fig. 4 and 5). First, the standard deviations are greater between words than within words ([a-i] 18.5 vs. 11.6 ms; [i-a] 18.0 vs. 12.1 ms). Second, at least for [a-i], the first consonant following the open syllables occurs later than the consonant in the closed syllables (open: 286 ms vs. closed: 256 ms), as would be expected from acoustic measurements [Fowler, 1983; Maddieson, 1985]. For [i-a], however, the relation is reversed (open: 205 ms vs. closed: 223 ms). In general, then, interword relations are more variable than intraword relations [see also Hardcastle and Roach, 1979], at least when the two words are in different syntactic phrases as in the utterances we examined (noun phrase-verb phrase). Such an effect is partially consistent with the model of the control of speech production proposed by Sternberg et al. [this volume], which suggests that the effects of selecting a unit should be localized at the end of the preceding unit; however, Sternberg et al. predict the effect should always be that of increasing the duration.

To summarize, it seems that within words or syllables, postvocalic (syllable-final) consonants may behave differently than prevocalic (syllable-initial) ones. While initial consonants are related to their words in terms of a single global metric for the entire cluster, the C-center, final consonants appear to be related to their words in terms of the local metric of achievement of target (left edges in the figures) rather than in terms of C-centers. This difference, if

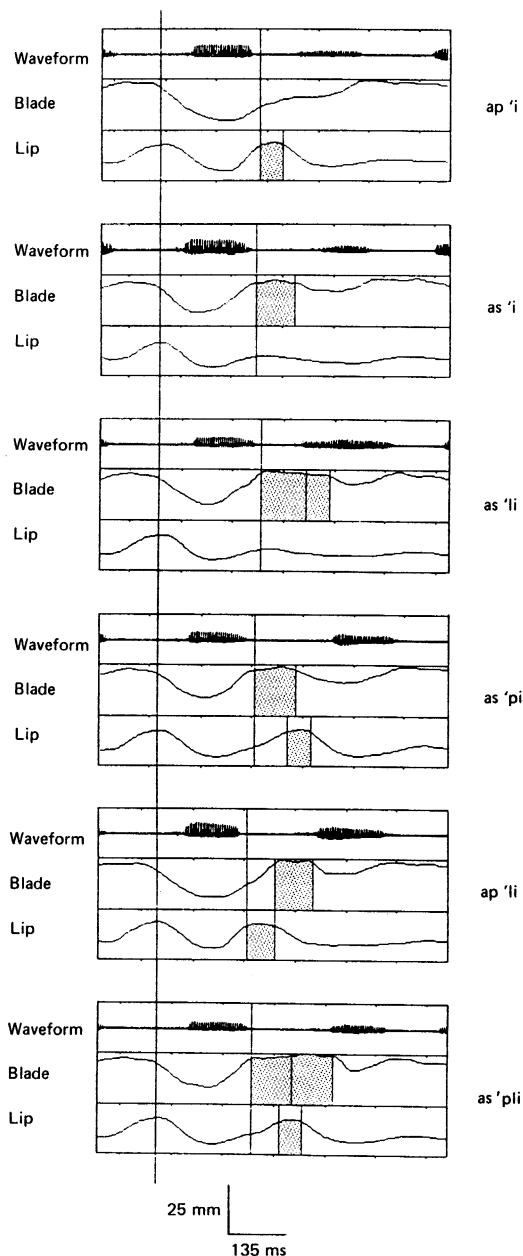


Fig. 7. Acoustic waveform and vertical displacement of pellets attached to the blade of the tongue and the lower lip for utterances /pap# 'its/, /pas# 'its/, and /paC# 'C(C)idz/, where C=/s/, /p/ or /l/, with the beginning of the plateau of peak displacement for the leftmost consonantal gesture of the consonant sequences marked in each display. The utterances are aligned with respect to the anchor point for the first word, the line extending the length of the figure on the left.

confirmed by further analyses, may account for some of the differing phonological properties of initial and final consonant sequences.

Possible Implications for Perceptual and Phonological Patterns

The relevance of the C-center in intrasyllabic articulatory organization, particularly when contrasted with the differing metric for postvocalic consonants, might help explain a variety of apparently disparate phenomena that have been observed in the perceptual and phonological patterning of speech. In order to explore these possible implications, let us make the following assumption: the (temporal) interval from the C-center to the final consonant anchor point is a measure of the activation interval of the vocalic gesture, where (a) the C-center corresponds to a fixed point early in the vocalic activation and (b) the final consonant anchor point, which is the effective achievement of the target for the final consonant, corresponds to the end of activation of the vowel, i. e., articulatory vowel offset [Browman and Goldstein, in press]. We also assume that the actual movement for the vocalic gesture begins at the achievement of target of the first consonant in a possible initial cluster [Browman and Goldstein, in press; Borden and Gay, 1979; Kozhevnikov and Chistovich, 1965]. Under these assumptions, then, the C-center serves to coordinate an initial consonant cluster with the vowel, but for final consonants, it is the achievement of target of the leftmost consonant (the anchor point in the above analyses) that is coordinated with the vowel. The consonants in an initial cluster are dis-

persed around the C-center, thereby overlapping the vowel on the one hand, and pushing the onset of the syllable to the left of the C-center on the other hand. The more consonants in the initial cluster, the shorter the acoustic realization of the vowel and the longer the whole syllable.

The basic relation among the consonants would be that proposed in Browman and Goldstein [in press], although that rule ($C1(\text{offset}) = C2(\text{onset})$) needs to be refined to include compression and syllable affiliation effects, as well as possible effects of the specific articulators used. The C-C relations in the current data appear to be a combination of several tendencies that cannot be sorted out without more data.

Note that the C-center organization only accounts for acoustic shortening of the vowel that is attributable to word-initial consonants. That is, since the first postvocalic consonant always achieves its target at the same point with respect to the vowel in our current account, there is no automatic increase in the amount of overlap as the number of consonants in the final cluster increases. However, Fowler et al. [1986] have shown that, as the number of final consonants increases from one to two (e.g., [aep] vs. [aepts]), the target of the first consonant is achieved earlier. In our current formulation, this timing shift would be handled similarly to the timing difference between closed and open syllables, that is, within and between words. It would be intriguing to see if such timing differences are similar to the phrase-final patterns reported by Edwards and Beckman (this volume).

C-Centers and p-centers.

The C-center pattern that is observed in these data is reminiscent of the perceptual-center (p-center) effects that have been re-

ported [Morton et al., 1976; Fowler and Tassinary, 1981]. When asked to align a list of words with different initial consonants in a regular rhythm, listeners do not align the successive words at their acoustic onsets. Conversely, words that are actually aligned at their onsets sound anisochronous to listeners. The perceptually synchronous lineup point seems to occur during the initial consonant sequence, though the particular point varies with the length of the sequence. In particular, the longer the initial consonant sequence, the later the lineup point in that sequence. Figure 8 [adapted from Fowler and Tassinary, 1981] shows a set of syllables of the form [C(C)(C)ad] that talkers lined up with a metronome beat. The vertical line in the center indicates this beat, while the connected points on the far right are the closure for the syllable-final [d]. This perceptual p-center lineup point bears a close qualitative resemblance to the articulatory C-centers of figures 2 and 3, in that both the p-center and C-center are equidistant from the closure of the syllable-final stop, regardless of the number of consonants in the initial sequence. Moreover, the longer the consonant sequence, the later both the p-center and the C-center occur with respect to the sequence onset.

Fowler's [1983] analysis of these data is that the p-center corresponds to some fixed point in the articulatory cycle of the vowel. This is consistent with the analysis of C-centers we have given above, namely, that the C-center is fixed with respect to the vowel gesture. If this is indeed the case (recall that we did not directly measure the vowel gesture in our data), then (some of) the details of the p-center effect would follow directly from the C-center organization. Moreover, it is interesting to note that

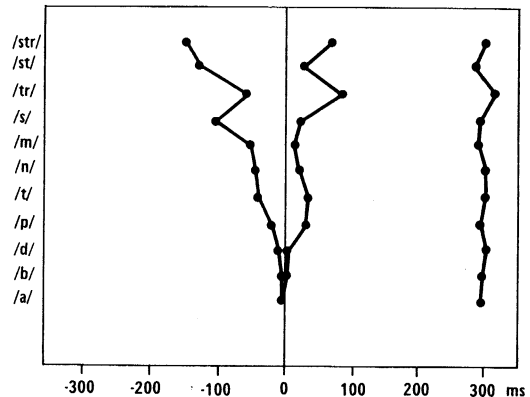


Fig. 8. Measures of syllables produced by talkers in time with a metronome. The vertical line at zero represents the metronome pulse. Different syllables are plotted top to bottom in the figure. The points generally to the left of the line represent the onset of acoustic energy for each syllable relative to the metronome pulse. Points generally just to the right of the pulse represent the measured vowel onset, that is, the onset of voiced oral formants for the vowel. Points to the far right represent measured vowel offset – the beginning of closure for final /d/. Adapted from Fowler and Tassinary, 1981.

changing the syllable-final consonants has a much smaller effect on the p-center [Marcus, 1981]. This might be accounted for by the V-C organization we have proposed here. Since the nature of the final consonant does not affect the relation of the vowel to the first following consonant, changes in final consonant duration would not be expected to change the relation between the acoustic closure for that consonant and the articulatory cycle of the vowel. Thus, assuming again that something about the articulatory vowel gesture is important to p-centers, the final consonants do not interact with this vowel gesture in the way that initial consonants do.

search are suggested by these proposals. The basic patterns proposed for C-V and V-C organizations need to be further tested, particularly by analyzing syllable-final consonant sequences, both with and without phonetic affixes. The apparent similarity of initial C-centers and p-centers bears further investigation, as does the prediction of universality of the C-center metric for syllable-initial consonants. Finally, there were several small but apparently regular differences in the data associated with the phonetic identity of both consonantal and vocalic gestures. For example, the bilabial gestures in initial position occur a little early, the lateral gestures a little late – perhaps the former means that glottal gestures should be included in the determination of the C-center, and the latter reflects the cohesiveness of laterals with vowels. And there appear to be systematic differences in the timing from [a] to [i] and [i] to [a], with the movement from [a] to [i] starting earlier relative to the intervocalic consonant [this has been observed in Japanese as well by Kiritani et al., 1977]. Thus, more work is needed, both to test and to refine the generalizations proposed here.

Acknowledgements

Our thanks to Carol Fowler, Eric Vatikiotis-Bateson, Eliot Saltzman, Caroline Smith, Kathy Harris, and Osamu Fujimura, who critiqued versions of this paper, but who should not be held responsible for any remaining conceptual problems, and to Caroline Smith and Mark Tiede for help in manuscript preparation. This work was supported by NSF grant BNS 8520709 and NIH grants HD-01994 and NS-13617 to Haskins Laboratories.

References

- Abbs, J.H.; Gracco, V.L.; Cole, K.J.: Control of multimovement coordination: sensorimotor mechanisms in speech motor programming. *J. Motor Behav.* 16: 195–231 (1984).
- Borden, G.; Gay, T.: Temporal aspects of articulatory movements for /s/-stop clusters. *Phonetica* 36: 21–31 (1979).
- Browman, C.P.; Goldstein, L.: Towards an articulatory phonology. *Phonol. Yb.* 3 (1986).
- Browman, C.P.; Goldstein, L.: Tiers in articulatory phonology, with some implications for casual speech; in Beckman, Kingston, *Papers in laboratory phonology I: Between the grammar and the physics of speech* (Cambridge University Press, Cambridge, in press).
- Browman, C.P.; Goldstein, L.; Saltzman E.; Smith C.: GEST: computational model for speech production using dynamically defined articulatory gestures (Abstract). *J. acoust. Soc. Am.* 80: S97 (1986).
- Clements, G.N.; Keyser S.J.: CV phonology: a generative theory of the syllable (MIT Press, Cambridge 1983).
- Fowler, C.A.: Converging sources of evidence on spoken and perceived rhythms of speech: cyclic production of vowels in sequences of monosyllabic stress feet. *J. exp. Psychol.* 112: 386–412 (1983).
- Fowler, C.A.; Munhall, K.; Saltzman, E.; Hawkins, S.: Acoustic and articulatory evidence for consonant-vowel interactions (Abstract). *J. acoust. Soc. Am.* 80: S96 (1986).
- Fowler, C.A.; Tassinari L.: Natural measurement criteria for speech: the anisochrony illusion; in Long, Baddeley, *Attention and performance*, vol. IX (Erlbaum, Hillsdale 1981).
- Fujimura, O.; Kiritani, S.; Ishida, H.: Computer controlled radiography for observation of movements of articulatory and other human organs. *Comput. Biol. Med.* 3: 371–384 (1973).
- Fujimura, O.; Lovins J.: Syllables as concatenative phonetic units; in Bell, Hooper, *Syllables and segments* (North Holland, Amsterdam 1978).
- Halle, M.; Vergnaud, J.R.: Three-dimensional phonology. *J. ling. Res.* 1: 83–105 (1980).
- Hardcastle, W.J.; Roach, P.J.: An instrumental investigation of coarticulation in stop consonant sequences; in Hollien, *Current issues in*

- the phonetic sciences (Benjamins, Amsterdam 1979).
- Hayes, B.: Extrametricality and English stress. *Ling. Inquiry* 13: 227–276 (1982).
- Hoequist, C.E.: The perceptual center and rhythm categories. *Lang. Speech* 26: 367–376 (1983).
- Hyman, L.M.: A theory of phonological weight (Foris Dordrecht 1985).
- Keating, P.A.: CV phonology, experimental phonetics, and coarticulation. *UCLA Working Papers Phonet.* 62: 1–13 (1985).
- Kelso, J.A.S.; Tuller, B.: Intrinsic time in speech production: theory, methodology, and preliminary observations. *Haskins Lab. Status Rep. Speech Res.*, SR-81, pp. 23–39 (Haskins Laboratories, New Haven 1985).
- Kelso, J.A.S.; Tuller, B.; Vatikiotis-Bateson, E.; Fowler, C.A.: Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. *J. exp. Psychol. hum. Perception Performance* 10: 812–832 (1984).
- Kiritani, S.; Itoh, K.; Hirose, H.; Sawashima, M.: Coordination of the consonant and vowel articulations: X-ray microbeam study of Japanese and English. *Annu. Bull. RILP* 11: 11–21 (1977).
- Kozhevnikov, V.; Chistovich, L.: Speech: articulation and Perception. *Transl. of Rech', Artikulyatsiya, i. Vospriyatiye. Report* 30, p. 543 (Joint Publications Research Service, Washington 1965).
- Ladefoged, P.: A course in phonetics; 2nd ed. (Harcourt Brace Jovanovich, New York 1982).
- Maddieson, I.: Phonetic cues to syllabification; in Fromkin, *Phonetic linguistics: essays in honor of Peter Ladefoged* (Academic Press, New York 1985).
- Marcus, S.M.: Acoustic determinants of perceptual centers (P-center) location. *Perception Psychophysics* 30: 247–256 (1981).
- McCarthy, J.J.: A prosodic theory of nonconcatenative morphology. *Ling. Inquiry* 12: 373–418 (1981).
- Miller, J.E.; Fujimura, O.: Graphic displays of combined presentations of acoustic and articulatory information. *Bell Syst. tech. J.* 61: 799–810 (1982).
- Morton, J.; Marcus, S.; Frankish, C.: Perceptual centers (P-centers). *Psychol. Rev.* 83: 405–408 (1976).
- Nittrouer, S.; Munhall, K.; Kelso, J.A.S.; Tuller, B.; Harris, K.S.: Patterns of interarticulator phasing relations (Abstract). *J. acoust. Soc. Am.* 80: S109 (1986).
- Öhman, S.E.G.: Coarticulation in VCV utterances: spectrographic measurements *J. acoust. Soc. Am.* 39: 151–168 (1966).
- Öhman, S.E.G.: Numerical model of coarticulation. *J. acoust. Soc. Am.* 41: 310–320 (1967).
- Saltzman, E.: Task dynamic coordination of the speech articulators: a preliminary model; in Heuer, Fromm, *Generation and modulation of action patterns. Experimental Brain Research Series* 15 (Springer, New York 1986).
- Saltzman, E.; Kelso, J.A.S.: Skilled actions, a task dynamic approach. *Haskins Lab. Status Rep. Speech Res.*, SR-76, pp. 3–50 (Haskins Laboratories, New Haven 1983).
- Saltzman, E.; Rubin, P.E.; Goldstein, L.; Browman, C.P.: Task-dynamic modeling of interarticulator coordination (Abstract). *J. acoust. Soc. Am.* 82: S15 (1987).
- Sussman, H.M.; MacNeilage, P.F.; Hanson, R.J.: Labial and mandibular dynamics during the production of bilabial consonants: preliminary observations. *J. Speech Hear. Res.* 16: 397–420 (1973).
- Tuller, B.; Kelso, J.A.S.: The timing of articulatory gestures: evidence for relational invariants. *J. acoust. Soc. Am.* 76: 1030–1036 (1984).
- Tuller, B.; Kelso, J.A.S.; Harris, K.S.: Interarticulator phasing as an index of temporal regularity in speech. *J. exp. Psychol. hum. Perception Performance* 8: 460–472 (1982).

Catherine P. Browman
Haskins Laboratories
270 Crown Street
New Haven, CT 06511 (USA)