The Timing of Lip Rounding and Tongue Backing for /u/

Gina M. Lee

Department of Linguistics, The Ohio State University

Abstract

A small corpus of X-ray microbeam data was examined to test the predictions made by two well-known views of anticipatory coarticulation: time locking and feature spreading. Lip rounding and tongue backing associated with English /u/ were investigated. The token types were VCnV sequences where V = /i, u/ and C = /s, t/, consonants which are assumed to be neutral with respect to the features under study. The lip and jaw pellets were rotated in such a way that their principal component of movement was in the vertical dimension. In the /iCnu/ tokens there was no evidence for either time locking or feature spreading of the onset of lip rounding. However, other articulatory phenomena are temporally regulated. The timing of both the maximum point of lip rounding and of tongue backing was fixed relative to the acoustic onset of /u/. These results may provide evidence for the idea of fixed targets in speech production.

Introduction

Different accounts of the timing of gestures in coarticulation have been put forth. Under a time locking hypothesis, such as that proposed by Bell-Berti and Harris (1979, 1982), timing is an intrinsic part of the speech motor plan. They found that the onset of EMG activity for lip rounding begins at a fixed interval relative to the onset of the voicing of a rounded vowel. Regardless of the number of (or duration of) preceding consonants, anticipatory activity begins at a point that is temporally fixed. Under this view, units of speech production are dynamic gestures.

Under a feature spreading hypothesis, such as that first proposed in Henke (1966), timing is not specified in the segmental description, but, rather, is extrinsic to the execution of the speech motor plan. Anticipatory activity begins at a point where there are no contradictory demands on the articulators. For /u/ speakers begin the anticipatory rounding gesture at the earliest point possible. The longer the cluster of feature-neutral consonants (or the larger the number of consonants) preceding the rounded vowel, the earlier the onset of rounding with respect to the acoustic period of the vowel. The feature is said to "spread" to the beginning of the consonant string, even as early as six consonants preceding the acoustic onset of the vowel. Under this view, the units of speech production are invariant in nature.

Figure 1, adapted from Perkell (1986: 49), illustrates the difference between these views. Lip protrusion (rounding) is shown as a function of time. The onset of rounding is indicated by a tick mark. The dashed lines represent acoustic boundaries of the intervocalic consonant(s). In (A) time locking, the point of onset relative to the voicing for the rounded vowel is the same regardless of the number of consonants. In (B) feature spreading, the points of onset differ. If the rounding begins at the end of the period for the preceding unrounded vowel, then the longer the string of intervocalic consonants, the earlier the onset of rounding.

To test the predictions made by these two views of coarticulation, the anticipatory gestures associated with /u/ were examined using X-ray microbeam data. Both lip rounding and tongue backing were investigated.

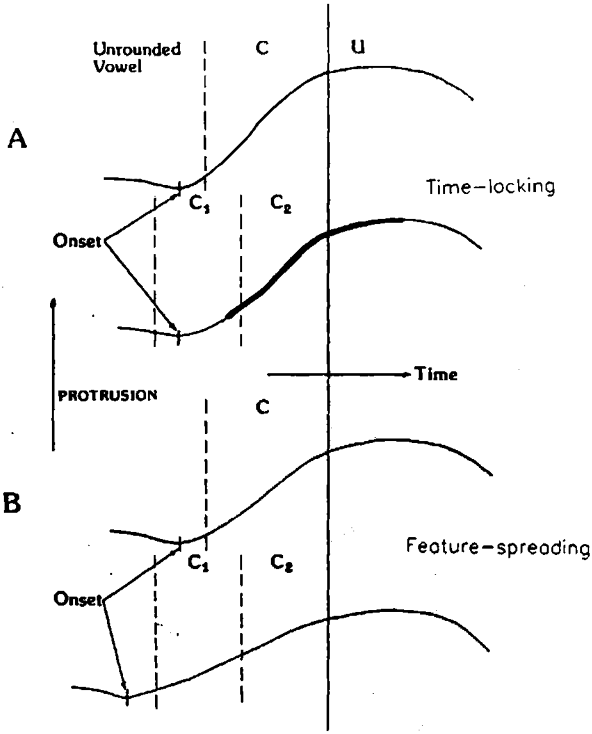


Figure 1. A comparison of time locking and feature spreading (Perkell 1986).

Methods

The tokens were taken from a set of X-ray microbeam data bases (Kiritani et al. 1975, Kiritani 1986) made under the direction of Osamu Fujimura at the University of Tokyo. The utterance types were identical to those used in Bell-Berti and Harris (1979); the data base had been made according to their specifications. The subject was a female native speaker of English. As indicated in Table 1, the token types consisted of VCnV sequences where each vowel was either /i/ or /u/. The consonants were /s/, /t/ and various combinations of the two up to a maximum of four consonants. For some token types only the placement of the word boundary differed, as in lee stool /li#stul/ and lease tool /lis#tul/. All sequences were embedded within actual English words, and were placed in the frame sentence 'It's a _______ again.' Each token type was repeated at least twice. There were thirteen /iCnu/ tokens and six of each of the other token types /iCni/, /uCni/ and /uCnu/ for a total of thirty-one tokens. Each sentence was read slowly with preservation of word boundaries, sometimes to the extent that [ ] could be heard between them, as in one reading of leased tool [list tul]. Across utterances, there was no difference in rate.

Table 1. Token types

lee tool	i # tu	loo tool	u # tu
lee stool	i #stu	loose stool	us #stu
lease tool	is # tu	loosed stool	ust#stu
lease stool	is #stu		
leased tool	ist# tu		
leased stool	ist#stu		
		loo teal	u # ti
lee teal	i # ti	loose steel	us #sti
lease steel	is #sti	loosed steel	ust#sti

I examined the tokens on an AT&T PC-6300 using XD, a display program developed by Joan E. Miller. Figure 2 contains a trajectory display which depicts the movement of the pellets on an x,y grid. The movements of the lower lip and jaw pellets were examined. To ensure that the principal component of movement was in the vertical dimension, I rotated the pellet trajectories 47 degrees. The top portion shows the original display; the bottom portion shows the display after the rotation of the pellets.

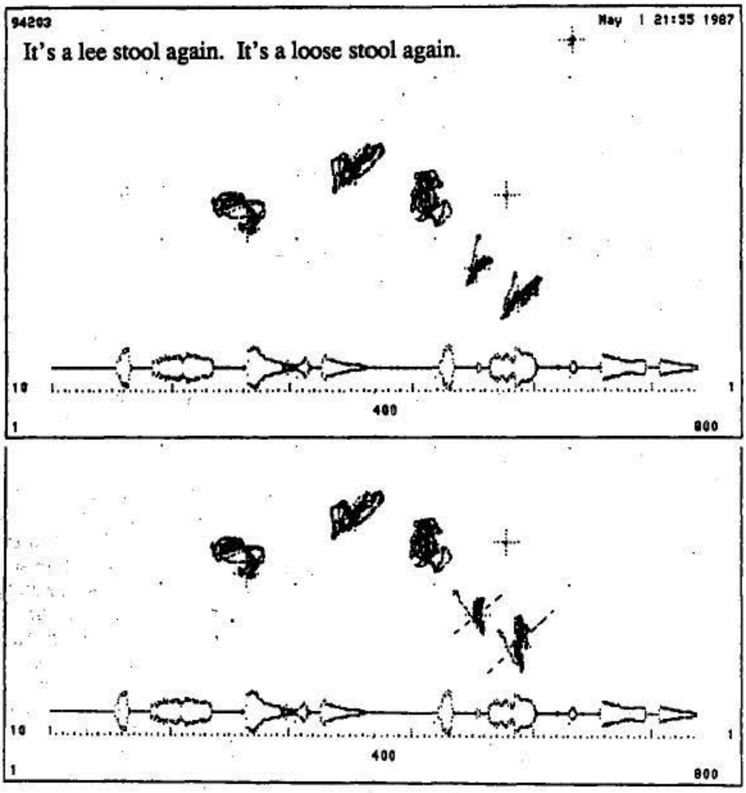


Figure 2. A trajectory display from the XD program. "Before" and "after" diagrams show the effects of rotating the lip and jaw pellets.

Figure 3 contains a trace display which depicts the vertical movement of the rotated pellets as a function of time. Since the lower lip and jaw are mechanically linked, there is a possible effect of confoundment of lip and jaw raising in the movement of the lower lip pellet. To determine the movement of the lower lip alone, I subtracted the value of jaw movement from the combined trajectory value of the jaw and lower lip movement, resulting in a trace which reflected, as much as possible, pure lip movement. In the diagram, this trace is represented as 'U1'.

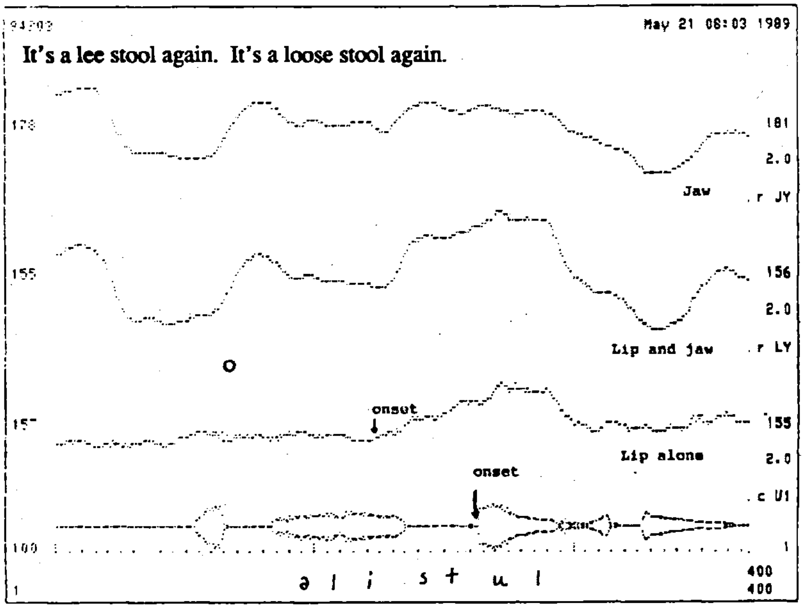


Figure 3. Traces showing the movement of the lip and jaw pellets, and a user-defined trace showing movement of the lip alone. The utterance is lee stool.

Results

Lip Rounding

Three kinds of measurement were made in the /iCnu/ tokens: the onset of lip rounding, defined as the point in u1 at which the trace begins to rise; the acoustic onset of /u/; the rounding peak within /u/; and the acoustic length of the intervocalic consonant(s). The first three are labeled in Figure 3.

In Figure 4 the onset of lip rounding is plotted against consonant length for the thirteen /iCnu/ tokens. The results showed no evidence for time locking of lip rounding. If the onset is time locked, one would expect to find values clustering within a narrow range. In this study, the values vary substantially in the vertical dimension. Moreover, the relative onsets vary in some tokens of the same consonant length and consonant type. Both /i#stu/ type tokens, labeled 1 and 2, show widely differing relative onsets (134 and 341.7 ms.) but the length of the /st/

cluster is virtually the same (273.6 and 270.0 ms. respectively). Both /is#stu/ type tokens, labeled 3 and 4, likewise show quite different relative onsets (127.3 and 314.9 ms.) but almost the same consonant length (532.8 and 529.2 ms. respectively). In a case of feature spreading, the relative onset is directly proportional to consonant length, and so we would expect a regression line of the form y=x. But the results clearly do not support this view either.

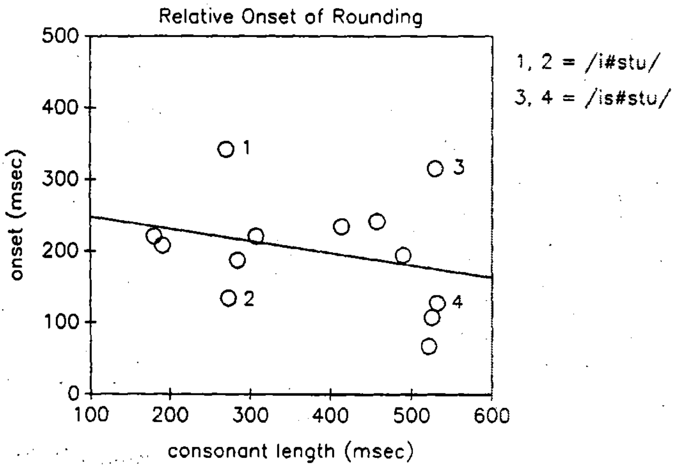


Figure 4. The relative onset of lip rounding plotted against the length of the consonant cluster for the /iCnu/ tokens.

As seen in past studies, the utterances of /uCnu/ type show so-called "troughs" in movement. Figure 5 contains an example. In terms of pure lip movement, the speaker momentarily diminished the degree of lip rounding during the production of the consonants. There were no troughs in the /utu/ type tokens. This may be due to the duration of the intervocalic consonant; the consonant was shorter in /utu/ than in the other /uCnu/ types. It would have been interesting to examine /usu/ type tokens for comparison, but since /sul/ and /ul/ are not actual English words, there were no tokens of this type.

The occurrence of troughs in the conditions specified is consistent with what has been found in earlier investigations. This is problematic for the notion of feature spreading in that there is no clear reason why lip rounding activity during this sequence should diminish during the production of consonants compatible with (and presumably) neutral for lip rounding.

Although no connection between rounding onset and acoustic onset can be seen here, there is a relationship between the acoustic onset of the vowel and the peak value of lip rounding. The peak value was the first cursor position at which the trace reached its maximum value during the production of /u/, as shown in Figure 6. Only twelve /iCnu/ tokens were examined. In one token, the timing of the peak could not be determined. The trajectory display ended shortly after the acoustic onset of the /u/, making it impossible to identify the point of maximum protrusion.

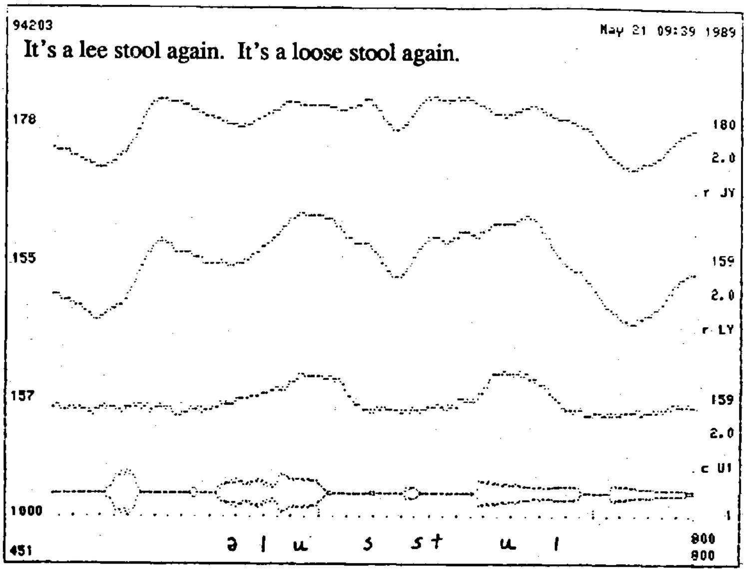


Figure 5. **A "trough" in a /uCnu/ utterance.**

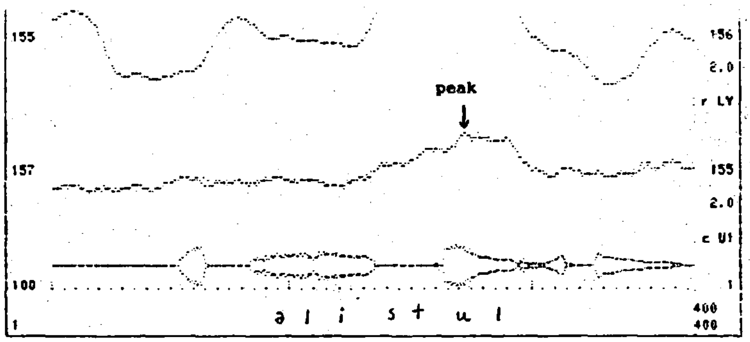


Figure 6. **Measuring the peak of lip rounding.**

As Figure 7 shows, there is no relationship between the peak distance and consonant length. The values of the peak distance range from 87.1 ms. to 180.9 ms. following the onset of voicing. The values generally fall within a narrow range; eleven of the twelve tokens are between 80-140 ms. This suggests that the occurrence of the peak is regulated relative to the acoustic onset of the vowel. The timing of the onset of lip rounding does not matter, as long as the maximum amplitude is reached at a relatively fixed time.

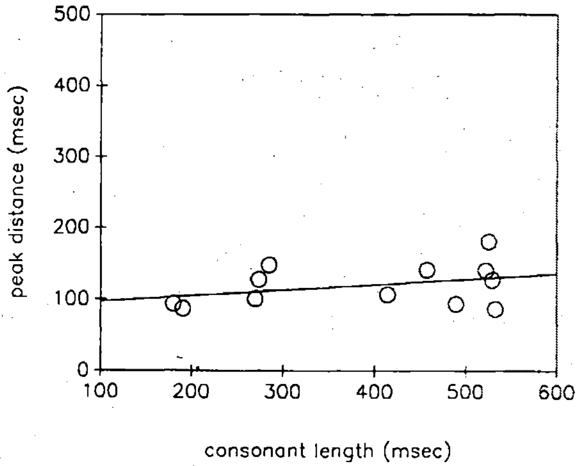


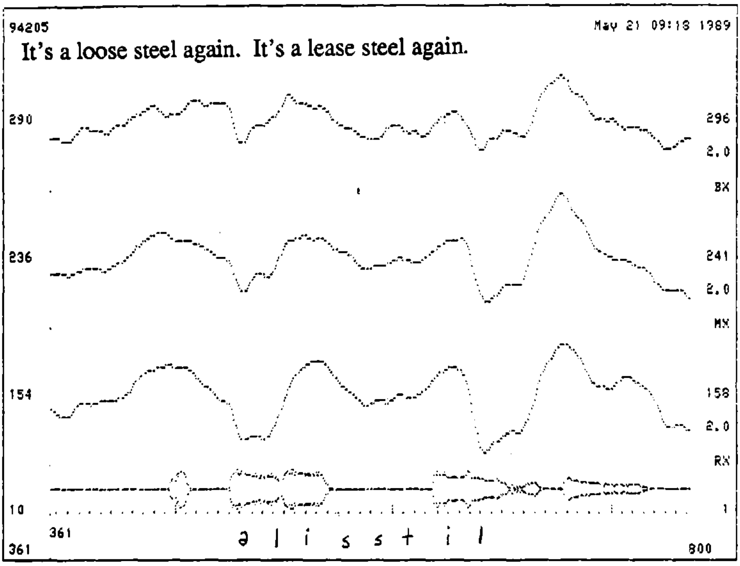
Figure 7. Timing of peak values of lip rounding relative to onset of voicing of the vowel.

<u>Tongue Backing</u>

Because the maximal point of lip rounding for /u/ was found to be "time locked", the question arises as to whether the same phenomenon might be seen in tongue backing as well. Figure 8 contains traces showing horizontal movement of the blade, middle and rear portions of the tongue in an /iCnu/ type utterance. Tongue backing is represented by a lowering of the trace.

Some /iCni/ type tokens showed troughs during the production of the consonants. The tokens with the longer consonant lengths showed more clearly defined troughs than the shorter consonant length. As in the case of the /uCnu/ troughs, these dips are problematic because there is no reason why the trace should lower and then rise during the production of consonants that are neutral with respect to backing of the middle and rear portions of the tongue. Interestingly, the /uCnu/ type tokens showed the opposite: some tokens had "humps", where the middle and rear portions moved slightly more forward in the production of the consonants.

In Figure 9 the timing of the maximum point of backing relative to the acoustic onset of /u/ is plotted against consonant length. As mentioned earlier, one of the tokens could not be analyzed, and only twelve were used. The maximum point was defined as the first position at which the trace was at its lowest. For all /iCnu/ tokens, the maximum point of backing is temporally fixed following the acoustic onset. The values have no connection with consonant length, and fall within a narrow range. The minimum is 187.6 ms. following the onset of voicing; the maximum is 301.5 ms. Eleven of the twelve tokens fall within the range of 210-300 ms. Thus, just as in rounding, the timing of maximal tongue backing was also consistent.



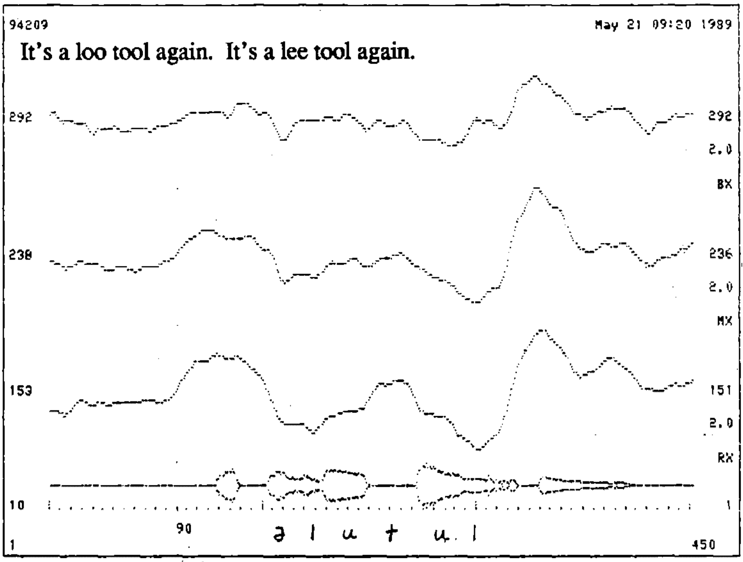


Figure 8. Movement traces of tongue backing in the utterances *lease steel* (top) and *loo tool* (bottom).

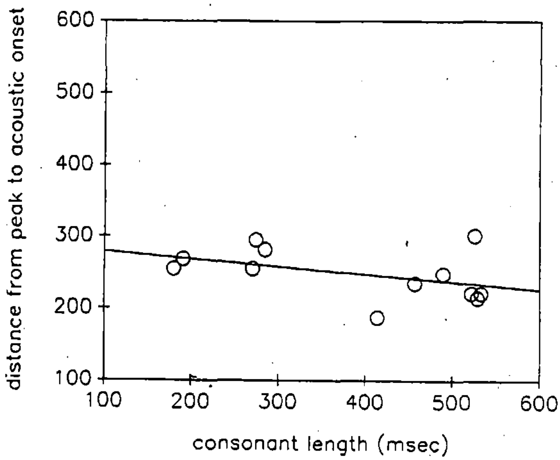


Figure 9. Timing of peak backing relative to onset of voicing of the vowel.

Conclusion

This study is, of course, preliminary in nature. Clearly the small number of tokens makes it difficult to form strong generalizations. Since only one speaker was examined, the results may reflect idiosyncratic characteristics. Further investigation should examine a larger data base from a larger number of speakers and, if possible, include an examination of acoustic and perceptual correlates.

Nonetheless, at least for the speaker in this study the timing of the onset of rounding is not relevant to the organization of the speech motor plan. What is regulated is the timing of the maximum amount of rounding. As was pointed out to me by John Ohala, this result is consistent with earlier EMG studies (Hirose et al. 1968, Hirose et al. 1969) in which EMG peaks occurred at relatively fixed intervals relative to the acoustic period for /u/. In these studies, one speaker consistently had EMG peaks approximately 90 ms. prior to the center of the acoustic period of /u/. If there is a consistent correlation between EMG peaks and movement peaks--with movement peaks typically occurring after EMG peaks--then the results of the current study are not surprising.

These results suggest that there is time locking in articulatory gestures, but of a different sort than the type previously suggested. For the speaker in this study it is not the onset of movement that is time locked, but, rather, the occurrences of the peak values. The results from this study also suggest, too, that speakers aim for fixed articulatory targets, which in this case involves the attainment of maximal points.

Acknowledgments

Earlier versions of this paper were presented at the 113th meeting of the Acoustical Society of America in Indianapolis, IN (1987) and at the 117th meeting in Syracuse, NY (1989), where I received many useful comments. I gained much from discussion with Mary Beckman and other participants in the seminar on articulation and coarticulation. I am also grateful to Jan Edwards for her suggestions.

References

JASA = Journal of the Acoustical Society of America

Bell-Berti, Fredericka and Katherine S. Harris. 1974. RMore on the motor organization of speech gesturesS, *Haskins Laboratories: Status Report on Speech Research* SR 37/38: 73-77.

__________. 1979. RAnticipatory co-articulation: some implications from a study of lip roundingS, *JASA* 65.3: 1268-1270.

__________. 1981. RA temporal model of speech productionS, *Phonetica* 38: 9-20.

__________. 1982. RTemporal patterns of coarticulation: lip roundingS, *JASA* 71.2: 449-454.

Fowler, Carol A. 1980. RCoarticulation and theories of extrinsic timingS, *Journal of Phonetics* 8: 113-133.

Gelfer, Carole E., Katherine S. Harris and Gary Hilt. 1981. RTemporal constraints on anticipatory coarticulationS, Abstract, *JASA Supplement 1*, Vol. 69, Spring: S55.

Gelfer, Carole E., Fredericka Bell-Berti and Katherine S. Harris. 1985. RDetermining the extent of coarticulation: effects of experimental designS, *Haskins Laboratories: Status Report on Speech Research* SR-82/83: 19-31.

Harris, Katherine S., Betty Tuller and J.A. Scott Kelso. 1986. RTemporal invariance in the production of speechS, *Invariance and Variability in Speech Processes*, Joseph S. Perkell and Dennis H. Klatt (eds.): 243-252. Hillsdale, N.J.: Lawrence Earlbaum Associates.

Henke, William L. 1966. *Dynamic articulatory model of speech production using computer simulation*. Ph.D. dissertation, MIT.

Hirose, H., S. Kiritani and S. Shibata. 1968. RAn electromyographic study of articulatory movementsS, *Annual Bulletin No. 2, Research Institute of Logopedics and Phoniatrics*, Univ. of Tokyo.

Hirose, H., Z. Simada and S. Kiritani. 1969. RAn electromyographic study of articulatory movementsS, *Annual Bulletin No. 3, Research Institute of Logopedics and Phoniatrics*, Univ. of Tokyo.

Kiritani, S., K. Itoh and O. Fujimura. 1975. RTongue-pellet tracking by a computer-controlled X-ray microbeam systemS, *JASA* 57.6.II: 1516-1520.

Kiritani, Shigeru. 1986. RX-ray microbeam method for measurement of articulatory dynamics-techniques and resultsS, *Speech Communication* 5: 119-140.

Lubker, J. and Thomas Gay. 1982. RAnticipatory labial coarticulation: experimental, biological, and linguistic variablesS, *JASA* 71.2: 437-448.

Perkell, Joseph S. 1980. RPhonetic features and the physiology of speech productionS, *Language Production* . Brian Butterworth (ed.), Vol. 1: 337-372. Academic Press.

__________. 1986. RCoarticulation strategies: preliminary implications of a detailed analysis of lower lip protrusion movementsS, *Speech Communication* 5: 47-68. North Holland.

Sussman, Harvey M. and John R. Westbury. 1981. RThe effects of antagonistic gestures on temporal and amplitude parameters of anticipatory labial coarticulationS, *Journal of Speech and Hearing Research* 46: 16-24.