Journal of Phonetics (1974), 2, 47–63

Effect of speaking rate on labial consonant-vowel articulation

Thomas Gay,* Tatsujiro Ushijima,† Hajime Hirose† and Franklin S. Cooper

Haskins Laboratories, New Haven, Connecticut 06510, U.S.A.

Received 12*th November* 1973

Abstract: The purpose of this experiment was to study the effect of speaking rate on the articulation of the consonants /p, w/, in combination with the vowels /i, a, u/. Two subjects read a list of nonsense syllables containing /p, w/, in all possible VCV combinations with /i, a, u/, at both moderate and fast speaking rates. EMG recordings from muscles that control movements of the lips, tongue and jaw were recorded simultaneously with high speed lateral view X-ray films of the tongue and jaw, and high speed full-face motion pictures of the lips. For labial consonant production, an increase in speaking rate is accompanied by an increase in the activity level of the muscle (orbicularis oris) and slightly faster rates of lip movement (both closing and opening). Vowel production, however, shows opposite effects: an increase in speaking rate is accompanied by a decrease in the activity level of the genioglossus muscle and, as shown by the X-ray films, evidence of target undershoot. Jaw movement data show more variable, context-dependent effects on speaking rate. Observed differences are explained in terms of the muscle systems involved.

Introduction

The way in which a speaker produces a given string of phones will show a good deal of variability depending upon, among other things, the suprasegmental features of stress and speaking rate. The control of speaking rate is a good illustration of the complex nature of these allophonic variations. For example, it is commonly known that during faster speech, a vowel tends to change colour toward the neutral schwa (Lindblom, 1963, 1964). Lindblom's original model posits that this neutralization is a consequence of the shorter duration of the vowel and is caused by a temporal overlap of motor commands to the articulators. In other words, the articulators fail to reach, or undershoot, their targets because the next set of motor commands deflects them to the following target before the first target is reached. This phenomenon implies further that both the rate of movement of the articulators (specifically, the tongue) and the activity levels of the muscles either remain unchanged, or are decreased during faster speech. Although similar undershoot effects have been observed for other phones (Gay, 1968; Kent, 1970), a general model of speaking rate control based on timing changes alone is too simple. For example, in a recent study of labial consonant production (Gay & Hirose, 1973), it was shown that an increase in speaking rate is accompanied by both an increase in the rate of movement of the lips and an increase in the activity levels of the muscles. Although changes in the timing of commands to the

---

*Also, University of Connecticut Health Center, Farmington, Connecticut 06032, U.S.A.

†Also, University of Tokyo, Tokyo, Japan.

muscles do occur, the production of labial consonants during faster speech is characterized primarily by an increase in articulatory effort.

The implication that more than one mechanism operates to control speaking rate is not necessarily unexpected. Whereas vowel production involves a movement toward a spatial target, the production of most consonants involves a movement towards constrictive or occlusal targets. Thus, in a strict sense, the concept of undershoot itself cannot be easily applied to consonant production. Of course, too, the phenomena described above are based on a number of different experiments; it is quite conceivable that some of the differences observed are individual ones.

This paper represents an attempt to further describe these phenomena by studying the effect of speaking rate on the articulation of labial consonants and both preceding and following vowels. The specific purpose of the experiment was to study the effect of speaking rate on the coordination of lip, tongue, and jaw movements during the production of the labial consonants, /p, w/ in combination with the vowels, /i, a, u/. The experiment utilized the combined techniques of electromyography, cinefluorography and direct high speed motion picture photography.

Method

Subjects and speech material

Speakers were two adult males, both native speakers of American English. The speech material consisted of the consonants /p, w/ and the vowels /i, a, u/ in a trisyllable nonsense word of the form /kV_1CV_2pə/ where V_1 and V_2 were all possible combinations of /i, a, u/ and C was either /p/ or /w/. An additional set of trisyllables, of the form /kutVpa/ (V = /i, a, u/) were also constructed. These stimuli were incorporated into the EMG part of the run for purposes of obtaining lip rounding data for /u/ in a non-labial consonant environment. The utterance types were random-ordered into a master list. The carrier phrase, "It's a . . .," preceded each utterance. Two speaking rates were studied: slow (normal) and fast. Each speaking rate was based on the subject's own appraisal of comfortable slow and fast rates. A brief practice session preceded each run. The subjects were also instructed to produce the first two syllables of the utterance with equal stress, with the final syllable unstressed. The subjects' performances were monitored continuously throughout the run.

Electromyography

For both subjects, conventional hooked-wire electrodes were inserted into muscles that control movements of the lips, tongue, and jaw. These muscles are listed in Table I.

Table I EMG electrode locations

Subject F.S.C.	Subject T.G.
Orbicularis Oris (OO)	Orbicularis Oris (OO)
Genioglossus (GG)	Genioglossus (GG)†
Internal Pterygoid (IP)*	Superior Longitudinal (SL)†
Anterior Belly Digastric (AD)	Internal Pterygoid (IP)†
	Anterior Belly Digastric (AD)

*Analyzed for combined run only.

†Not usable.

Although all muscle locations showed adequate firing levels at the time of electrode insertion, some locations deteriorated at one time or another during the run. The extent to which this occurred is also indicated in Table I.

The basic procedure was to collect EMG data for a number of tokens of a given utterance and, using a digital computer, average the integrated EMG signals at each electrode position. The EMG data were recorded on a fourteen channel instrumentation tape recorder together with the acoustic signal and a sequence of digital code pulses. These pulses are used to identify each utterance for the computer during processing. A more detailed description of the various aspects of the experimental procedure can be found elsewhere (Hirose, 1971; Kewley-Port, 1971, 1973).

Cinefluorography

Lateral view X-ray films were recorded with a 16 mm cine camera set up to run at 64 ft/s. The X-ray generator delivered 1 ms pulses to a 9 in image intensifier tube. The subject was seated with his head positioned in a standard headholder. A barium sulphate paste was used as a contrast medium on the tongue and tantalum was applied along the midline of the nose, lips, and jaw to outline those structures. The X-ray film records were synchronized with the other records by a pulse train generated by the camera and recorded on the data tape.

High speed motion picture photography

High speed motion pictures of lip movements were recorded with a 16 mm Milliken camera set up to run at 128 ft/s. Because of space constraints the full-face motion pictures of the lips were recorded through a mirror. The motion picture and EMG data were synchronized by an annotation system that displayed the octal code pulses on an LED device placed in the path of the camera. This display was also driven by a signal from the camera to count individual frames between octal codes. Prior to the beginning of the run, white reference dots were painted on the subject's lips, at the midline. A scale was fixed to the mirror for calibration of lip movement measurements. A block diagram of the recording system is shown in Fig. 1.

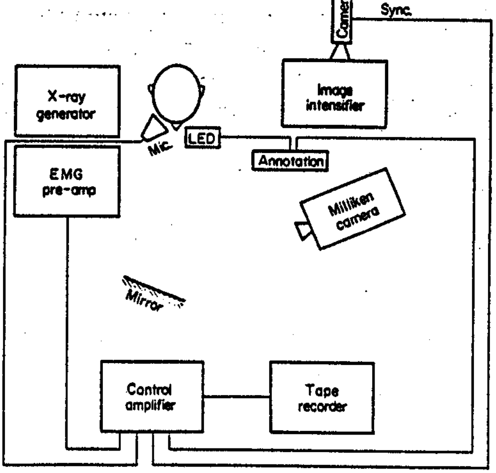

Figure 1 Block diagram of the electromyographic, cinefluorographic and direct, high speed motion picture recording systems.

Data recording and analysis

The combined EMG-cinefluorographic-high speed motion picture data were recorded at the beginning of the run, after which the EMG part of the experiment continued. For this segment of the run (EMG only), the word list was repeated 10 times at each of the two speaking rates. The EMG data from the combined-techniques part of the run were processed separately from the remainder of the run. This allowed comparisons to be made between the individual (combined run) tokens and the averaged data.

The X-ray films were analyzed by frame-by-frame tracings to obtain the outline of the surface of the tongue as well as a direct measurement of jaw displacement (vertical distance between the incisors). The direct view motion picture films were analyzed by frame-by-frame measurements of vertical lip opening at the midline. All film measurements were made on a Perceptoscope film analyzer.

Duration measurements were made from the Visicorder tracings. The mean durations of the utterances (token + carrier) were 980 ms and 650 ms for Subject FSC, and 1030 ms and 670 ms for Subject TG, for the slow and fast speaking rates, respectively.

Results

Lip movement

Results of the electromyographic analyses for both subjects are summarized in Table II. This table shows the peak activity levels of the orbicularis oris muscle for all utterances and both speaking rate conditions. The orbicularis oris muscle is largely responsible for a closing gesture of the lips and its activity, as shown here, is associated with lip closure during the production of the consonant. C_1 in the table represents the first consonant (either /p/ or /w/) and C_2 represents the final /p/ in the utterance.

Generally speaking, the data summarized in this table show that the peak muscle activity levels of the orbicularis oris are greater during the fast speaking rate condition than during the slow speaking rate condition. These differences, with only a few exceptions,[1] hold up for both C_1 and C_2 and for the single tokens (from the combined run) as well as the averaged tokens. The magnitude of these increases, however, varies from nil to over 100%; the only consistent trend is for C_1 differences to be greater than C_2 differences. This, of course, is a probable stress effect. On the whole, these data, like those from a previous study (Gay & Hirose, 1973) demonstrate that the major effect of an increase in speaking rate on labial consonant production is an increase in articulatory effort.

A compatible result was obtained from the frame-by-frame analysis of the direct view, high speed motion pictures. Figure 2 shows the vertical lip opening measurements for the /a/ series of utterances. These data show that the rates of lip closing and opening for the consonant are slightly faster for the faster speaking rate condition. These effects occur for both C_1 and C_2 and all vowels, with the exception of the /upu/ and /uwu/ series for Subject TG. The increase in articulatory speed for the consonant is also carried over to the adjacent vowel as greater lip opening. Lip closure duration is somewhat variable across changes in speaking rate (Table III), although in most cases it decreases with an increase in speaking rate.

The effects that occur for /p/ also occur for /w/, i.e. an increase in speaking rate is accompanied by both an increase in the activity level of the muscle and an increase in the

---

[1] The instances where an increase in muscle activity level was associated with the slower speaking rate condition occurred for 11 utterances out of a total of 144. Reversals of the expected result were generally small and occurred only for the final /p/ in the /w/ utterances.

Table II Averaged and combined run (in parentheses) peak EMG values (μV) for the orbicularis oris muscle. Values for the slow speaking rate are in the left column and values for the fast speaking rate are in the right column of each cell. An asterisk (*) indicates higher values for the slow speaking rate condition

	Subject FSC		Subject TG	
	C_1 S F	C_2 S F	C_1 S F	C_2 S F
ipip	120–185 (110–195)	130–140 (90–110)	205–280 (220–270)	200–210 (180–240)
ipap	150–205 (115–120)	125–165 (80–90)	225–325 (210–285)	240–300 (225–290)
ipup	175–215 (150–175)	135–150 (145–170)	215–345 (190–305)	185–345 (165–315)
apip	140–220 (130–145)	140–145 (80–85)	235–270 (235–265)	215–225 (220–235)
apap	145–225 (120–125)	130–155 (120–190)	220–270 (215–295)	245–260 (240–255)
apup	170–205 (140–150)	145–145 (120–200)	220–270 (205–310)	175–255 (175–250)
upip	130–185 (120–180)	130–145 (100–160)	135–290 (155–240)	210–260 (200–250)
upap	150–240 (95–220)	120–150 (90–95)	155–245 (145–230)	100–210 (105–195)
upup	110–265 (115–205)	140–170 (95–95)	165–245 (145–250)	180–195 (180–195)
iwip	150–230 (115–175)	130–170 (115–95)*	175–325 (180–300)	225–235 (230–230)
iwap	150–240 (150–160)	130–165 (95–95)	195–315 (175–295)	250–215* (240–230)*
iwup	200–275 (130–220)	165–170 (95–110)	160–315 (160–295)	200–290 (190–290)
awip	155–195 (130–225)	125–155 (95–70)*	130–245 (140–240)	220–205* (235–215)*
awap	175–205 (165–175)	125–155 (120–85)*	135–240 (120–235)	230–190* (235–225)*
awup	190–220 (160–200)	165–165 (155–155)	155–275 (165–230)	225–255 (225–240)
uwip	155–190 (120–240)	125–140 (95–105)	105–205 (95–180)	215–205* (205–190)*
uwap	150–245 (150–265)	125–170 (105–90)*	100–235 (115–220)	240–215* (255–220)*
uwup	160–180 (160–220)	140–150 (95–85)*	85–240 (110–205)	190–270 (200–250)

speed of movement of the lips. Further, the target configuration of the lips (minimum lip opening) remains constant across both speaking rates. In other words, the lips do not undershoot the /w/ target during faster speech. An example of the /w/ curves is shown in Fig. 3.

The EMG data also show both contextual and individual differences in labial consonant production. For those utterances where both consonants are /p/, the EMG data for Subject FSC show higher peaks for p_1 than p_2; however, for Subject TG, the opposite is usually the case: p_2 peaks are usually (although not always) higher than p_1 peaks. For both subjects,

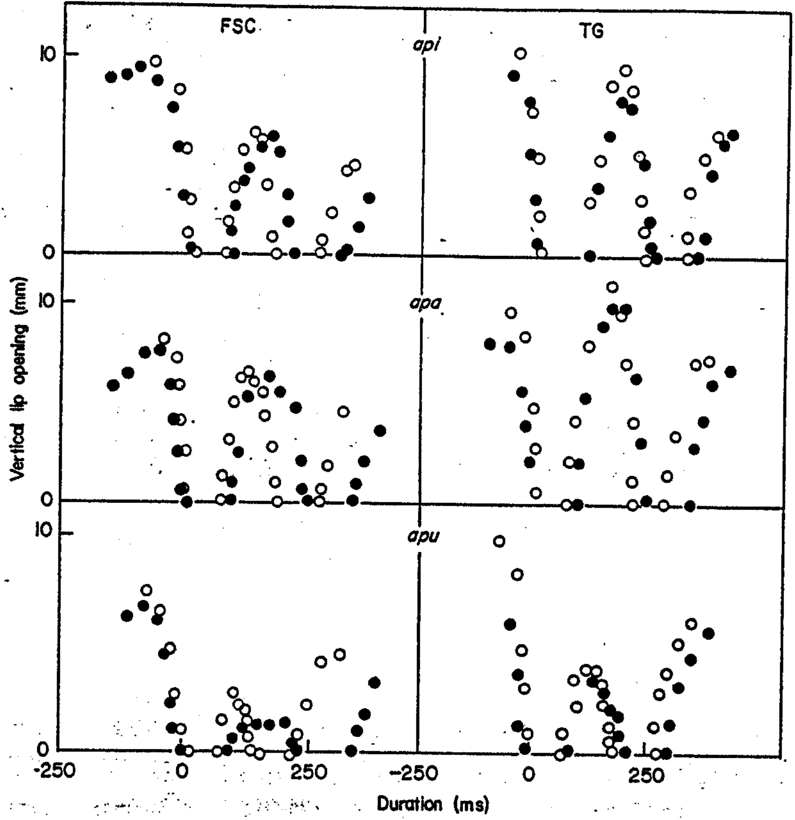

Figure 2 Vertical lip opening measurements. "O" on the abscissa represents the time of lip closure for /p/. The slow speaking rate is represented by filled circles and the fast speaking rate by unfilled circles.

Table III Lip closure durations (ms) for p_1, rounded to the nearest 5 ms

	Subject FSC		Subject TG	
	Slow	Fast	Slow	Fast
ipi	80	60	80	60
ipa	100	100	100	80
ipu	90	60	100	70
api	100	70	60	60
apa	70	70	80	80
apu	110	80	80	70
upi	120	80	110	70
upa	110	90	100	80
upu	80	70	90	60

muscle activity differences are conditioned more by speaking rate than position. In other words, the differences in muscle activity levels as a function of speaking rate are greater than differences in muscle activity levels associated with position (C_1 *vs.* C_2). Further, these data do not show any consistent vowel effects on muscle activity levels for consonant

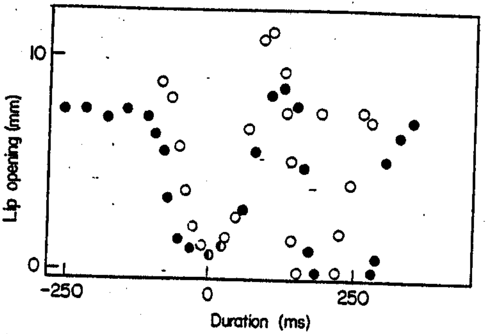

Figure 3 Vertical lip opening measurements for /iwap/ (Subject TG). "O" on the abscissa equals point of minimum lip opening. The slow rate is represented by filled circles and the fast rate by unfilled circles.

closure, although displacement differences were evident in the films (lip opening was greatest for /a/, less for /i/, and least for /u/). This, of course, might be attributed to a trade off between displacement (/a/ and /i/) and degree of rounding (/u/), or to a trade-off between jaw opening *vs.* lip opening.

In summary, both the EMG and motion picture data show that the major effect of an increase in speaking rate on the production of a labial consonant is an increase in articulatory effort and a corresponding increase in the speed of articulatory movement. Both effects imply a reorganization of the commands to the articulators as well as a change in the timing of those commands.

Tongue movement

The EMG data for the genioglossus muscle of Subject FSC, are summarized to Table IV.[2] The genioglossus muscle, which makes up the bulk of the tongue body, is responsible for both protruding and bunching movements of the tongue.[3] The data in this table show that the activity levels of the genioglossus muscle decrease during faster speech. This decrease occurs for all utterances (except for /a/, where the genioglossus muscle shows only resting potentials). The magnitude of these differences is clearly vowel dependent with the greatest differences occurring for /i/, less for /u/, and, of course, none for /a/. Of course, the activity patterns of the muscle at each speaking rate show large and consistent vowel effects in the same directions. This latter finding is a rather common pattern that has been shown before (Smith, 1970; Harris, 1971). Both the vowel and speaking rate effects occur systematically and hold up across /p/ and /w/ as well as both the first and second vowels.

The X-ray films clearly reflect the decrease in muscle activity during faster speech. Figure 4 shows the position of the tongue (Subject FSC) at the V_1 target, time of lip closure for /p/, and V_2 target, for the /a/ series of utterances at both speaking rates. It is evident

---

[2] This electrode was unusable for Subject TG.

[3] The genioglossus muscle is classically divided into two muscle groups. The anterior fibres course fan-like from the mandibular tubercle to subsurface points along the length of the tongue. The posterior fibres run longitudinally from the mandibular tubercle to the hyoid bone. Depending upon the precise location of the recording electrodes, different response patterns can be observed for different gestures. Based on the specific patterns of activity observed for this location, it is assumed that the electrodes were placed in the posterior or lower anterior fibres of the muscle.

Table IV Averaged and combined run in parentheses) peak EMG values (μV) for the genioglossus muscle for subject FSC. Values for the slow speaking rate condition are in the left column and values for the fast speaking rate condition are in the right column of each cell

	V_1	V_2
	S F	S F
ipip	355–205	335–275
	(450–405)	(490–370)
ipap	325–105	—
	(440–195)	—
ipup	360–180	155–125
	(495–285)	(310–200)
apip	—	350–265
	—	(495–480)
apap	—	—
	—	—
apup	—	160–130
	—	(220–210)
upip	230–125	315–300
	(275–180)	(425–310)
upap	185–105	—
	(240–130)	—
	(240–130)	—
upup	195–105	165–165
	(205–150)	(200–195)
iwip	295–180	305–285
	(380–340)	(270–360)
iwap	295–110	—
	(360–210)	—
iwup	305–175	165–145
	(410–390)	(265–230)
awip	—	310–300
	—	(490–350)
awap	—	—
	—	—
awup	—	165–165
	—	(420–400)
uwip	130–90	350–315
	(195–175)	(485–460)
uwap	140–110	—
	(160–150)	—
uwup	125–85	200–150
	(175–145)	(245–190)

from these tracings that the tongue does not extend as far for the vowel during fast speech as it does during slow speech; it clearly undershoots its target. Tongue undershoot occurs consistently for both subjects (although with a greater magnitude for FSC) and for both V_1 and V_2. This shortened course of movement is also reflected in the position of the tongue at the time of /p/ closure. During fast speech the tongue is in a lagging (or more neutral) position with respect to the tongue during slow speech. These positional relationships are most obvious for the /i–a/ contrasts, apparently because of their more directly opposite target positions.

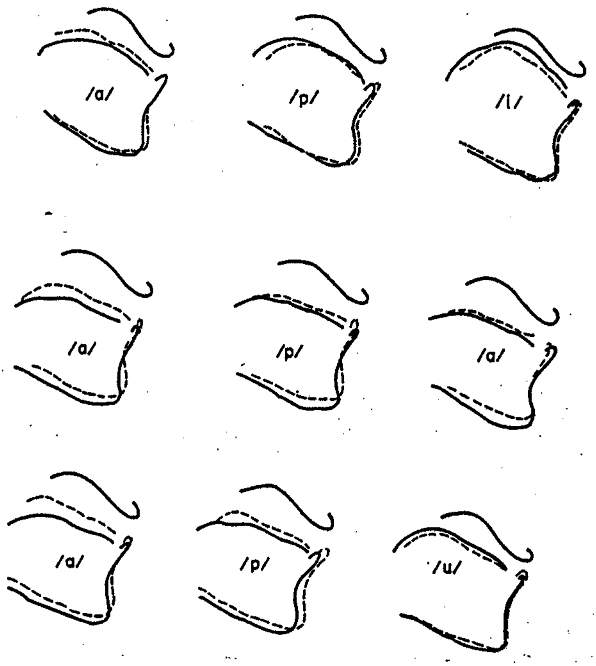

Figure 4 X-ray tracings (Subject FSC) for V_1 target, time of lip closure for /p/, and V_2 target, for slow (solid line) and fast (dashed line) speaking rates.

Although it is apparent that during faster speech, the tongue follows a shorter, more restricted course from vowel to vowel, because of the measurement technique used, the film data could not be used to quantify articulatory rates of movement. However, coupled with the EMG data for Subject FSC, the film data suggest that articulatory rates of movement might very well remain constant across changes in speaking rate. This assumes, of course, that the decrease in EMG activity reflects only a decrease in articulator displacement and not a concurrent decrease in articulator velocity. This suggestion is in agreement with the findings of Lindblom (1964) and Kent (1970), both of whose data show little effect of speaking rate on articulatory velocity. However, such a statement is not necessarily universal. For example, Kuehn (1973) has shown that different speakers seem to use different strategies in the control of speaking rate. His data show that some speakers increase speaking rate by increasing articulatory velocity (with a corresponding decrease in amount of undershoot), and others decrease articulatory displacement (with a corresponding decrease in articulatory velocity).

In this experiment, the data for tongue movement differ markedly from those for lip movement. Whereas the tongue shows a decrease in muscle activity and target undershoot during faster speech, lip movement is characterized by an increase in muscle activity and an increase in articulatory speed. The question then arises as to whether such differences are phoneme related or muscle system related. The EMG data on lip rounding for /u/ bear directly on this question. Table V shows the averaged peak muscle activity levels of the orbicularis oris during lip rounding for /u/. These figures show that, for both subjects and all utterances, the lip rounding gesture for /u/ is characterized by consistently higher peaks of muscle activity during the faster speaking rate condition. Thus, it would appear that for

Table V Averaged peak EMG values (μV) of the orbicularis oris muscle for lip rounding of /u/

	Subject FSC		Subject TG	
	Slow	Fast	Slow	Fast
uti	120	150	80	145
uta	115	165	80	150
utu	145	185	75	140

the phoneme sequences studied here, the observed vowel-consonant differences are, in fact, tongue-lip differences.

The X-ray tracings can also serve to illustrate the precision with which vowel targets are attained at a given speaking rate, irrespective of preceding or following vowels. Figure 5 shows vowel target positions of the tongue for the /a/ series of utterances. Although the routes taken by the tongue toward these targets vary considerably, the final target position is consistently stable. This finding, which is in agreement with those of Houde (1967) and MacNeilage & DeClerk (1969), characterizes the V$_1$ and V$_2$ targets for /p/ as well as for /w/.

Jaw movement

The electromyographic data for the anterior belly of the digastric muscle are summarized in Table VI.[4] The anterior belly the digastric acts to open the jaw; thus, the values shown in this table are associated with jaw opening for V$_1$ and V^2.

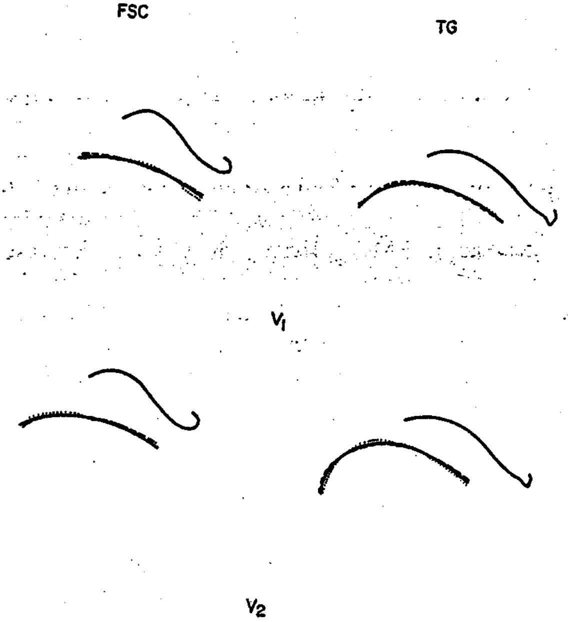

Figure 5 X-ray tracings for the /a/ series of utterances. V$_1$ represents /apV/ and V$_2$ represents /Vpa/, both at slow speaking rates. /i/, ———; /a/ ----, /u/ -----.

---

[4] The internal pterygoid muscle was also studied for both subjects. However, for Subject TG, this muscle did not show any activity during speech (although it was active for clenching), and for Subject FSC, activity was present only during closure for initial /k/.

Table VI Averaged and combined run (in parentheses) peak EMG values (μV) for the anterior belly of the digastric muscle. Values for the slow speaking rate condition are in the left column and values for the fast speaking rate condition are in the right column of each cell. An asterisk (*) indicates higher values for the slow speaking rate condition

	Subject FSC		Subject TG	
	V_1	V_2	V_1	V_2
	S F	S F	S F	S F
ipip	110–135	—	120–280	15–15
	(90–135)	—	(110–130)	(30–40)
ipap	115–120	—	120–240	30–75
	(130–135)	—	(120–260)	(25–30)
ipup	110–110	—	115–265	23–25
	(125–180)	—	(100–265)	(30–35)
apip	105–170	—	165–400	30–45
	(145–220)	—	(170–420)	(25–25)
apap	135–155	—	140–375	25–35
	(120–165)	—	(130–370)	(20–30)
apup	120–130	—	200–385	30–30
	(95–180)	—	(185–375)	(30–35)
upip	115–125	—	130–290	25–95
	(115–125)	—	(125–265)	(30–40)
upap	120–30*	—	120–305	30–85
	(105–25)*	—	(115–275)	(30–75)
upup	115–140	—	125–285	30–50
	(125–140)	—	(125–295)	(25–30)
iwip	125–145	—	135–255	30–60
	(110–140)	—	(110–230)	(30–45)
iwap	120–125	—	115–250	35–70
	(125–200)	—	(140–230)	(30–60)
iwup	110–125	—	125–260	30–35
	(115–160)	—	(145–255)	(45–50)
awip	125–125	—	190–390	30–55
	(135–140)	—	(170–370)	(25–60)
awap	115–140	—	160–385	25–45
	(125–130)	—	(135–350)	(25–40)
awup	115–110*	—	175–390	25–35
	(120–150)	—	(170–410)	(35–40)
uwip	105–130	—	135–375	25–95
	(115–115)	—	(125–410)	(25–90)
uwap	110–115	—	135–300	25–70
	(95–115)	—	(140–300)	(25–90)
uwup	105–135	—	140–330	25–50
	(100–110)	—	(165–335)	(35–40)

The muscle activity levels of the anterior belly of the digastric show an increase during fast speech. This increase is consistent (small in magnitude for Subject FSC, large in magnitude for Subject TG), but occurs only for the first vowel. For both subjects, this muscle does not show much more than resting potentials for V_2. It is possible, of course, that these peaks reflect, at least in part, a stabilizing gesture for /k/. However, the timing of the peaks is compatible with the vowel, and further, the data for TG show a consistent vowel effect, i.e. large peaks and greater speaking rate differences for /a-u-i/, in that order.

Figure 6 illustrates the effect of speaking rate on jaw movement. This figure shows that jaw displacement is usually greater during slow speech than during fast speech. Jaw movement during fast speech mirrors jaw movement during slow speech, but from a more closed position. These data are not consistent with recent findings of Abbs (1973), whose measurements for /pip/ and /pæp/ show little effect of speaking rate on jaw displacement.

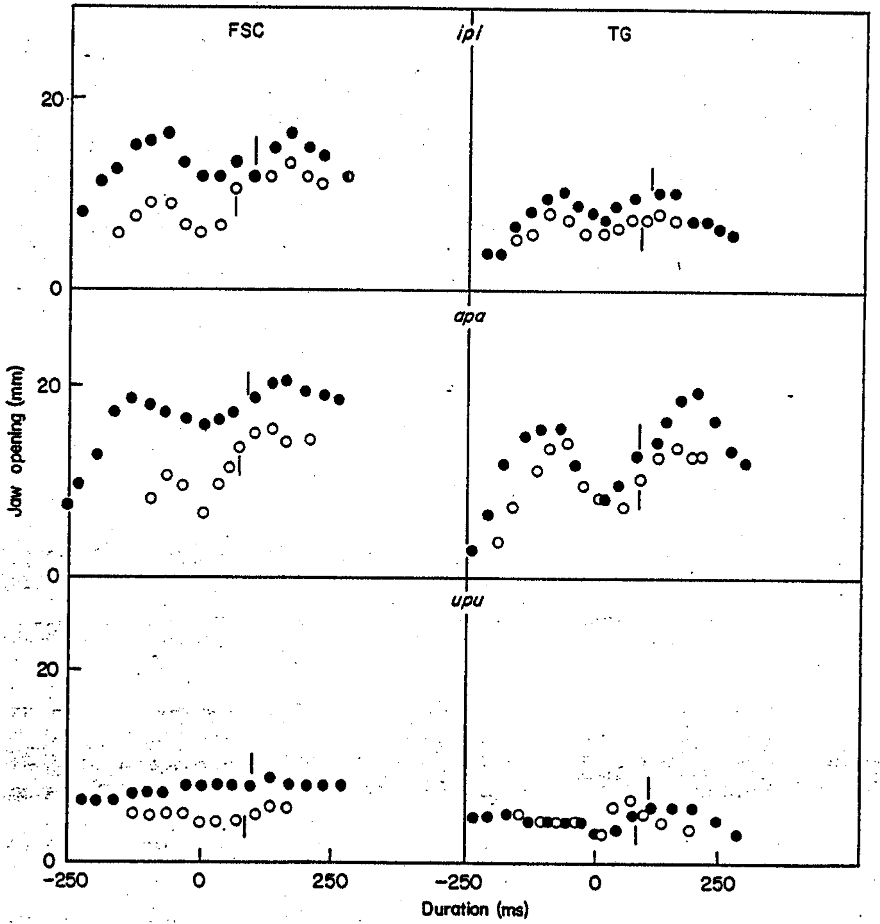

Figure 6 Jaw displacement measurements for slow (filled circles) and fast (unfilled circles) speaking rates. "O" on the abscissa refers to time of lip closure for /p/ and the short vertical lines indicate time of lip opening for /p/.

This figure also shows that for the opening segment of the consonant, jaw opening consistently leads lip opening (lip opening is shown by the vertical lines of each graph). This lead effect, which is evident for both subjects, will be described more fully in the following section. Finally, this figure illustrates that the effect of speaking rate on jaw velocity varies with context. Although a trend seems to exist for rate of movement to increase during faster speech for /a/, and sometimes /i/, no such trend is apparent for /u/. The effects for /a/ occur the most consistently while effects for /i/ are more variable. The absence of any effect for /u/ is obviously due to the fact that jaw movement for /u/ is minimal at both speaking rates.

Jaw movement is also subject to certain anticipatory coarticulation effects. Figure 7 shows the jaw displacement curves for all /p/ utterances at the slow speaking rate. These curves are arranged so that the data for the first vowel in the VCV sequence can be compared with respect to differences in the second vowel. Except for the /u/ series for Subject TG, jaw displacement for V_1 is least when V_2 is /u/. Also, in most cases, the first vowel has relatively little effect on jaw position at the time of /p/ closure for Subject FSC, and almost no effect for Subject TG. Displacement differences for the vowel do not occur consistently between /i/ and /a/, although there is a trend for greater jaw displacement for V_1 when V_2 is /a/ rather than /i/.

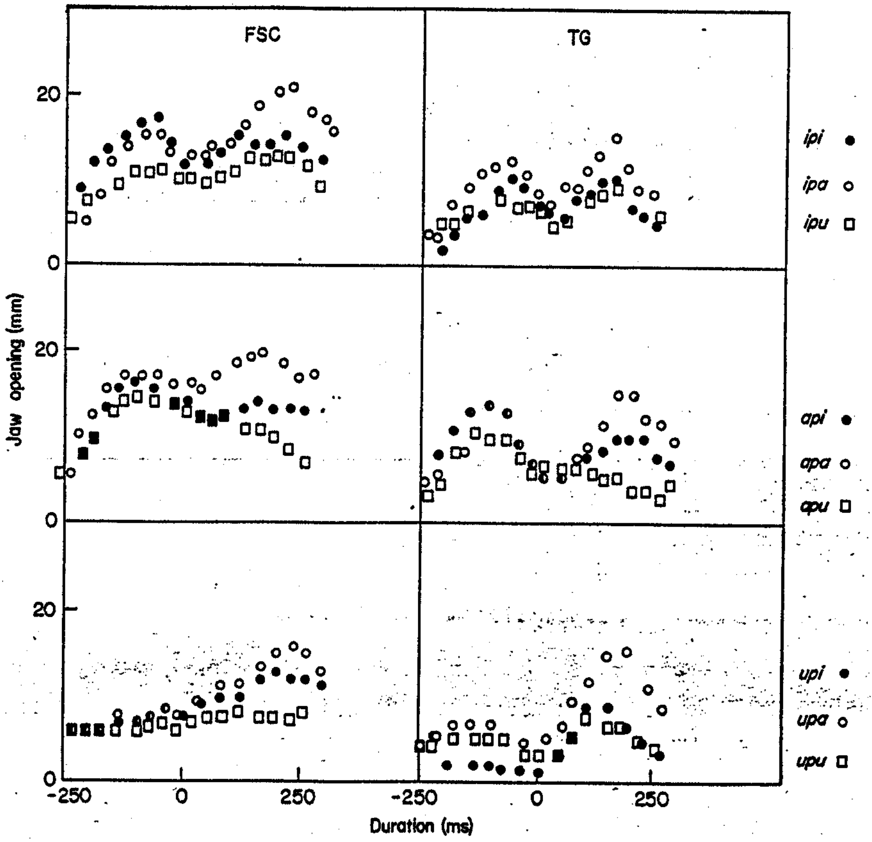

Figure 7 Jaw displacement measurements for all /p/ utterances at the slow speaking rate. "O" on the abscissa indicates time of lip closure for /p/.

Figure 8 shows these data replotted so that second vowel comparisons can be made, i.e. the effect of a different first vowel on the displacement for the same second vowel. For these comparisons, individual differences are evident. For Subject FSC, large /u/ effects on jaw displacement for the second vowel occur consistently, i.e. jaw displacement for the second vowel is least when the first vowel is /u/. These curves also show an effect of the first vowel on jaw closure for /p/. Here, systematic differences in degree of jaw closing for the consonant occur, with greater closing for /u-i-a/, in that order. These effects are not so apparent for the second vowel (the comparisons in Fig. 6). For subject TG, on the other hand, there are essentially no effects of the first vowel on jaw displacement for the second vowel. Also, there do not seem to be any anticipatory effects for the second vowel at the time of /p/ closure. The jaw movement data for /w/ are essentially the same as for /p/, i.e. reduced displacement for the vowel and increased closing for the consonant during faster speech, and decreased displacement for a vowel when either preceded or followed by /u/.

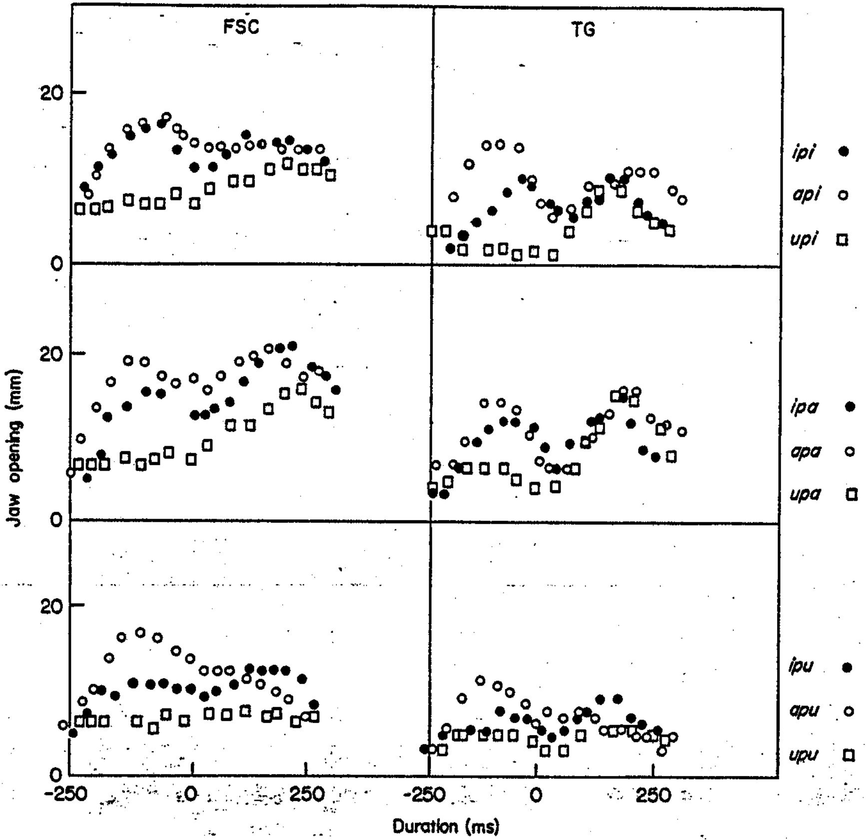

Figure 8 Jaw displacement curves for /p/, replotted for second vowel comparisons.

In summary then, it would seem that jaw movement during fast speech is characterized by a pattern similar to that for slow speech, but with a decrease in overall displacement. Changes in the velocity of jaw movement, when they occurred, did so primarily for /a/. The movement data were only partially supported by the EMG data (for Subject TG, where greater activity levels for the anterior belly of the digastric muscle correlated with the increase in velocity of the jaw for /a/). Of course, too, this experiment sampled only two of the muscles involved in movements of the jaw; this, in itself, plus the absence of any activity for the second vowel in the utterance clearly indicates that a more complete muscle inventory is needed for an adequate description.

Coordination of lip, tongue and jaw movements

Table VII summarizes timing information for movements of the tongue, jaw and lips during the closing, closed, and opening segments of the consonant. The data entered in this table

Table VII Relative onset times (in ms) of tongue, lip and jaw movements (lip closure for /p/ or minimum opening for /w/ = 0) for the closing, closed, and opening segments of the intervocalic consonant. Values for the fast speaking rate appear below those for the slow speaking rate. All values are averaged over the second vowel

	Subject FSC							Subject TG						
	Closing			Closed		Opening		Closing			Closed		Opening	
	Tongue	Lips	Jaw	Lips	Jaw	Lips	Jaw	Tongue	Lips	Jaw	Lips	Jaw	Lips	Jaw
ip	−80	−80	−45	0	15	90	35	−100	−95	−55	0	10	95	20
	−50	−75	−35	0	0	75	20	−60	−70	−45	0	0	70	20
ap	−80	−95	−85	0	20	95	45	−80	−75	−70	0	10	75	20
	−40	−70	−50	0	10	70	25	−55	−65	−50	0	10	70	20
up	−70	−165	−30	0	0	100	20	−95	−175	−25	0	0	100	30
	−50	−135	−30	0	0	80	20	−45	−100	−20	0	0	75	20
iw	−155	−135	−100	0	−10	45	40	−105	−125	−80	0	−20	45	35
	−110	−95	−75	0	−10	30	35	−100	−100	−75	0	−10	20	25
aw	−120	−155	−120	0	0	70	55	−125	−135	−95	0	0	40	45
	−65	−110	−90	0	0	55	45	−75	−90	−80	0	−10	30	30
uw	−110	−210	−95	0	0	120	60	−120	−195	−75	0	0	45	40
	−75	−140	−80	0	0	80	40	−75	−130	−65	0	0	25	30

are the relative onset times (lip closure for /p/ or minimum lip opening for /w/ = O) of tongue movement from the first vowel, and lip and jaw closing and opening for the consonant, at both speaking rates.

For both subjects and both speaking rates, the onset of jaw closing for /p/ lags the onset of both tongue movement and lip closing. The onset time of jaw closing from /a/ is earlier than the onset of either /i/ or /u/. This is most likely because jaw displacement for /a/ is greater than that for the other two vowels. The lips begin to close approximately 75 ms ahead of tongue movement when the vowel preceding the consonant is /u/. This is apparently part of the lip rounding gesture for /u/.

Jaw closing is usually completed at or slightly after lip closure for /p/, while jaw opening for the following vowel precedes lip opening by approximately 50 ms. Except for shorter lead and lag times, similar patterns emerge for fast speech.

Segment durations are somewhat longer for /w/ than /p/ (50 ms for Subject FSC, 25 ms for Subject TG). Also, the onset times of lip movement for /w/ (with two exceptions) are earlier than those for tongue movement. Jaw closing is completed at or slightly ahead of minimum lip opening, while jaw opening begins at about the same time as lip opening. Closure duration is shorter for /w/ than /p/.

The absence of an anticipatory movement of jaw opening during closure for /w/ contradicts recent data of Gay & Hirose (1973), who showed that for /w/, jaw movement was independent of lip movement, anticipating a following vowel by opening for it during lip closing. This effect was not evident, of course, in the present X-ray data, the examination of which easily explains the discrepancy. In their experiment, Gay & Hirose measured jaw displacement by the movement of a marker painted on the chin. The present X-ray films, however, show that flesh points directly over the mandible can move independently from and even in opposite directions to the mandible, itself. Thus, indirect measurements such as the above can produce seemingly accurate, but in fact, erroneous data. It is also possible that similar errors are inherent in strain gauge measurements, especially if the transducer is positioned at the level of the mandibular protuberance.

The major results of this experiment can be summarized as follows. For lip movements associated with either labial consonant production or rounding for the vowel, an increase in speaking rate is accompanied by an increase in the activity level of the muscle and slightly faster rates of movement. For tongue movement during vowel production, an increase in speaking rate has the opposite effect: a decrease in the activity level of the muscle and a decrease in articulatory displacement. For jaw movement, the major effect of an increase in speaking rate is a decrease in the displacement of the jaw throughout the course of the utterance, i.e. for both the vowel and consonant. Jaw movement is also more sensitive to changes in phonetic context, than either lip or tongue movements, and shows a lag effect for consonant closing and a lead effect for vowel opening.

Discussion

The results of this experiment show that the control of speaking rate cannot be accounted for by one simple mechanism. As was shown in a previous experiment (Gay & Hirose, 1973) and confirmed here, the major effects of an increase in speaking rate on the production of a labial consonant is an increase in the activity level of the muscles and an increase in the rate of movement of the lips. Both of these effects are apparent consequences of an increase in articulatory effort. Clearly, a strategy of this type is not necessarily an unexpected one for a consonant gesture that involves an occlusal target. Since the articulators must approximate for a stop consonant,[5] it is reasonable to assume that under the constraints of an increase in speaking rate, they would do so somewhat faster.

The data for the tongue (and to some extent the jaw), however, cannot be explained in the same way. For tongue movement, an increase in speaking rate is accompanied by a decrease in displacement (undershoot) and a decrease in the activity level of the muscle. The decrease in the activity level of the genioglossus muscle shows that undershoot is not, as Lindblom (1963) originally suggested, a consequence of an overlap in the timing of commands to the muscle.[6] The decrease in muscle activity for the tongue during faster speech may reflect only the decrease in overall displacement of the tongue (and not any changes in its speed of movement); nevertheless, the lessened activity is indicative of some reorganization taking place at the level of the muscle commands.

The electromyographic data for the tongue indicate that vowels are characterized by different targets for slow and fast speech. In other words, a vowel target cannot be operationally defined by a set of *invariant* spatial coordinates. Rather, a vowel target should be defined by either a multiple coordinate system or by an articulatory field (with limits). For slow speech, and perhaps for stressed vowels, one given set of coordinates is aimed for, while for fast speech, where articulatory expediency or the constraints of decreased jaw displacement place additional demands on the mechanism, a different set of coordinates is aimed for.

As was mentioned before, Kuehn (1973) has shown that different speakers use different strategies to increase speaking rate, i.e. by trade-offs in displacement *vs.* velocity. These differences might also be explained in terms of a target field or multiple coordinate system. Some individuals might be able to produce a given vowel with a greater degree of freedom than other speakers; that is, the acoustical properties of a given vocal tract might be such that a wider range of formants can produce the same perceptual result. Other tracts might

[5] For /w/, the articulators apparently must reach an invariant target position in order to produce an acoustic steady state.

[6] Lindblom's original model implies that only the timing, not the size, of the EMG signal would change during faster speech.

not have these characteristics and, thus, the articulators must (by increasing velocity) attain a more strictly defined set of target coordinates. However, irrespective of both the strategy employed and the effect observed, the crucial point is that the control mechanism for speaking rate is one that involves changes in both the organization and timing of commands to the muscles.

One final observation is the contrast between the variability of target position across changes in speaking rate *vs.* the precision in attaining that same target in varied phonetic contexts. The larger effects due to speaking rate can be interpreted to mean that the greater challenge to the system lies not in the organization of a target-directed movement but rather in the rapid sequencing of such movements.

In summary, the data of this experiment show that speaking rate is controlled by a mechanism that involves more than a simple change in the timing of motor commands. Reorganization of the gesture for fast speech involves changes in both the duration and size of the muscle contraction. However, an adequate description of the mechanism will require additional information about how the tongue handles vowel sequences which are separated by both lingual and velar consonants, along with a more detailed description of the variability of target position as a function of a greater number of speaking rate variations.

The authors wish to thank Dr J. Daniel Subtelny, Head, Department of Orthodontics, Eastman Dental Center, Rochester, New York, U.S.A. for the use of his cinefluorographic facilities, and Mrs Constance Christian for her help in analyzing the motion picture data. This research was supported, in part, by Grant No. DE01774, National Institute of Dental Research, National Institutes of Health.

References

Abbs, J. (1973). The influence of the gamma motor system on jaw movements during speech: A theoretical framework and some preliminary observations. *Journal of Speech and Hearing Research* 16, 175–200

Gay, T. (1968). Effect of speaking rate on diphthong formant movements. *Journal of the Acoustical Society of America* 44, 1570–1573.

Gay, T. & Hirose, H. (1973). Effect of speaking rate on labial consonant production. *Phonetica* 27, 44–56.

Harris, K. S. (1971). Action of the extrinsic musculature in the control of tongue position: A preliminary report. *Haskins Laboratories Status Report on Speech Research* SR-25/26, 87–96.

Hirose, H. (1971). Electromyography of the articulatory muscles: Current instrumentation and technique. *Haskins Laboratories Status Report on Speech Research* SR-25/26, 73–85.

Houde, R. A. (1967). *A Study of Tongue Body Motion during Selected Speech Sounds*. Ph.D. Thesis, University of Michigan.

Kent, R. D. (1970). *A Cinefluorographic–Spectrographic Investigation of the Component Gestures in Lingual Articulation*. Ph.D. Thesis, University of Iowa.

Kewley-Port, D. (1971). The EMG data system. *Haskins Laboratories Status Report on Speech Research* SR-26, 67–72.

Kewley-Port, D. (1973). Computer processing of EMG signals at Haskins Laboratories. *Haskins Laboratories Status Report on Speech Research* SR 33, 173–184.

Kuehn, D. P. (1973). *A Cinefluorographic Investigation of Articulatory Velocities*. Ph.D. Thesis, University of Iowa .

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 34, 1773–1781.

Lindblom, B. (1964). Articulatory activity in vowels. *Quarterly Progress Status Report* STL/RIT2, 1–5.

MacNeilage, P. F. & DeClerk, J. L. (1969). On the motor control of coarticulation in CVC monosyllables. *Journal of the Acoustical Society of America* 45, 1217–1233.

Smith, T. (1970). *A Phonetic Study of the Function of the Extrinsic Tongue Muscles*. Ph.D. Thesis, U.C.L.A.