Haskins Laboratories Status Report on Speech Research
1992, SR-111/112, 43-62

Acoustic Evidence for Gestural Overlap in Consonant Sequences*

Elizabeth C. Zsiga†

Acoustic evidence for temporal overlap of the two closure gestures in the environment VC#CV was investigated. It was hypothesized that evidence of C2 would be found in the VC formant transitions and would increasingly dominate the transitions as rate (and by hypothesis, overlap) increased. Twenty repetitions (ten at a normal rate and ten at a rapid rate) of word pairs where the first word ended in /d/ and the second began with /p/, /t/, or /k/ were elicited in a sentence context from four subjects. F2 and F3 transitions from the midpoint of V1 to just before closure were then measured. In all environments, C2 had a clear influence on the VC formant transitions. The rate effects were less clear. For the /d#k/ environment, a significant correlation was found between more prominent velar transitions and increasing ratio of vowel duration to consonant closure duration, which may be considered a measure of increasing consonant overlap. The acoustic influence of C2 on V1 suggests considerable temporal overlap of the two closure gestures, and at least for the d#k case, increasing overlap as a function of fluency.

1. INTRODUCTION

The framework of articulatory phonology (first described in Browman & Goldstein, 1986 and elaborated on in their subsequent papers (1988, 1989, 1990a, b)) has drawn attention to the importance of understanding the patterns of gestural overlap in speech. The theory has focused in particular on the fact that there is a significant amount of overlap between sequential consonant gestures. Such overlap, discussed for example by Catford (1977), was first demonstrated instrumentally by Hardcastle and Roach (1977), using electropalatographic data from -VCCV- utterances. X-ray microbeam studies (Browman & Goldstein, 1988, 1990b) confirmed that in utterances like "perfect memory," or "seven plus" the lips begin moving toward the labial closure beginning the second word before the tongue tip closure ending the first word is released. Articulatory phonology proposes that gestural overlap increases in casual, rapid speech, and that this increased overlap may account for the apparent consonant deletions and assimilations characteristic of such speech. Assimilation of final alveolars is particularly common: for example, in the fluent pronunciation of "that boy and that girl" (two of the examples discussed by Gimson, 1962). The hypothesis is that in fast or casual speech gestural overlap may increase to the point where the labial gesture completely masks the preceding alveolar. The alveolar closure may still be made, but its acoustic effects will be largely hidden. As this proposal makes crucial reference to the acoustic consequences of gestural overlap, it is important to investigate those consequences. This experiment investigates whether there is acoustic evidence for the proposed patterns of gestural overlap in the formant transitions of vowels that precede consonant sequences.

A preliminary version of this paper was presented at the 120th meeting of the Acoustical Society of America, November, 1990. Dani Byrd and I collected and began the analysis of the data together, and without her, this experiment would never have been done. While taking responsibility for all the findings and interpretations in this paper, I would like to acknowledge the insightful comments of Michael Studdert-Kennedy, Bruno Repp, and Doug Whalen, and especially the profound influence and considerable assistance, at all stages of this work, of Louis Goldstein. This research was supported by NIH Grant HD-01994 to Haskins Laboratories and an NSF graduate fellowship to the author.

In the act of speaking, articulations overlap. Fowler (1980, p. 114) describes the complete lack of "temporal discreteness" in both articulatory and acoustic records: "The different kinds of gestures go on simultaneously, and thus there are no borders perpendicular to the time axis in an articulatory or acoustic record to separate one segment from another." Instrumental studies demonstrating overlap among speech gestures include Hardcastle (1985), Hardcastle and Roach (1977),and Marchal (1988) Öhman (1966), Perkell (1969),. These studies have shown, as Marchal (p. 287) puts it, that "an inescapable feature of speech production is the well-attested overlapping of speech segments." Articulatory phonology has taken the overlapping nature of speech to be basic to the formulation of phonetic and phonological generalizations. The temporal phasing, and changes in phasing, of articulatory gestures play a crucial role in this framework. The gesture, defined as "an abstract characterization of coordinated task-directed movements of articulators within the vocal tract" (Browman & Goldstein, 1989, p. 206), is the basic constituent of phonetic and phonological description. For example, a wide glottal opening gesture and a tongue tip raising gesture that approximates the alveolar ridge constitute an /s/. Through rules of temporal phasing, gestures become organized into larger contrastive units and into meaningful utterances: "The pattern of organization, or constellation, of gestures corresponding to a given utterance is embodied in a set of phasing principles...that specify the spatiotemporal coordination of the gestures" (1989, p. 211). The movements of gestures are not timed with respect to an external clock but only with respect to the internal stages of some other gesture. The effective temporal phasing between gestures may change, however, especially in fast or casual speech, with important consequences.

Many researchers (e.g., Barry, 1985, Kaisse, 1985) have described the differences between careful "canonical" pronunciation and pronunciation in "connected" fast or casual speech. Catford (1977) discusses the various consequences that arise when two consonants (especially those made at different places of articulation) become adjacent in fluent pronunciation. Word-final alveolars have been a particular focus of discussion as the consonants most likely to undergo deletion or assimilation when followed by another consonant (Avery & Rice, 1989, Byrd, 1991, Gimson, 1962, Guy, 1980, Paradis & Prunet, 1991). Articulatory phonology proposes that many of the processes of deletion, insertion, and assimilation described by these authors are due to changes in the temporal relations between gestures. There will be different consequences (hiding, revealing, or blending of gestures) depending on the gestures involved and the extent of the overlap, yet the theory proposes that "all result from two simple kinds of changes to the gestural score: (1) reduction in the magnitude of individual gestures (in both time and space) and (2) increase in overlap among gestures" (Browman & Goldstein, 1989, p. 214). An example (from Browman & Goldstein, 1990b) of gestural hiding as a result of increased overlap among gestures is the fluent pronunciation of the phrase "perfect memory," in which the final /t/ of the first word is apparently deleted. X-ray microbeam tracings of the utterance show that an alveolar closure *was* produced. It could not be heard, however, because the gesture was completely overlapped by the preceding velar and following labial stops, so that any acoustic effect was hidden by the other closures. Articulatory phonology predicts that other apparent deletions or assimilations of final consonants are the result of increased gestural overlap in fluent speech. That prediction, particularly with respect to the deletion or assimilation of final alveolar stops (as in "that boy") is examined in this paper.

There are two parts to the prediction. The first is that increased gestural overlap in such consonant sequences will result in an acoustic output consistent with the percept that the first consonant has been deleted.[1] Both the duration of closure and the characteristic vowel-to-consonant formant patterns will be affected by increased overlap. Concerning changes in duration, Repp (1978) found that as the intervocalic period of closure is shortened in VCCV sequences, listeners perceive a single consonant rather than two, even though the onset and offset vowel formant patterns indicate two different places of articulation. Typically, it is the second consonant that is reported (see also Abbs, 1971, Ohala, 1990). Concerning the formant transitions from vowel to consonant, however, Repp (1983) concluded that in VCCV utterances there is no perceptible influence of C2 on the first vowel: listeners could not identify C2 on the basis of the transitional vowel formants preceding the consonant closure. Although he found no perceptual evidence of overlap, Repp did find some statistically significant differences in the formant patterns; in particular, for one speaker F2 was higher preceding /bg/ sequences than preceding /bd/. In addition, Repp used carefully articulated speech in this experiment, in which overlap is predicted to be minimal. Byrd (1991) used speech synthesized by the Haskins computational gestural model, which allows precise control of gestural coordination (Browman, Goldstein, Saltzman, & Smith, 1986), to investigate directly the acoustic and perceptual consequences of changes in gestural overlap. She found that as gestural overlap was increased in the sequence /bæd bæn/, vowel formant transitions into the closure gradually became more labial in character and listeners were increasingly likely to report hearing /bæb bæn/. The experiment to be reported here examines the evidence for such gestural overlap in vowel-to-consonant formant transitions in natural speech. In particular, it investigates whether transitions into a word-final alveolar stop differ as a function of a following word-initial /p/, /t/, or /k/.

The formant transitions that are expected before a single labial, alveolar, or velar stop following a low or mid front vowel (the vowel contexts used in this experiment) are known: F2 and F3 falling for a labial, F2 level and F3 level or slightly rising for an alveolar, and for a velar, F2 rising and F3 falling (Delattre, Liberman, & Cooper, 1955; Fant 1970; Klatt, 1980; Stevens & Blumstein, 1978). Before a consonant sequence such as /dp/ or /dk/ there are two possibilities. If movement toward closure for /p/ or /k/ did not even begin until after closure for the /d/ was reached, there could be no influence of the second consonant on the acoustics of the vowel, and transitions into /dp/ or /dk/ would be identical to transitions into /dt/. If, on the other hand, movement toward /p/ or /k/ began before closure for the /d/ was reached, formant transitions into /dp/ and /dk/ would be expected to differ from those into /dt/. Any influence from a following labial stop would be seen as transitions that are more labial in character: both F2 and F3 falling or falling more sharply. Any influence from a following velar stop would be seen as transitions that are more velar in character; again, F3 would be expected to fall, while F2 would be expected to rise. It was hypothesized that these differing patterns would be found in the formant transitions preceding the different consonant sequences, indicating a substantial amount of overlap.

The second part of the prediction of articulatory phonology concerning casual speech processes is that overlap increases in casual, fast speech. While not all fast speech is casual speech (or all casual speech fast), researchers have found a relationship between an increase in speaking rate and changes in gestural organization in the direction of greater overlap. Engstrand (1988) found that an increase in speaking rate resulted in "active motor reorganization" such that "at the faster speaking rate, vowel- and consonant-related gestures were coproduced to a greater extent than at the slower rate" (pp. 1872, 1863). Similar results were obtained by Gay (1978, 1981). Although the relationship between rate and gestural organization is complex (for example Kuehn & Moll, 1976, Ostry & Munhall, 1985, and Fowler, 1980 found that speakers differed in the way an increase in rate affected gestural organization; see the discussion section below), for this experiment it was hypothesized that the manipulation of rate would result in greater temporal overlap between gestures, consistent with the findings of Engstrand and of Gay.

Acoustically, evidence consistent with an increase in overlap at a fast rate of speech would be found in a greater divergence of the formant patterns before the different consonant sequences. Patterns characteristic of the second consonant are predicted to increasingly dominate the transitions as rate increases. If F2 and F3 are expected to be lower for /dp/ than for /dt/, the formants would be expected to fall more sharply at the fast rate than at the slow. Similarly, for /dk/ the convergence of F2 and F3 would be expected to be more pronounced at the fast rate than at the slow. It is not predicted that the formant transitions in the /dt/ sequence would be affected by a greater or lesser amount of overlap between the two consonantal gestures, because /d/ and /t/ involve the same tongue tip gesture at the same place of articulation.

To test these two predictions, this experiment examines tokens of utterances that differ in containing /dt/, /dp/, and /dk/ sequences, produced at fast and slow rates. Formant transitions are analyzed to determine if there are significant differences between transitions into the three sequences, and whether rate has any effect on the amount of divergence.

2. Method

2.1 Materials. Three series of word pairs juxtaposing a final /d/ and an initial /p/, /t/, or /k/ were constructed. The nine word pairs consisted of a single-syllable modifier followed by a single-syllable noun. The three sets differed in phonological context; that is, in the surrounding vowels and consonants, and in the stress pattern. Two of the sets had stress on the second word, the other set had stressx on the first. The nine word pairs were:

A. bad pick	B. bad pen	C. bed pan
bad tick	bad ten	bed tan
bad kick	bad ken	bed can

Groups A, B, and C will be referred to subsequently as the bad*C*ick, bad*C*en, and bed*C*an contexts. Only front vowels were used, in order to avoid any complications due to rounding: the lowering of the formants associated with rounding might be confounded with any effect of the labial consonant. Different combinations of low and high vowels were chosen to allow for possible effects of vowel-to-vowel coarticulation. Additionally, the phonological contexts were chosen so that for each word pair a plausible meaning could be constructed. Each word pair was placed in a sentence, in which the syntactic structure was kept as constant as possible. Each sentence was then placed in a paragraph, with the test utterances in the final clause. The paragraphs were designed to provide a plausible context for the appearance of each word pair. The full paragraphs, with the test utterances shown in italics, are given in Table 1.

2.2 Procedure. The subjects were four undergraduates, two men and two women, all native speakers of American English, who volunteered their time. The subjects were taped in a sound-treated room. In order to familiarize the subjects with the utterances, each subject was first given a set of nine cards, with one paragraph on each card. The tape recorder was turned on and the experimenter asked the subject to read each paragraph aloud at a conversational rate. The order of the cards was randomized, with the condition that the two stress patterns were kept separate: group C was not mixed in with groups A and B. Half the subjects read group C first and half read it last. After reading the paragraphs the subject was given a second set of cards. On each of these cards one of the target sentences was printed ten times. Again, order was randomized, with the separate stress patterns presented in the same order as in the paragraphs. The experimenter instructed the subject to repeat each sentence ten times, reading at a normal, conversational rate. After having read ten repetitions of all the sentences, the subject was given the same set of cards in the same order and was asked to read each sentence (again repeating it ten times) at a very rapid rate. In total, twenty-one utterances of each word pair were obtained from each of the four subjects: one in a paragraph, ten at a normal rate of speech, and ten at a rapid rate of speech. As the paragraphs were used only to familiarize the subjects with the utterances, tokens from the reading of the paragraphs were not used in the analysis.

Table 1. *Paragraphs used in the experiment.*

Molly liked her job as a nurse at the hospital. Her duties were usually interesting, but *she hated when she had to clean a bed pan for a patient.*

Mary and Jim were planning to go to Hawaii for their vacation, but Mary got sick. She went to the doctor, and *he told her to stay home and get a bed tan for a change.*

A man walked into the furniture store. Bob, the salesman, recognized immediately that he was a foreigner because of his clothes. His suspicions were confirmed when the man asked him for a bed can. After several moments, Bob realized that what he wanted was a waste basket. *The man left the store very happy to have found a bed can for his house.*

Susan is an architect who specializes in drawing blueprints. When the apprentice she was training came and asked why his drawings were always smudging, *she explained that he had chosen a bad pen for the job.*

Mary is in a calligraphy class. The teacher asked the class to practice writing the numerals one through ten across a sheet of paper. When she showed the teacher her paper, *he told her there was a bad ten in the set.*

George works in the Barbie doll factory. He works in the inspection department, inspecting Ken dolls. If any of the workers find a defective doll, they leave it for him. On Wednesday after lunch, *he was annoyed to find a bad Ken on his desk.*

The nominating committee met last Wednesday to discuss candidates for the new position. They considered resumes from Smith, Johnson, and Jones. The committee decided that Smith and Johnson were qualified, but *they all agreed that Jones would be a bad pick for the job.*

David hasn't really recovered from his car accident last spring. He was in the hospital for a long time. Even now he has to take a muscle relaxant, because *the doctors say he still has a bad tic on one side.*

Jim is the second-string field goal kicker for his high school football team. Though he comes to all the practices, he hasn't played since the Thanksgiving game. He isn't allowed to play, because *the coach thinks he made such a bad kick in the game.*

2.3 Analysis. For each talker, eight tokens of each word pair at the slow rate and eight at the fast rate[2] were digitized at a 10-kHz sampling rate, and analyzed by LPC analysis in the ILS program. For each analysis frame, a 20 ms Hamming window was used, with twelve filter coefficients for the female talkers and fourteen for the males. Ten milliseconds separated successive analysis frames.

In order to quantify the formant transitions, F2 and F3 were measured at two points: in the middle of the vowel preceding the consonant sequence, and immediately before closure. The point immediately before closure was defined as the last analysis frame in which three distinct formants were visible in the spectrogram, that coincided with a steeply declining amplitude envelope in the vocalic portion of the waveform, and that had a residual energy 10 to 25% of that found in the steady state portion of the vowel but still greater than that found during closure. The residual energy is an approximate estimate of source amplitude, determined by LPC analysis through inverse filtering. In nearly all cases the three selection criteria picked out the same point. In the few cases where they did not, two of the three criteria were considered sufficient to determine the endpoint.[3] The first frame after the release of the initial /b/ in "bad" or "bed" was chosen as the beginning of the vowel, and the frame halfway between release and closure was chosen as the midpoint. (For an even number of frames the frame nearer the end of the vowel was chosen as the midpoint.) The formant transition was taken to be the difference between the final value and the midpoint value.[4] Analysis of variance was carried out on F2 and F3 to determine if the transitions differed significantly according to following consonant sequence.

In order to quantify the rate effects, the durations of the vowel preceding the consonant sequence and of the consonant closure were measured in the digitized waveform of each token. The vowel duration was measured from the beginning of the release burst to the end of the vocalic portion of the waveform, which was defined by a sharp reduction in the amplitude envelope, and the consonant closure was measured from the end of the vowel to the release of the second consonant.

3. Results

3.1 Differences in transitions preceding the consonant sequences. The predicted contrasts were indeed found in the formant transitions before the consonant sequences in these data. The measured values of the change in F2 and F3 for each subject and phonological context are given in Table 2:

Table 2. *Mean measured values for F2 and F3: vowel midpoint, vowel offset, and change from midpoint to offset.*

	F2 midpoint	final	Δ F2	F3 midpoint	final	Δ F3
1 adpi slow	1642	1586	-56	2495	2544	49
fast	1569	1544	-25	2500	2490	-10
adti slow	1649	1623	-26	2504	2628	124
fast	1607	1588	-19	2508	2573	65
adki slow	1675	1801	126	2551	2589	38
fast	1644	1796	152	2476	2470	-6
1 adpe slow	1632	1611	-21	2503	2591	88
fast	1600	1575	-25	2521	2584	63
adte slow	1610	1627	17	2513	2675	162
fast	1595	1555	-40	2532	2583	51
adke slow	1636	1658	22	2486	2561	75
fast	1620	1654	34	2490	2543	53
1 edpa slow	1711	1700	-11	2585	2683	98
fast	1670	1618	-52	2568	2597	29
edta slow	1769	1702	-67	2644	2694	50
fast	1711	1671	-40	2634	2686	52
edka slow	1785	1802	17	2603	2616	13
fast	1734	1738	4	2580	2572	-8

Table 2. *(continued).*

	F2			F3		
	midpoint	final	ΔF2	midpoint	final	Δ F3
2 adpi slow	1915	1773	-142	2828	2878	50
fast	1932	1767	-165	2882	2874	-8
adti slow	1928	1850	-78	2840	2955	115
fast	1902	1792	-110	2890	2939	49
adki slow	2023	2114	91	2836	2977	141
fast	2021	2115	94	2844	2799	-45
2 adpe slow	1924	1792	-132	2868	2923	55
fast	1918	1749	-169	2859	2875	16
adte slow	1933	1855	-78	2910	3046	136
fast	1908	1808	-100	2808	2909	101
adke slow	1981	2051	70	2841	2896	55
fast	1994	2011	17	2920	2905	-15
2 edpa slow	2042	1999	-43	2889	2966	77
fast	2028	1959	-69	2922	2952	30
edta slow	2026	1991	-35	2924	2944	20
fast	2023	1984	-39	2926	2977	51
edka slow	2049	2144	95	2916	2964	48
fast	2059	2144	85	2916	2952	36
3 adpi slow	1730	1760	30	2769	2887	118
fast	1783	1777	-6	2790	2815	25
adti slow	1772	1875	103	2843	2887	44
fast	1794	1852	58	2824	2944	120
adki slow	1821	2241	420	2840	2707	-133
fast	1888	2476	588	2851	2765	-86
3 adpe slow	1764	1790	26	2871	2906	35
fast	1736	1708	-28	2850	2877	27
adte slow	1800	1924	124	2859	2863	4
fast	1772	1886	114	2855	2964	109
adke slow	1776	2128	352	2806	2855	49
fast	1807	2118	311	2854	2744	-110
3 edpa slow	1924	1884	-40	2887	2913	26
fast	1844	1804	-40	2866	2903	37
edta slow	1992	1995	3	2937	2997	60
fast	1944	1939	-5	2911	2958	47
edka slow	1968	2191	223	2914	2873	-41
fast	1965	2260	295	2896	2819	-77
4 adpi slow	1509	1456	-53	2398	2512	114
fast	1552	1468	-84	2406	2426	20
adti slow	1517	1476	-41	2424	2558	134
fast	1544	1486	-58	2409	2486	77
adki slow	1577	1602	25	2402	2321	-81
fast	1698	1865	167	2412	2241	-171
4 adpe slow	1483	1430	-53	2416	2518	102
fast	1528	1454	-74	2411	2432	21
adte slow	1506	1463	-43	2456	2575	119
fast	1565	1485	-80	2424	2464	40
adke slow	1557	1570	13	2442	2327	-115
fast	1618	1716	98	2339	2258	-81
4 edpa slow	1579	1491	-88	2454	2468	14
fast	1598	1529	-69	2379	2441	62
edta slow	1577	1513	-64	2505	2588	83
fast	1616	1555	-61	2421	2501	80
edka slow	1650	1636	-14	2482	2416	-66
fast	1631	1696	65	2410	2287	-123

A negative value indicates a falling transition, a positive value a rising transition. Figure 1 shows the mean formant transitions preceding /dp/, /dt/, and /dk/ for all subjects across both rates. The plot shows the formant frequency value at the vowel midpoint connected by a line to the value at the vowel offset. For the transitions into /dk/ F3 fell and F2 rose sharply. In the transitions into /dp/, while the differences were smaller, both F2 and F3 ended lower than F2 and F3 preceding /dt/.

To test for the significance of these patterns, analyses of variance were carried out on the differences between the final value and midpoint value for each token. An overall ANOVA, including data from the four subjects in three phonological contexts at both rates, reveals a significant main effect of consonant sequence for both F2 and F3 (for F2, $F_{2,6} = 10.48$, $p = .011$; for F3, $F_{2,6} = 6.05$, $p = .036$). However, for both F2 and F3 there was also a highly significant main effect of subject ($F_{3,503} = 251.27$ for F2 and 15.578 for F3, $p < .001$) and a highly significant interaction of subject and consonant ($F_{6,503} = 53.94$ for F2 and 18.402 for F3, $p < .001$), as well as several other significant interactions. These statistics indicate that the formant transitions for each subject differed in some respects, and that the relationships between the transitions into /dt/, /dp/, and /dk/ also differed for each subject. When the subjects are considered separately, these differences become apparent. Figure 2 a — d plots the transitions for each of the four subjects, again including the three phonological contexts and both rates. The subjects differed in the extent to which effects of consonant sequence were seen in both F2 and F3, and in the extent to which /dp/ and /dk/ were distinct from /dt/. For subject 1, the F3 values at the vowel offset are very similar for the three consonants, but when the midpoint values are considered, F3 for /dk/ is seen to fall, distinguishing it from the slightly rising patterns of both /dp/ and /dt/. While subject 2 shows little or no separation by consonant sequence in F3, subjects 3 and 4 show a clear divergence in F3 in the predicted directions. For all four subjects, the endpoints for /dp/ and /dk/ are lower than those for /dt/, as expected. In F2, for all subjects the transition into /dk/ ends higher than the transitions into /dp/ and /dt/, but /dt/ and /dp/ are widely divergent only for subject 3. Overall, subject 3 fits the hypothesis of differing formants almost perfectly, showing a clear separation of both formants in the predicted directions for the three consonant sequences. The plots of the data for the other subjects do not show all of these distinctions as clearly, especially the distinction between /dt/ and /dp/ in F2.

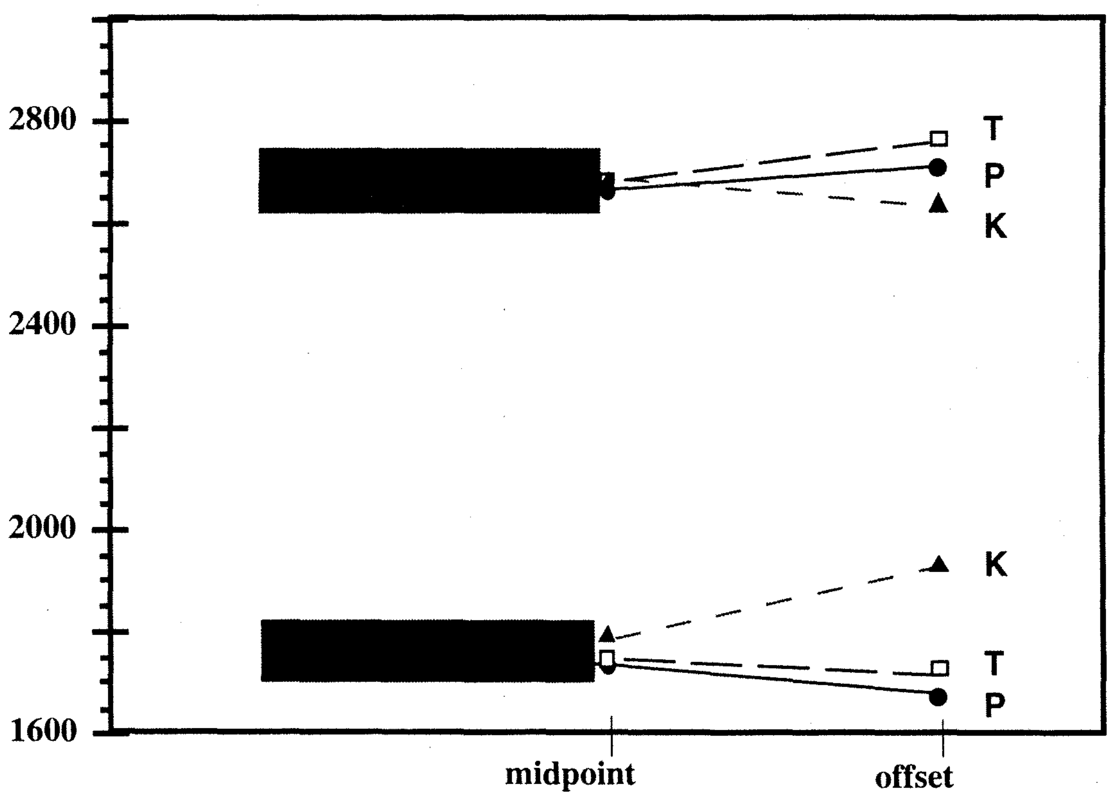


Figure 1. **Formant transition: Mean for all subjects.**

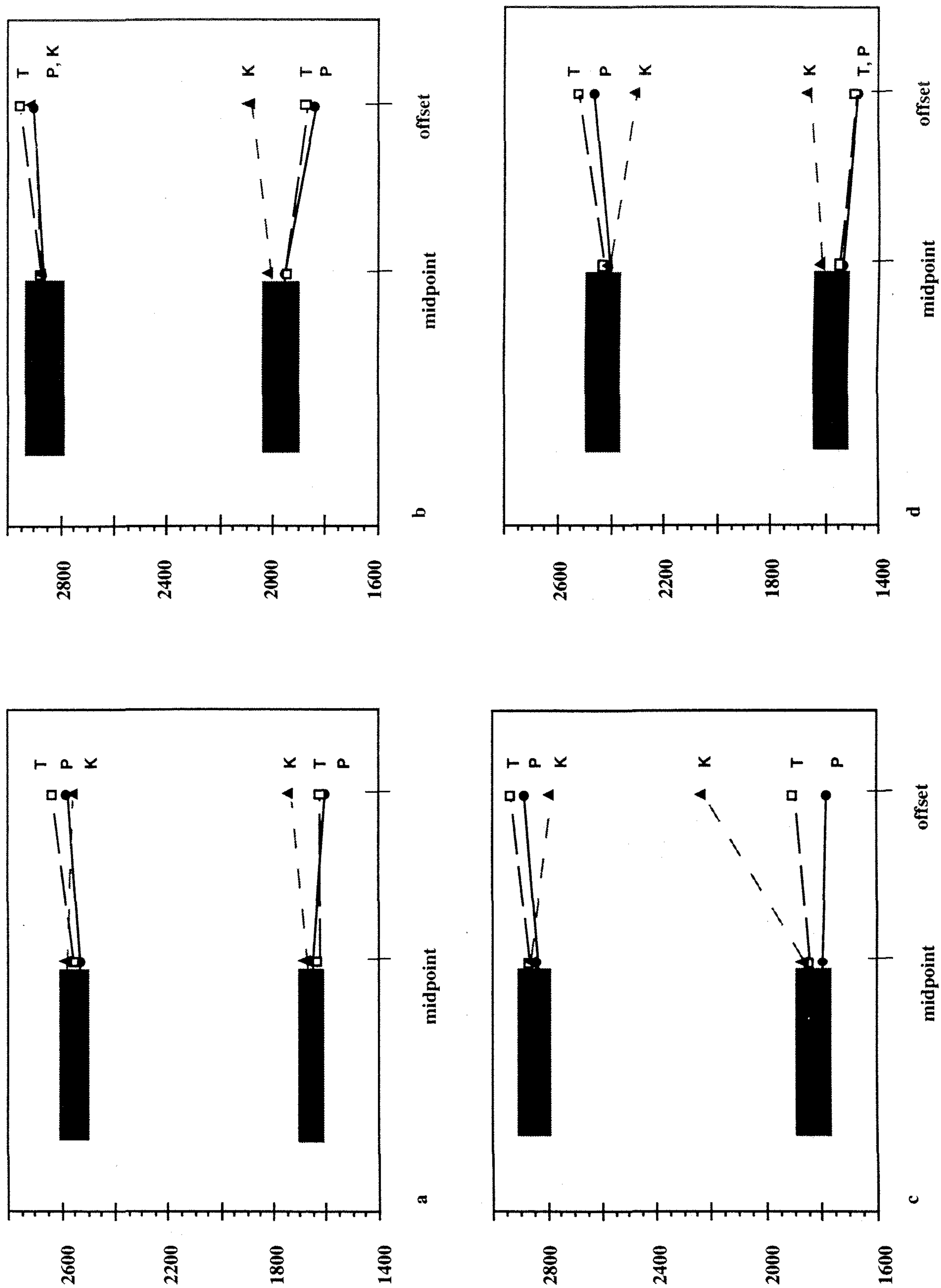


Figure 2. **Mean formant transitions. a. Subject 1. b. Subject 2. c. Subject 3. d. Subject 4.**

To simplify the statistical analysis, the data from each subject were analyzed separately. For each of the four subjects an analysis of variance (Table 3) revealed highly significant effects on both F2 and F3 due to the word-initial consonant. Although for each subject the consonant effect interacted with phonological context, or phonological context and rate, simple main effects revealed that the consonant effect was robust across these other factors.[5]

These effects were further analyzed by post-hoc comparisons, to determine if both initial /p/ and initial /k/ differed significantly from initial /t/ in the expected direction, or whether the significance came solely from the difference between /t/ and /k/. Newman-Keuls analysis was used to compare the means of the formant transitions both between /t/ and /k/ and between /t/ and /p/. In each case, the post-hoc test was done separately in the smallest cell for which analysis of variance showed no interactions. Thus for each subject each phonological context was considered separately by an analysis of variance. When the ANOVA for a given context indicated a significant interaction of consonant sequence and rate, the two rates were also considered separately.

Figure 3 shows the results of these tests. The hypothesis being tested predicts that both between /t/ and /k/ and between /t/ and /p/ there will be significant differences. Both F2 and F3 are predicted to be lower for /p/ than for /t/. For /k/, F3 should be lower than for /t/, F2 higher. In the figure, the diagonally shaded blocks mark a difference in the expected direction; the gray blocks a *significant* difference ($p < .05$) in the expected direction. There were no significant differences in the direction opposite to that predicted. As was seen in the graphs of these data, the distinction between /t/ and /k/ is the more robust. In F2, the formant transitions were significantly higher (more sharply rising) for /k/ than for /t/ in almost all cases. In F3, /k/ was significantly lower than /t/ in a majority of cases. While both F2 and F3 were lower for /p/ than for /t/ in almost every case, the difference was significant at the .05 level in half or fewer cases.

Overall, however, these analyses of variance indicate that all subjects and phonological contexts show a highly significant variation due to following consonant sequence. The formant transitions in a vowel preceding a consonant sequence do differ depending on the second consonant, in the predicted direction. The influence of a following velar consonant is especially clear.

Table 3. *Subject by subject ANOVA on formant transitions, showing only significant effects.*

Source	df	F-ratio	Probability
Subject 1			
F2			
Consonant	2,126	61.08	0.000
Phol. Context	2,126	14.40	0.000
Ctx*Cns	4,126	13.49	0.000
Ctx*Cns*Rate	4,126	2.34	0.059
F3			
Consonant	2,126	16.28	0.000
Phol. Context	2,126	11.46	0.000
Ctx*Cns	4,126	3.63	0.007
Rate	1,126	31.57	0.000
Ctx*Cns*Rate	4,126	2.84	0.027
Subject 2			
F2			
Consonant	2,125	147.27	0
Phol. Context	2,125	14.35	0.000
Ctx*Cns	4,125	4.48	0.002
F3			
Consonant	2,125	5.67	0.004
Ctx*Cns	4,125	4.33	0.003
Rate	1,125	19.50	0.000
Ctx*Rat	2,125	4.45	0.014
Subject 3			
F2			
Consonant	2,126	241.11	0
Phol. Context	2,126	24.08	0.000
Ctx*Cns	4,126	8.16	0.000
Cns*Rat	2,126	4.61	0.017
F3			
Consonant	2,126	29.47	0.000
Cns*Rat	2,126	4.68	0.011
Ctx*Cns*Rate	4,126	4.16	0.003
Subject 4			
F2			
Consonant	2,126	132.93	0
Phol. Context	2,126	6.36	0.002
Ctx*Cns	4,126	2.82	0.028
Rate	1,126	12.05	0.001
Cns*Rat	2,126	29.34	0.000
F3			
Consonant	2,126	106.67	0.000
Rate	1,126	12.50	0.001
Ctx*Rat	2,126	3.39	0.037
Ctx*Cns*Rate	4,126	2.99	0.022

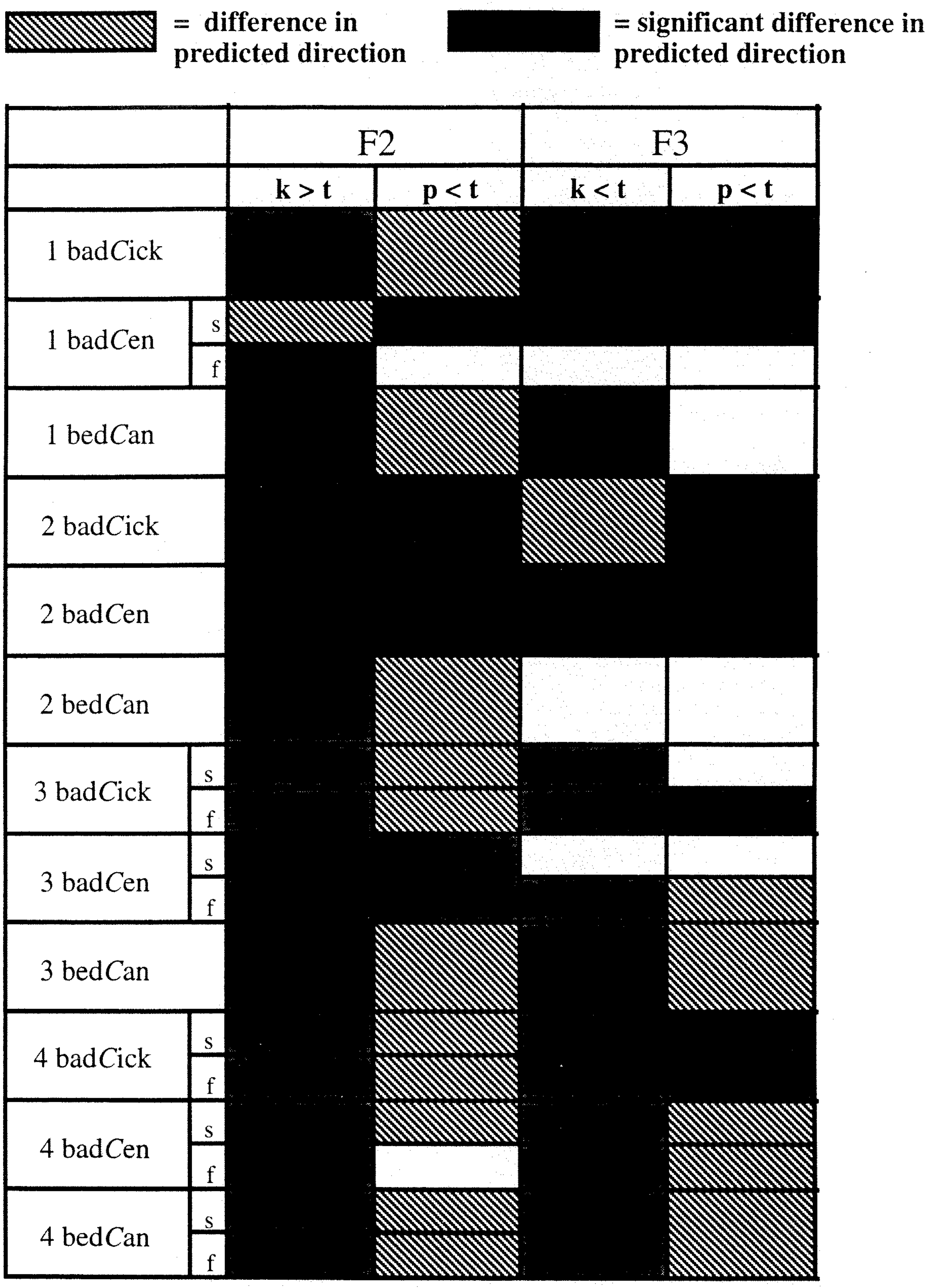


Figure 3. **Results of the post-hoc tests. The distinction between fast and slow is made only when there is a significant interaction of rate and consonant sequence within each phonological context. There were no significant differences in the direction opposite to that predicted.**

3.2 Rate. When asked to speak "at a rapid rate," the subjects did indeed speed up. They differed however, in the extent to which their rate of speech changed. Phonological context also had an effect. Table 4 shows, for each subject and phonological context, the mean duration of the vowel in the first word of the pair (either "bed" or "bad") at both the conversational and rapid rates, the difference in mean vowel duration between the two rates, and the ratio of mean duration at the fast rate to mean duration at the slow rate. Consonant closure duration for the two rates is also reported, as well as the ratio of mean vowel duration to mean consonant closure duration. The actual vowel length is of little interest, except to show that there is no overlap between the rates: for each phonological context, the slowest talker at the fast rate shows a shorter vowel duration than the fastest talker at the slow rate. The concern here is not how fast in absolute terms the talker spoke, but rather how the talker's rate of speech changed between the two conditions. It was hypothesized that a greater change in rate would lead to more overlap and greater acoustic influence of the second consonant in the sequence. The ratio of durations at the fast and slow rates is thus more interesting, as it may be taken as a measure of the change in rate from fast to slow for each subject.[6] An analysis of variance on the mean ratios of fast to slow vowel duration for each subject, phonological context, and consonant revealed that with respect to changes in rate subjects again behaved differently.

Table 4. *Change in rate: Duration measurements.*

		Mean vowel duration slow rate	Mean vowel duration fast rate	Difference in mean vowel duration slow - fast rate	Ratio of mean vowel dur. fast rate / vowel dur. slow rate
1	bed*C*an	111.5	94.9	16.6	.851
	bad*C*en	156.9	106.9	50.0	.681
	bad*C*ick	148.8	108.2	40.6	.727
2	bed*C*an	113.6	87.4	26.2	.769
	bad*C*en	172.7	132.6	40.1	.767
	bad*C*ick	174.7	123.7	51.0	.705
3	bed*C*an	101.2	90.1	11.1	.890
	bad*C*en	170.8	132.0	38.8	.770
	bad*C*ick	154.2	126.0	28.2	.817
4	bed*C*an	111.7	85.7	26.0	.767
	bad*C*en	171.2	98.7	72.5	.577
	bad*C*ick	157.1	93.5	63.6	.595
		Mean C closure dur. slow rate	Mean C closure dur. fast rate	Ratio of mean V dur. / C dur. slow rate	Ratio of mean V dur. / C dur. fast rate
1	bed*C*an	108.3	82.1	1.03	1.16
	bad*C*en	114.7	88.2	1.37	1.21
	bad*C*ick	114.8	84.2	1.30	1.28
2	bed*C*an	134.3	98.0	.846	.891
	bad*C*en	138.2	112.0	1.25	1.18
	bad*C*ick	137.8	99.6	1.27	1.24
3	bed*C*an	87.4	73.3	1.16	1.22
	bad*C*en	84.0	69.3	2.03	1.90
	bad*C*ick	94.8	71.1	1.63	1.77
4	bed*C*an	106.4	66.0	1.05	1.30
	bad*C*en	106.2	64.6	1.61	1.53
	bad*C*ick	98.8	60.4	1.59	1.54

The analysis indicated significant main effects for both subject ($F_{3,12} = 12.5$, $p = .005$) and phonological context ($F_{2,12} = 8.0$, $p < .02$). The subjects speeded up by different amounts: as can be seen from Table 4, subject 4 showed the greatest change from slow to fast, speaking, in one context, almost twice as fast at the rapid rate (indicated by a vowel duration only 58% as long), while subject 3 showed the least change, with a vowel duration at the fast rate in one context nearly 90% as long as the vowel duration at the slow rate. The main effect of phonological context indicates that the different vowels showed different amounts of change. This is presumably due at least in part to stress. Note that for each subject, it is the bed*C*an context, in which the vowel being measured is stressed, that shows the least change from fast to slow. Probably because they are less subject to reduction, the stressed vowels retain a longer duration at the fast rate. The analysis of variance found no significant interactions, and no significant effect of consonant sequence.[7] This indicates that, for each subject, the consonant to be articulated did not affect the way the change in rate was implemented.

3.3 Rate and Formant Transitions. For these data, the hypothesized relationship between rapid speaking rate and increased formant change held for some subjects and contexts, but not for all. Statistically, this effect should be evidenced as an interaction of rate and consonant sequence in the analyses of variance for each subject (Table 3).[8] The interaction of consonant and rate was rarely significant, however: F2 and F3 for subject 3, and F2 for subject 4. For subject 1, however, the three-way interaction of consonant*rate*phonological context is significant for F3 and just misses significance for F2, and for subject 4 this three-way interaction was significant for F3 . Subject 2 shows no interactions between consonant and rate at all. Where the interaction of consonant*rate*context was significant for a given subject, the consonant*rate interaction was tested separately for each phonological context. The results are given in Table 5. A significant interaction of consonant and rate was found in fewer than half of the cases analyzed. For subject 1, the interaction in the bed*C*an context approached significance, but for bad*C*en, the post-hoc analysis showed that the effect of rate was the opposite of that predicted: there was a significantly greater difference between the formant transitions at the slow rate than at the fast (see Figure 3).[9] It is clear that a change in rate produced different results for different subjects and different phonological contexts.

Table 5. *The interaction of rate and consonant sequence compared to the effect of rate on the ratio of vowel duration to consonant duration.*

		Interaction of Consonant and Rate				Main Effect of Rate on V duration / C duration	
		F2		F3			
		F(2,42)	*p*	F(2,42)	*p*	F(1,42)	*p*
Subject 1	bed*C*an	3.08	**.056**	2.51	**.094**	10.3	**.002**
	bad*C*en	4.48	**.017***	5.70	**.001***	29.1	**.001***
	bad*C*ick	.176	.8394	.110	.896	.009	.922
Subject 2	bed*C*an	(1,125)		(1,125)		2.68	.108
	bad*C*en	.411	.663	1.98	.142	.103	.593
	bad*C*ick					.737	.396
Subject 3	bed*C*an	(1,126)		.829	.443	5.63	**.022**
	bad*C*en	4.61	**.016**	5.10	**.010**	1.34	.253+
	bad*C*ick			4.26	**.021**	5.62	**.022**
Subject 4	bed*C*an	(1,126)		1.53	.228	17.2	**.001**
	bad*C*en	29.3	**0.00**	5.72	**.006**	.018	.893
	bad*C*ick			.311	.734	.054	.816

*This indicates a significant change in the direction opposite to that predicted: both the influence of the second consonant and the ratio of vowel duration to consonant duration were larger at the slow rate than at the fast.

+The interaction of consonant and rate was significant for this cell: $F(2,42) = 6.34$, $p < .004$.

A comparison of Tables 4 and 5 shows that whether or not there was a greater effect of consonant sequence at the fast rate is not directly related to how much the talker sped up at the fast rate. A talker could speak very quickly and still not show any interaction of rate and consonant sequence. Subject 2, for example, showed a large difference in rate, but the change in rate did not produce significant changes in the pattern of formant transitions preceding the different consonant sequences.

Another factor, however, *was* correlated with a significant difference in formant transitions between slow and fast rate: the ratio of vowel duration to consonant closure duration. This ratio may be an indication of the style a speaker uses when speaking quickly. Two possibilities are diagrammed in Figure 4. Figure 4a illustrates a hypothetical articulatory relationship between a vowel and two consonants in a cluster at a conversational rate of speech. The shaded blocks indicate the intervals of closure for the two consonants, showing some overlap between the two. (The vowel is represented by a line only. It will overlap considerably with the articulations of the consonants, although the extent of that overlap is not the focus here. The right-hand end of the vowel line in this figure is not assumed to be meaningful.) Below the diagram the acoustic results of this articulatory organization are indicated: the ratio of acoustic vowel duration to acoustic consonant closure duration is 1:1 (within the range of values for these data). Figures 4b and 4c indicate two possible strategies for increasing rate. When speaking at a fast rate, one possibility would be for a speaker to execute each articulation twice as fast, without changing the relative temporal relationship between the articulations. If this were the case, there would be no difference between the ratio of acoustic vowel duration to acoustic consonant closure duration at the slow and fast rates. This strategy is diagrammed in 4b. The time taken for each articulation is half that taken in 4a, but the percentage of overlap remains constant and the vowel to consonant ratio remains 1:1. Using another strategy, a speaker, in addition to speeding up each individual gesture, might change the temporal relationship between them. If this change produced increased overlap between the consonants, as shown in 4c, the actual duration of the consonant closure would be shorter, and the ratio of acoustic vowel duration to acoustic consonant closure duration would be larger.

A relationship was found to hold between a larger ratio of vowel duration to consonant closure duration and a greater formant change due to following consonant sequence. This relationship is suggested by a comparison of the contexts in which formant transitions show a significant interaction of consonant effect and rate with those contexts in which there was a significant effect of rate on the ratio of vowel to consonant duration. (Analysis of variance on the ratio of vowel duration to consonant closure duration for each token shows highly significant main effects of subject ($F_{3,499} = 174.9$, $p < .001$) and phonological context ($F_{2,499} = 14.7$, $p < .005$). Each subject and context was thus analyzed separately.) A significant effect of rate indicates that the ratio of vowel duration to consonant closure duration was different at the two rates. The absence of a significant effect indicates that the ratio remained the same. Table 5 compares those contexts in which the interaction of rate and consonant sequence is significant with those in which the interaction of rate significantly affects the ratio of vowel to consonant duration. In many cases, those subjects who showed an interaction of consonant effect and rate in some contexts also showed an effect of rate on the ratio of vowel to consonant duration in those contexts. Where there was no interaction of consonant and rate, rate also had no effect on the vowel to consonant ratio. In the one condition where an increase in rate led to a significant *reduction* in the effect of following consonant sequence, the ratio of vowel duration to consonant closure duration was found to be significantly smaller at the fast rate. This suggests a relationship between a change in articulatory organization (evidenced by a change in vowel-to-consonant ratio) and a change in the effect a following consonant sequence has on formant transitions.

The relationship between the hypothesized spectral and temporal indices of overlap was tested directly by correlating the mean transition in F2 and F3 for each subject, phonological context, and rate with the mean ratio of vowel to consonant duration in each of these environments. The scatterplots in Figure 5 plot the relationship between formant transition and the ratio of vowel to consonant duration for /dk/. The data points plot the means for a given rate*subject*consonant condition. A significant correlation was found between a larger vowel to consonant ratio and a steeper rise in F2 and a steeper fall in F3.

That is, over all subjects and vowels, F2 rose more and F3 fell more when the duration of the /dk/ sequence was shorter with respect to the vowel duration. As the regression lines show, the change in F2 becomes more and more positive, and the change in F3 more and more negative, as vowel to consonant ratio increases.

For /dp/, the correlation between formant change and the ratio of vowel to consonant duration was not significant, as shown in Figure 6. The flat regression line in 9a shows that there was no relationship between a larger vowel to consonant ratio and the amount of change in F3. Although the regression line for F2 in 6b is slightly rising, this trend is not significant.

These analyses show that a simple increase in speaking rate is not correlated with a greater acoustic influence of the second consonant in a sequence. Speakers differ in the effect that change in rate has on formant transitions. The strategy a speaker uses in speeding up, as evidenced in the ratio of vowel duration to consonant duration, seems to determine whether or not the formant transitions into a consonant sequence differ more at a rapid speaking rate than at a slow, at least for a following /k/.

One further test of the influence of a following consonant sequence was conducted. In a small perceptual experiment, the word "bad" was excised from the badCick utterances of subject 4. No final release burst was included in the excised syllables. The syllables were randomized and played back to four phoneticians, who were asked to transcribe the words they heard. The results of this informal test are given in Table 6. The subjects overwhelmingly reported the final consonant to be alveolar. The results of this perceptual experiment must be reconciled with the measured differences found in the formant transitions. Although acoustic measurements showed changes toward formant transitions that had labial or velar characteristics in /dp/ and /dk/ sequences, listeners still perceived alveolar stops.

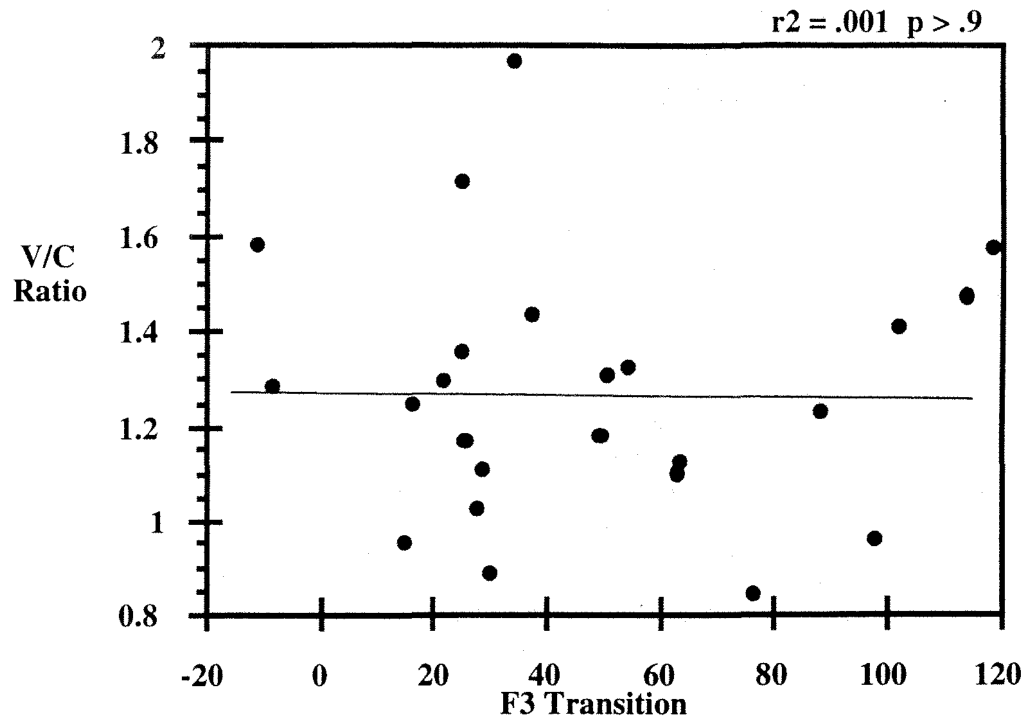
4. Discussion

As was stated earlier, the prediction of articulatory phonology that apparent consonant deletions in fluent speech may be the result of increased gestural overlap has two parts: first, that there is a substantial degree of overlap between two sequential consonant gestures, and that this overlap will be evidenced in the acoustic output as influence of the second consonant on the vowel-to-consonant formant transitions; second, that overlap will increase in more fluent speech, with a concomitant increase in the acoustic influence of the second consonant. The acoustic evidence examined in this experiment is consistent with the hypothesis of substantial gestural overlap, the first part of the prediction. The formant transitions into a word-final alveolar stop were found to differ significantly as a function of a following word-initial /p/, /t/, or /k/. These differences are consistent with the hypothesis that movement of the lips toward closure for the /p/, and movement of the tongue body toward closure for the /k/, begin, in /dp/ and /dk/ sequences, before closure for the /d/ is achieved. Comparison of consonant effect in slow and fast tokens of the same utterances, however, revealed that there is no evidence of a direct relationship between increased speaking rate and increased gestural overlap. For some subjects in some contexts there was a significant interaction of rate and consonant effect; for other subjects and contexts there was no significant difference in formant transitions at the slow and fast rates; for a few contexts, the relationship was the opposite of that predicted, showing a greater effect of consonant sequence at the slow rate.

These results add to the evidence that the relationship between fast speech and casual speech is not straightforward. One can speak very quickly and yet very precisely. The formal, experimental setting in which this speech was recorded may very well have influenced at least some of the talkers to speak carefully. Researchers have found that asking talkers in an experimental situation to speak rapidly will not necessarily result in fluent speech. Rather, speakers will use different strategies in speeding up articulation (Engstrand, 1988; Gay, 1978, 1981; Kuehn & Moll, 1976; Ostry & Munhall, 1985). For example, Kuehn and Moll (1976, p. 320) report that

> The subjects were found to use different physiological methods of changing speaking rate. With an increase in speaking rate each subject reduced transition time by the same amount but the velocity and displacement variables were changed in different proportions to each other depending on the individual speaker.

In this experiment it was found that an increase in the ratio of vowel duration to closure duration is a better indicator of an increase in gestural overlap than is a simple increase in rate. As was shown in Figure 4, as overlap between consonant closure gestures increases, the duration of the consonant closure will be shorter, and the ratio of vowel duration to closure duration will increase.



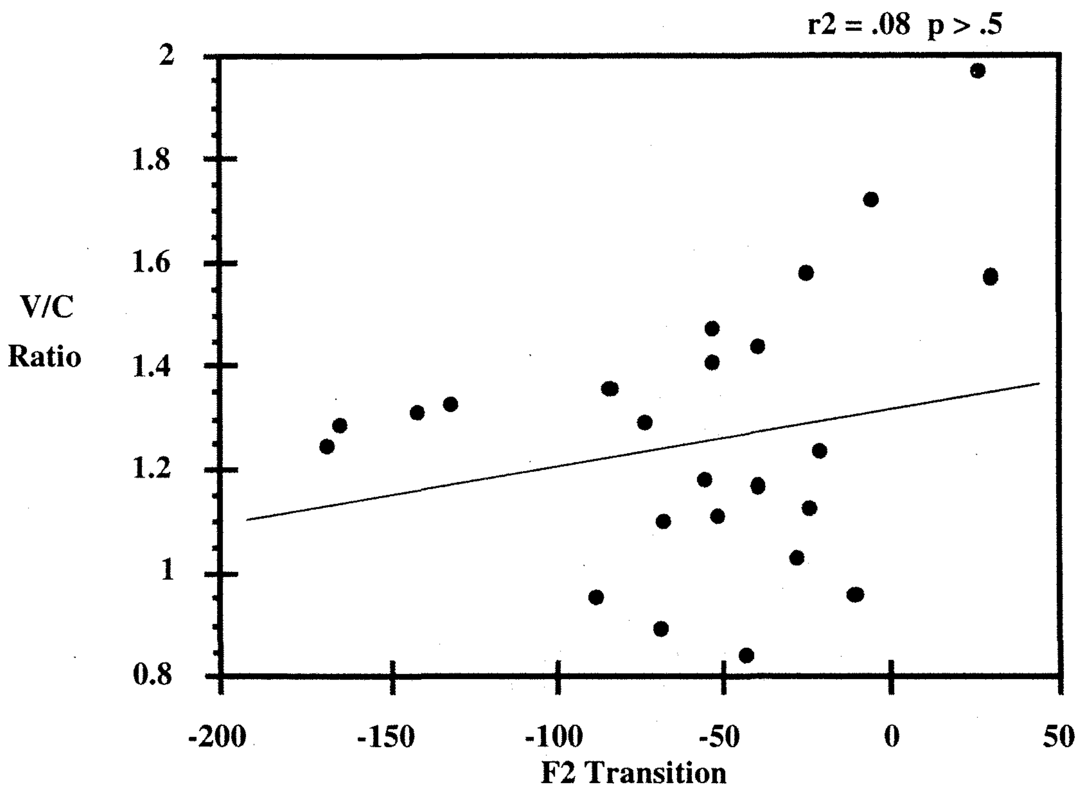


Figure 6. Correlation for /dp/ between increased formant change and ratio of vowel duration to consonant closure duration.

Table 6. *Results of the perceptual experiment.*

	Consonant Reported (out of 48)		
Subject	Labial	Alveolar	Velar
AF	2	42	4
CS	2	44	2
DW		48	
JK	2	44	2

In this experiment it was found, at least for the /dk/ sequences, that an increase in vowel to consonant ratio did correlate significantly with an increased influence of the second consonant on the vocalic formants. If a larger ratio of vowel duration to consonant duration is taken as a measure of increased overlap in fluent speech, as indicated by Figure 4, this correlation provides evidence for a relationship between increasing fluency and formant transitions increasingly characteristic of the overlapping consonant.

For the /dp/ sequences, however, no significant correlation was found. In the analysis of variance on the effects of consonant sequence as well, there were fewer instances of significant differences between /dp/ and /dt/ sequences than between /dt/ and /dk/ sequences. There are two possible explanations for the difference in results between /p/ and /k/. The first is that there is less overlap in alveolar-labial sequences than in alveolar-velar sequences. This would mean that, in the temporal coordination of consonant gestures, movement of the tongue toward closure for a /k/ begins sooner with respect to a preceding consonant than does movement of the lips toward closure for a /p/. While the claim that there is less overlap in /dp/ than in /dk/ sequences accounts for why the difference in formant transitions before /dp/ and /dt/ sequences was less often significant than the difference in transitions before /dk/ and /dt/ sequences (Figure 3), it does not account for the lack of correlation between temporal and spectral measures of overlap in the /dp/ sequences (Figure 6). As Figure 6 shows, the ratio of vowel duration to consonant closure duration did change; no direct relationship could be found, however, between this change and the measured values of the formant transitions.

A second (and more likely) possibility is that the temporal coordination of labial and velar articulations is not different, but that the effects of increased overlap of a labial gesture are less evident in these contexts than the effects of increased overlap of a tongue body gesture. Lack of evidence in the acoustic record of increased labial/alveolar overlap would help account for the lack of a significant correlation between the temporal and spectral measures for /dp/. Possibly, movement of the lips has less of an effect on F2 and F3 in these environments than does tongue body movement; the acoustic effects of two simultaneous closures in the vocal tract has not been extensively investigated. (While Byrd, 1991 measured the effects of varying degrees of overlap between labial and alveolar closures in synthetic speech, she did not compare the timing or magnitude of these effects to the influence of a velar closure). It is also possible that, because the tongue dorsum and tongue tip are connected, movement of the dorsum has a greater articulatory influence on the tongue tip closure gesture than does movement of the lips, which are relatively independent.

In addition to the effects of the following consonants on transitional vowel formants, vowel-to-vowel coarticulation may have had an influence. In two of the three phonological environments examined (bad*C*ick and bad*C*en), the measured vowel was followed by a vowel higher and further forward. Overlap between the vowel articulations in these environments would result in a higher F2 in the first vowel (Choi and Keating 1991, Manuel and Krakow 1984). Such an influence might counteract any lowering of F2 induced by labial closure, but would enhance raising of F2 caused by movement toward velar closure. There is some statistical support in these data for the influence of a following vowel: in F2 all four subjects show a significant effect of phonological context and of the interaction of context and consonant sequence (Table 3). Vowel-to-vowel coarticulation cannot account for all of the discrepancy in results between /dp/ and /dk/, however. As Figure 3 shows, the results for /dp/ were no better in the bed*C*an context, where no vowel-induced raising of F2 is expected, than they were in the bad*C*ick context. A further complication is introduced by the observation that for subjects two and four the lack of a significant difference in the formant transitions preceding /dt/ and /dp/ sequences does not come from the fact that F2 fails to fall in the alveolar-labial sequences, but from the fact that F2 *does* fall in the alveolar-alveolar sequences (see Figure 2). This result was unexpected, and cannot be attributed to the effects of overlap of either the consonants or of the vowels. Given the several factors that might be involved—differences in acoustic influence, differences in articulatory influence, or differences in the influence of the vowel context—an explanation of the divergent results obtained for /dp/ and /dk/ sequences must await further study.

A final problem remains to be addressed. There is an alternative hypothesis that could account for the differences found between transitions into /dp/, /dt/, and /dk/. If the alveolar consonant had been completely deleted in some cases, leaving only the word-initial consonant, transitions into the closure would certainly be those characteristic of that consonant. The overall means would then combine cases where deletion occurred with those where it did not, resulting in intermediate formant transitions. There is evidence, however, that the alveolar consonant was not deleted. First, the /d/ is present perceptually, as was seen in the transcription experiment. An account that posits deletion of the alveolar consonant must explain why subjects still heard a /d/. Second, the influence of the word-initial consonant varies along a continuum, at least for the /dk/ case. The correlation plotted in Figure 5 shows that formant change increases gradually as the ratio of vowel to consonant duration changes.[10] Had the alveolar consonant been deleted in some cases, the points would cluster in two groups: one with a low ratio of vowel duration to consonant duration and little change in F2 or F3, representing cases where the /d/ was present, and the second with a much higher ratio of vowel duration to consonant duration and considerable change in F2 or F3, representing cases where only a /k/ was present. Rather than showing evidence of a sudden change, indicative of complete consonant deletion, the data are more consistent with the hypothesis of overlap. Overlap can vary in its extent, with the second consonant gradually showing more acoustic influence as overlap increases. Byrd (1991) found that, for synthesized speech, as overlap increases the formants gradually become more like those characteristic of the second consonant. Although changes in the formants began as soon as any movement of the articulator for the second consonant preceded the complete closure for the first, listeners continued to perceive the original final consonant until overlap was well advanced.

While the data are not consistent with abrupt deletion of the word-final consonant, reduction in its gestural magnitude may well be involved. The results described here are consistent with an alveolar closing gesture that is shorter in duration or perhaps incomplete before a competing velar closure. X-ray microbeam data for these utterances, which will allow direct measurement of the temporal relations between gestures, as well as of their relative magnitudes, have been collected. Further research, involving both acoustic and physiological measurements, is planned.

5. CONCLUSION

This experiment has shown that there is acoustic evidence for gestural overlap in consonant sequences. There are significant differences in the formant transitions into a word-final /d/ before initial /p/, /t/ or /k/. These differences are consistent with the hypothesis that the gestures for the second consonant begin before closure for the first consonant is reached. Further, comparisons of formant changes across different rates showed that while increased rate does not necessarily result in increased gestural overlap, overlap may increase with rate. A more pronounced gestural overlap, as evidenced, at least for velar consonants, in a larger vowel to consonant duration ratio, does result in more pronounced differences in the vowel to consonant transitions.

REFERENCES

Abbs, M. H. (1971). A *study of cues for the identification of voiced stop consonants in intervocalic contexts*. Doctoral dissertation, University of Wisconsin.

Avery, P., & Rice, K. (1989). Segment structure and coronal underspecification. *Phonology, 6(2)*, 179-200.

Barry, M. (1985). A palatographic study of connected speech process. *Cambridge Papers in Phonetics and Experimental Linguistics, 4*, 1-16.

Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3*, 219-252.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. In O. Fujimara (Ed.), *Articulatory organization—Phonology tospeech signals*. Basel: S. Karger.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6(2)*, 201-51.

Browman, C. P., & Goldstein, L. (1990a) Gestural structures and phonological patterns. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception*. Hillsdale, NJ: Lawrence Erlbaum.

Browman, C. P., & Goldstein, L. (1990b). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech*. Cambridge: Cambridge University Press.

Browman, C. P., Goldstein, L., Saltzman, E., & Smith, C. (1986). GEST: A computational model for speech production using dynamically defined articulatory gestures. *Journal of the Acoustical Society of America, 80*, S97.

Byrd, D. (1991.) Perception of assimilation in consonant clusters: a gestural model. *UCLA Working Papers in Phonetics, 78*, 97-126.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Bloomington: Indiana University Press.

Choi, J., & Keating, P. (1991). Vowel-to-vowel coarticulation in Slavic languages. *UCLA Working Papers in Phonetics, 78*, 78-86.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America, 27*, 769-73.

Engstrand, O. (1988). Articulatory correlates of stress and speaking rate in Swedish VCV utterances. *Journal of the Acoustical Society of America, 83,* 1863-75.

Fant, C. G. M. (1970). Analysis and synthesis of speech processes. In B. Malmberg (Ed.), *Manual of phonetics.* Amsterdam: North Holland.

Fowler, C. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics, 8,* 113-133.

Gay, T. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America, 63,* 223-30.

Gay, T. (1981). Mechanisms in the control of speech rate. *Phonetica, 38,* 148-58.

Gimson, A. C. (1962). *An introduction to the pronunciation of English.* London: Edward Arnold.

Guy, G. R. (1980). Variation in the group and in the individual: The case of final stop deletion. In W. Labov (Ed.), *Locating language in time and space.* New York: Academic Press.

Hardcastle, W. J. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication, 4,* 247-63.

Hardcastle, W. J., & Roach, P. J. (1977). An instrumental investigation of coarticulation in stop consonant seqences. *University of Reading Working Papers.*

Kaisse, E. (1985). *Connected speech: The interaction of syntax and phonology.* NY: Academic Press.

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America, 67,* 971-95.

Kuehn , D. P., & Moll, K. (1976). A cineflourographic investiagion of CV and VC articulatory velocities. *Journal of Phonetics, 3,* 303-20.

Manuel, S. Y., & Krakow, R. A. (1984). Universal and language-particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Report on Speech Research , SR77/78,* 69-78.

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

Ohala, J. J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech.* Cambridge: Cambridge University Press.

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

Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *Journal of the Acoustical Society of America, 77,* 640-48.

Paradis, C., & Prunet, J.-F. (Eds.). (1991). *The special status of coronals.* New York: Academic Press.

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

Repp, B. (1978). Perceptual integration and differentiation of spectral cues for intervocalic stop consonants. *Perception and Psychophysics, 24,* 471-85.

Repp, B. (1983). Bidirectional contrast effects in the perception of VC-CV sequences. *Perception and Psychophysics, 33,* 147-55.

Stevens, K. N., & Blumstein, S. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America, 64,* 1358-68.

FOOTNOTES

*To appear in *Journal of Phonetics.*

†Department of Linguistics, Yale University.

[1] That is, no alveolar stop is heard. In this context, the process might be described as either deletion or assimilation, and no distinction between the two will be made here.

[2] For subject 2, only 7 tokens of "bad kick" at the fast rate were clear enough to be used in the analysis.

[3] For example, formants of a very weak amplitude might be visible some frames into what the waveform indicated was closure. These frames were not counted as vocalic. In the course of the analysis, each token was independently measured twice. There was exact agreement on the frame chosen 90% of the time, and in no case did the points chosen differ by more than one frame.

[4] While defining the formant transitions in this way allowed for consistency across tokens, the shape of the transition could not be taken into account; that is, a smooth decline from midpoint to endpoint could not be distinguished from a steady state followed by a sharp fall. In the judgement of the investigator, however, the transitions were smooth and consistent.

[5] In F2, the consonant effect was significant for all cells except subject 1 adCe slow, where p =.06. In F3, the consonant effect was significant for all cells except subject 1 adCe fast, subject 2 edCa, and subject 3 adCe slow.

[6] The absolute difference between the two rates (column 3) is of course another measure of change in rate. For comparison between the talkers, however, the ratio is a more accurate measure, as it abstracts away from differences between subjects in overall rate of speech. The same absolute difference in msec means a greater relative change for a talker who speaks quickly than for one who speaks slowly. For example, a change of 40 msec translates into a larger percent change for the faster speaking subject 2 than for subject 1. For these data, however, the distinction appears to be small. In fact, statistics done on the difference and statistics done on the ratio show very similar results, with no disparity in the variables found to be significant.

[7] The effect of consonant approached significance ($F_{2,12} = 3.6$, $p = .0944$). This small effect may be due to the relative strangeness ("bed tan") or familiarity ("bad pick") of the phrases.

[8] In the overall ANOVA, the interaction of consonant sequence and rate was only marginally significant for F2 ($F_{2,6} = 5.31$, $p = .047$) and was not significant for F3 ($F_{2,6} = 1.06$, $p = .402$), although the triple interaction of subject*consonant*rate was significant for both F2 and F3 (for F2, $F_{6,503} = 2.88$ $p = .009$, for F3, $F_{6,503} = 2.47$, $p = .023$).

[9] It can be also be seen from Figure 3 (where each phonological context was examined individually for an interaction with rate, regardless of whether the rate*context interaction was significant overall for that subject) that for subject 4 badCen as well, the difference between /dt/ and /dp/ in F2 was greater at the slow rate than at the fast, although the difference was not significant at either rate.

[10] While each point on this graph represents the mean of eight tokens, a token by token analysis reveals the same pattern.