INVARIANCE AND VARIABILITY IN TONGUE CONTACT PATTERNS: ELECTROPALATOGRAPHIC EVIDENCE

Andrew Butcher

Department of Linguistics (Arts)
Australian National University
(on leave from University of Reading, U.K.)

ABSTRACT - An attempt was made to quantify the variability of tongue contact patterns at certain stages during the pronunciation of VCV sequences, as registered by an electropalatograph system. Of specific interest were: total contact area during the vowels and the rate of change of contact for consonant closures and releases. Results are discussed in the light of some current theories of coarticulation.

INTRODUCTION

Viewed at the acoustic level, it is obvious that the time structure of speech is 'elastic'. Occurrences of the same phoneme in different environments or repetitions of the same phoneme sequence at different tempi or with different stress patterns will show variable degrees of compression and expansion at certain points in the signal. This variability is itself variable, however, in that, in some instances its occurrence and extent is more predictable than in others. Viewed from an articulatory point of view, however, it is clear that this time structure is not only elastic but multi-dimensional. Some predictions might thus be made about the effects of one sound upon another based on whether or not anatomical and physiological connections place their component gestures in the same or neighbouring articulatory dimensions. It has been suggested that these component gestures - 'icebergs' (Fujimura, 1981) or 'stereotypic articulatory movements' (Kent, 1986) - such as the opening and closing gestures for consonants, are relatively invariant. If they occur within the same dimension - as when two adjacent sounds are made with the tongue body, or maybe one with the tongue body and one with the tongue tip - they may directly influence one another, but extension of either gesture forwards or back in time will be blocked. This has been termed 'hard' coarticulation. Sequences of gestures in different articulatory dimensions, however, - i.e. those requiring unconnected speech organs - may overlap one another in time and are more likely to show variation in the degree and extent of coarticulation. This has been termed 'soft' coarticulation, as between tongue and lip gestures or tongue and velic movement and may be language- or speaker-specific. Two crucial questions, then are: are there identifiable 'relatively invariant' articulatory gestures and to what extent may they be independent of one another? This paper reports some preliminary results of a study of tongue tip and tongue body articulations in the light of the above questions.

METHOD

Electropalatography (EPG) has proved a useful technique in the study of lingual coarticulation (Butcher & Weiher, 1976; Hardcastle, 1984), but little attempt has been made to quantify the effects observed. For the present study, three adult native speakers of English read a list of nonsense words of the form VCV, where V was either /i/, /u/ or /a/ (1) and C was /p/, /t/ or /k/. Five repetitions were produced with stress on the first vowel and five with stress on the second vowel (=810 tokens or approximately 81,000 data points). Subjects wore an artificial palate which registered tongue contact by means of 62 electrodes with a sampling frequency of 133 Hz (see Brasington & Clark, 1984). At the same time, in connection with another aspect of the study, oral and nasal air flow were measured by a pneumotachograph system, and the speech signal was recorded via a microphone in the rubber face mask. EPG data were processed by a microcomputer, which produced curves representing the total number of contacts registered in the anterior third and posterior two-thirds of the palate (roughly the alveolar and palatal regions respectively) as a function of time.

RESULTS

Results of the analysis of data from one speaker's production of items stressed on the second syllable are summarized in Table 1. In order to reduce the approximately 13,500 data points for the purpose of comparison, six crucial points have been identified for each token and the mean values taken over the five repetitions. These points are: V1, the beginning of the first vowel and V2, the end of the second vowel, both identified from the acoustic signal; A, the beginning of the closure phase, corresponding to a steep rise in the number of contacts, normally coincident with, but sometimes just prior to the offset of voicing; B, the point where maximum contact for the closure is first reached; C, the point where maximum contact for the closure begins to decline; and D, the end of the release phase, where the rapid decline in number of contacts is at an end, normally coincident with but sometimes just prior to the onset of voicing. In the case of /t/, these points were identified in the alveolar totals curve, and palatal totals are given for these same points in time. In the case of /p/ and /k/, few, if any, contacts were registered in the alveolar region, and these curves are ignored. Points for /k/ were identified in the palatal totals curve. For /p/, points B and C could obviously not be determined; points A and D were taken to be at the offset and onset of voicing respectively.

Two measures of vowel-to-vowel coarticulation were applied and gave very similar results. Firstly, as a measure of coarticulatory effect throughout the vowel, the total number of contacts were calculated over the (normalized) duration of the vowel - i.e. for the first vowel the total at V1 plus half the difference between totals at V1 and A, all multiplied by the mean number of frames for that vowel type in initial position, and mutatis mutandis for the second vowel. Secondly, as a measure of coarticulation at the vowel/consonant boundary, only the number of contacts at A or D was considered. Expressed as percentages of the total possible contact area in each case, figures show a clear hierarchy for the three vowel types, with /i/ normally taking up between 40 and 50% of the palatal area, /u/ between 20 and 35% and /a/ under 10%. Contact totals would therefore appear to be a legitimate measure of coarticulation between vowels, and values for tokens where V1 ≠ V2 can be compared with the corresponding tokens where V1 = V2 (regarded as the 'neutral' environment).

As a measure of variability in the formation and release of the stop closures, the rate of increase and decrease of contact totals between points A-B and C-D respectively were calculated in terms of the number of contacts gained or lost per frame.

DISCUSSION

Vowel-to-vowel coarticulation appeared to be very similar for /p/ and /t/

tokens, and was mainly 'right-to-left' or anticipatory in nature. In general the vowel /i/ in the second syllable influenced both /a/ and /u/ in initial position, in as much as these latter sounds were produced with a greater area of contact than in the V1 = V2 context. Similarly, initial /i/ was in turn influenced by following /u/ and /a/, in that these tokens tended to show less contact than for /ipi/ and /iti/. /a/ and /u/ did not appear to influence each other in terms of palatal contact. Some 'left-to-right' or perseverative effects may be noted, whereby final /i/ preceded by initial /a/ or /u/ has less contact than in /iCi/ contexts.

Figure 1 gives a schematic comparison of data for /ati/ and /ata/. It can clearly be seen that the amount of palatal contact for /a/ is greater before /i/ than before /a/ and that this effect extends throughout the vowel. Furthermore, there is a steady increase in palatal contact throughout /ati/ tokens (very similar to the pattern found for /api/) with a relatively independent rapid alveolar gesture superimposed upon it. The consonant gesture appears to be relatively unaffected by the adjacent vowels. The approach phases in the case of this pair are almost identical, although overall we find that /t/ closures out of /i/ as V1 and releases into /i/ as V2 are slightly slower than the corresponding phases for /a/ and /u/. Thus, while the /t/ data seem to support the notion of a degree of independence of tongue tip and blade from tongue body gestures, slight variations in the execution of these gestures seem to point to invariant targets rather than invariant trajectories.

For /k/ tokens quite a different picture is found. In the case of these sequences, there are no examples of vowel-to-vowel coarticulation in either direction. As the comparison of /aka/ and /aki/ in figure 2 shows, there is no smooth progression from V1 to V2. In fact the two initial /a/ curves are very similar and any coarticulatory effect from a following /i/ is blocked by the tongue body gesture required for the intervening consonant. In fact contact area differences are in the reverse direction to that found for /p/ and /t/ tokens. Whereas realizations of /u/ remain fairly stable as regards contact area, there is a discernible trend for /i/ to have <u>more</u> contact and /a/ <u>less</u> contact in V1 ≠ V2 tokens than in /iki/ and /aka/. One possible explanation of this would be to invoke some kind of 'overshoot' effect. Those /VkV/ sequences which contain only back vowels involve relatively short articulatory trajectories into and out of the consonant and therefore less exaggerated gestures, whereas sequences involving a transition between back and front vowels <u>via</u> the velar closure require longer trajectories. The effect of the adjacent vowels on this consonant are, after all, quite massive, with consistently around twice as much contact area for tokens with adjacent /i/ as for tokens with no /i/ in the sequence. Rates of change of contact for /k/ closures and releases are indeed faster into and out of /i/ vowels than into or out of the back vowels - with the exception of the /iki/ sequence. Thus, whilst /VkV/ sequences during which there is no change in backness seem to involve a short almost tap-like closing/opening gesture, sequences involving a change in tongue retraction seem to give rise to a degree of overshoot, whereby faster gestures and therefore more exaggerated articulations of both the consonant and the vowels are produced. Once again, this is indicative of trajectories varying according to the target required, rather than of invariant 'iceberg'-type gestures being independent of context.

There is, however, in these data considerable support for the idea of the independence of articulatory dimensions - even those corresponding to structures anatomically so closely connected as the tongue body and the tip and blade system. There is, furthermore, some indication that the closing and opening rates for particular VC- and -CV types respectively remain relatively consistent. This accords with the cinefluorographic data reported by Gay (1977) and would not exclude the possibility of postulating some kind of invariance of syllable onset and coda types (see Fujimura, 1981). Analysis of the entire corpus of data will show whether this context sensitivity extends to changes in stress pattern and, even more importantly, whether these findings can be replicated with other speakers.

NOTE

(1) For typographical convenience /a/ is used for /ɑ/ throughout.

REFERENCES

Brasington, R.W.P. & Clark, C.J. (1984) A micro-software system for electropalatography, Phon. Lab. Univ. Reading Work in Prog. 4, 49-66.

Butcher, A. R. & Weiher, E. (1976) An electropalatographic investigation of coarticulation in VCV sequences, J. Phonetics 4, 59-74.

Fujimura, O. (1981) Temporal organization of articulatory movements as a multidimensional phrasal structure, Phonetica 38, 66-83.

Gay, T. (1977) Articulatory movements in VCV sequences, J. Acoust. Soc. Am. 62, 183-193.

Hardcastle, W.J. (1984) New methods of profiling lingual-palatal contact patterns with electropalatography, Phon. Lab. Univ. Reading Work in Prog. 4, 1-40.

Kent, R.D. (1986) The iceberg hypothesis: the temporal assembly of speech movements, in J.S. Perkell & K.N. Stevens (eds), Invariance and Variability in Speech Processes, 234-242, (Lawrence Erlbaum Associates: Hilldale, NJ).

token	duration (ms)					peak contacts						rates (cont/fr)		% total cont in V		% total cont next fr	
	V1	A-D			V2	V1	A	B/C		D	V2	A-B	B-C	V1	V2	V1	V2
apa	194	212			390	0	0	0		0	0			0	0	0	0
ipi	176	220			354	16	17	13		14	17			51.6	48.4	53.1	43.8
upu	153	225			360	8	9	4		5	6			26.6	17.2	28.1	15.6
api	173	216			322	1	1	10		13	18			3.1	48.4	3.1	40.6
apu	215	165			340	0	0	5		7	7			0	**21.9**	0	21.9
ipa	161	199			396	15	15	0		0	0			46.9	0	**46.9**	0
ipu	173	217			357	14	15	7		4	6			**45.3**	15.6	**46.9**	12.5
upa	129	236			364	10	10	0		0	0			31.2	0	31.2	0
upi	170	193			350	11	11	10		14	17			**34.4**	48.4	**34.4**	43.8
px̄	172	209			359												
sd	23	20			22												
		A-B	B-C	C-D													
ata	180	44	93	55	419	0	3	7	8	7	0			4.7	10.9	9.4	21.9
						0	0	14	15	1	0	2.40	1.88				
iti	146	46	110	82	350	14	16	15	13	14	18			46.9	50.0	50.0	43.8
						6	6	17	20	5	4	1.80	1.35				
utu	149	60	43	111	360	10	12	12	9	9	9			34.4	28.1	37.5	28.1
						0	0	18	18	1	0	2.25	1.13				
ati	182	52	78	68	355	1	6	7	9	12	18			**10.9**	46.9	**18.8**	**37.5**
						0	0	16	16	5	3	2.33	1.20				
atu	182	19	105	66	360	1	3	5	8	9	11			6.3	31.3	9.4	28.1
						0	0	13	13	1	0	5.10	1.35				
ita	137	61	71	84	400	14	17	11	8	5	0			48.4	7.8	53.1	**15.6**
						7	7	19	16	0	0	1.50	1.43				
itu	125	45	148	53	390	14	15	14	8	9	7			45.3	25.0	46.9	28.1
						2	3	15	15	0	0	2.03	2.10				
uta	137	48	96	49	442	9	12	10	8	6	0			32.8	9.4	37.5	18.8
						0	0	17	13	0	0	2.63	2.03				
uti	156	54	71	90	385	7	14	12	9	11	16			32.8	**42.2**	**43.8**	**34.4**
						0	0	17	15	4	3	2.33	0.90				
tx̄	155	48	91	73	385							2.48	1.50				
sd	20	12	29	19	30							0.98	0.38				
aka	149	78	84	97	356	3	2	13	13	3	1	1.05	0.75	7.8	6.3	6.3	9.4
iki	113	76	126	62	385	14	13	24	24	13	15	1.05	1.35	42.2	43.8	40.6	40.6
uku	159	72	42	116	354	8	9	14	14	7	7	0.53	0.45	26.6	21.9	28.1	21.9
aki	155	106	47	51	391	2	0	19	21	13	16	1.35	1.20	**3.1**	45.3	**0**	40.6
aku	155	78	99	61	380	0	0	12	13	8	6	1.13	0.60	**0**	21.9	**0**	25.0
ika	114	39	128	105	360	21	18	22	12	1	1	0.75	0.75	**60.9**	3.1	**56.3**	**3.1**
iku	144	41	46	143	373	17	17	23	23	6	7	1.13	0.90	**53.1**	20.3	**53.1**	18.8
uka	135	48	105	123	327	9	9	12	9	2	1	0.45	0.45	28.1	4.7	28.1	6.3
uki	150	38	114	30	417	8	9	14	19	14	16	0.98	1.20	26.6	46.9	28.1	43.8
kx̄	142	64	88	88	371							0.90	0.83				
sd	16	22	33	36	24							0.30	0.30				

Table 1. Summary of EPG data.

V1 = first vowel, V2 = second vowel, A = beginning of stop closure, B = completion of stop closure, C = beginning of stop release, D = end of stop release. Peak contact totals for /p/ and /k/ refer to palatal area, for /t/ upper row in each case refers to palatal area and lower row to alveolar area. Rates of contact build-up and decay are in contacts (gained or lost) per frame (= 7.5 ms). % total contact in vowel is number of contacts registered in palatal area throughout vowel expressed as % of total possible contacts (= 32 X mean no. of frames for that vowel type); % total contact in next frame is number of contacts registered in frame immediately adjacent to consonant closure as % of total possible contacts (= 32); figures in bold type represent difference of >3% from corresponding V1=V2 token.

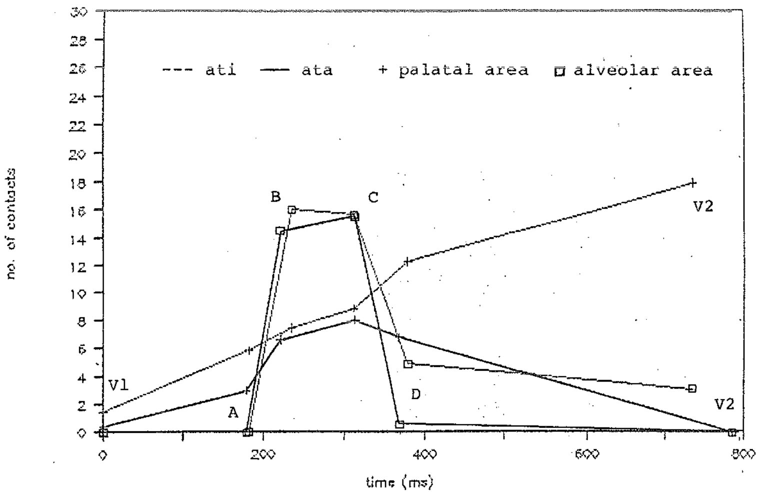

Figure 1. Contact totals as a function of time for the utterances /a'ti/ and /a'ta/ (mean of 5 repetitions).

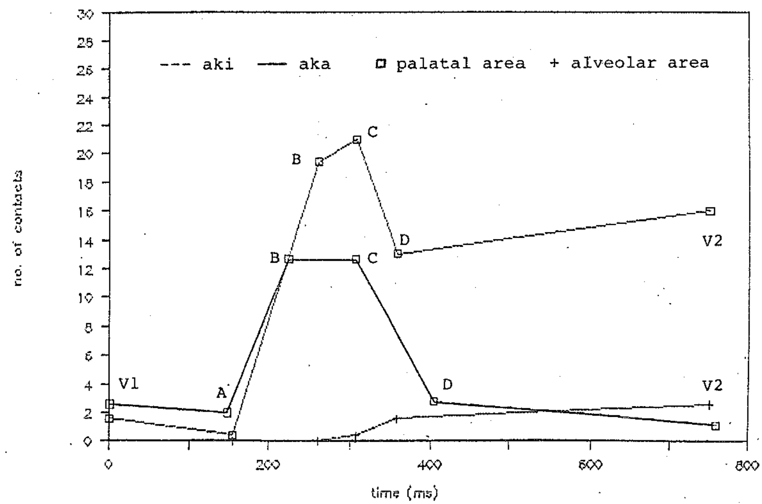

Figure 2. Contact totals as a function of time for the utterances /a'ki/ and /a'ka/ (mean of 5 repetitions).