Context Sensitivity in Some Electromyographic Signals from *M. Orbicularis Oris*

Mark Tatham
Kate Morton

This work was supported in part by grant T3/SR/6733 from the Science Research Council.

Reproduced from *Occasional Papers* 12, Language Centre, The University of Essex (1972).

INTRODUCTION

The object of this paper is to present a range of electromyographic data collected over a period of some two or three years in the Phonetics Laboratories of the Essex University Language Centre and which has not hitherto been published. Most of the experiments are small and exploratory — that is, not large enough to be presented individually. We hope, though, that since in no sense was any less care than usual taken with this work the findings will be seen as useful and contributing to the body of knowledge in the area of experimental phonetics.

EQUIPMENT AND DATA HANDLING

All electromyographic (EMG) data in this series of experiments was derived using surface electrodes from *m. orbicularis oris* of one subject (a speaker of southern British English). The surface electrodes, in the usual bipolar arrangement, were of silver cupped form, 2-3 mm in diameter affixed with Blenderm tape to the abraded upper lip approximately 0.5-1 cm apart near to and on one side of the midline. Conductivity was improved using electrode jelly. The ground electrode was a lead band approximately 1.5 cm wide strapped round the subject's wrist and assisted by electrode jelly.

Signals were amplified using high gain differential amplifiers with a frequency response of approximately 1 Hz - 10 kHz ±3 dB. Signals were recorded on either a standard direct record tape recorder 45 Hz - 18 kHz ±3d) or on a FM tape recorder (0 - 2.5 kHz ±3 dB). Those signals recorded FM were subsequently high pass filtered at 47 Hz (18 dB/octave cut off) to remove some electrode movement artefacts. Since rarely, if ever, such a method of deriving EMG signals results in a signal bandwidth extending beyond 2.5 kHz (see Mansell 1972) we may judge the two recording systems to be equivalent for all practical purposes. On all occasions an audio recording was synchronised with the EMG recording.

EMG signals were rectified and smoothed using a low pass filter providing an effective integration time of 50 ms, and then displayed on a Mingograf (frequency response 0-700Hz ±3 dB) running at a speed of l00 mm/s (accuracy ±5%).

[*note: the Mingograf was a four channel ink jet chart recorder commonly used at the time as an alternative to a pen recorder because of its wider and more linear frequency response — MT 1997.]

EXPERIMENT A

This experiment seeks to provide some answer to the question: does the vowel in a C1VC2 (where C1 is not the same segment as C2) monosyllable have any influence on the preceding of following consonant along the same (articulatory) parameter? The two pairs of (nonsense) words /bit/ *vs.* /but/ and /tib/ *vs.* /tub/ were chosen as suitable candidates, and 100 tokens of each were recorded in the frame [he-e] ([h]+schwa+segment+schwa) — assuming, of course, an English type phonology. Peak amplitudes and durations of the EMG signals were measured (see Fig. l), as being two of the least equivocal features of a rectified / smoothed EMG trace.

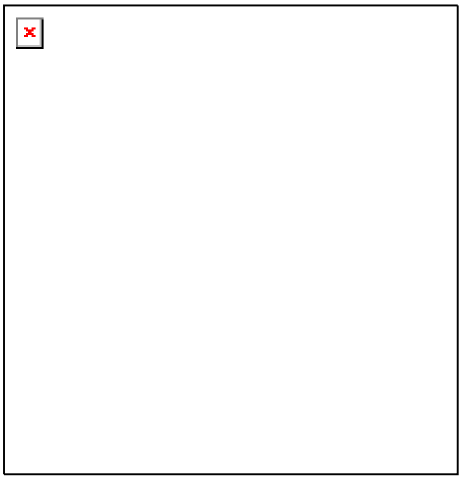

Fig. 1 Typical rectified / smoothed EMG signal, showing amplitude and duration measures taken.

Hypothesis A

There is a difference in amplitude and duration of muscle contraction (as reflected in EMG) for the *m. orbicularis oris* parameter in the pairs: [bit] — [but] and [tib] — [tub].

Results A

Durations in csec and peak amplitudes in linear arbitrary units:

	mean duration	sd	range	mean peak amplitude	sd	range
[bit]	27.98	3.59	18	67.91	9.12	35
[but]	60.31	9.77	42	64.22	10.73	45
[tib]	24.84	7.87	11	53.34	8.84	42
[tub]	47.83	7.87	32	47.39	9.32	50

Duration measurements of the EMG in [but] include signal associated with both [b] and [u] — they cannot he directly compared, therefore, with the corresponding figures in [bit]. Similarly duration measurements of the EMG in [tub] cannot be compared with the corresponding figure in [tib].

There is *no* significant difference between the mean amplitude of EMG associated with the initial plosive of [bit] and [but] (*t*-test, P=0.l).

There *is* a significant difference between the mean amplitudes of EMG associated with the final plosive of [tib] and [tub] (the null hypothesis was confirmed by *t*-test only at the P=0.001 level).

Inference A

The hypothesis is not confirmed for initial [b] — no significant difference possibly dependent upon the vowel could be found in the peak amplitudes of EMG signals from *m. orbicularis oris* associated with contraction for lip closure. However, a significant difference in EMG amplitude *was* detected associated with [b] in final position — when following [u] the amplitude was less (mean = 47.39) than when following [i] (mean = 53.34).

Thus the possible existence of a left-to-right effect is demonstrated — but not a right-to-left effect (*cf.* MacNeilage and Declerk [1968] where a right-to-left effect was observed).

Conclusion A

Two possible explanations (at least) exist for the observed left-to-right effect:

1. gamma loop or other feedback systems are signalling (in the case of final [b] preceded by [u]) that lip muscle contraction already exists and that a contraction less than for [b] following [i] is required (MacNeilage and Declerk [1968]), or
2. a context sensitive encoding has taken place at a higher level which results in a peripherally observable left-to- right effect (Wickelgren [1969], but see Tatham and Morton [1970]).

or both 1. and 2. are *both* operating.

EXPERIMENT B

If a left-to-right effect can be observed (Experiment A, this paper) in the peak amplitudes of *orbicularis oris* contraction associated with bilabial stop closure following a round vowel, then is a similar effect observed in *orbicularis oris* contraction associated with a round vowel when following a bilabial stop? Intuitively we would expect that the gesture *round —» closed* involves less effort for the closure, but that *closed —» round* involves more effort for the rounding. The experiment was conducted using items: [tut] (92 tokens), [but] (55 tokens) and [tub] (76 tokens) in the frame [he-e].

Hypothesis B

Peak amplitude of *orbicularis oris* contraction associated with [u] in the frame [tVt] will be less than in the frame [bVt] and no different in the frame [tVb].

Results B

Peak amplitude of EMG in arbitrary linear units:

	mean	sd	Range
[tut]	20.40	8.49	41
[but]	24.38	8.96	44
[tub]	19.16	5.65	32

The mean peak amplitude of the EMG from *orbicularis oris* associated with contraction for [u] in [tut] and [but] is different (*t*-test indicates them as same only at the P=0.02 level).

The mean peak amplitudes of EMG associated with [u] in [tut] and [tub] are the same (*t*-test at the P=0.1 level).

Inference B

The hypothesis is confirmed: peak amplitude of EMG from *m. orbicularis oris* associated with [u] in [tut] is less than in [but] and no different from [tub].

Conclusion B

Left-to-right effects are observable in a CVC monosyllable along the same parameter (in this case contraction of *m. orbicularis oris*) between an initial [b] and a following [tu]. A right-to-left effect is not observed between a final [b] and a [u] in [tub] compared with [tut]. No account was taken of the possibility that the [b] of [tub] could, in fact, have been initiating a following syllable ([-be]) because of the nature of the frame.

EXPERIMENT C

Following our work reported in Tatham and Morton [1968a, b] (also 1970) we have had occasion to check some of the results mentioned there. An important question which was not entirely clear then was: Is there a peak amplitude or durational difference in the EMG signal between initial [b] and final [b] on the *orbicularis oris* parameter when the intervening vowel does not employ this parameter? We examined the three utterances: [bib], [bit] and [tib] in 100 tokens of each in the fame [he-e] as in the cited reports. Once again we note the possible difficulty of the occurrence of a syllable boundary giving us either [be] or [te] as syllables, but we have chosen the frame as before. Informal experiments with frameless tokens produced similar results to the ones reported here but it was judged that the dangers of frameless tokens were too great for a formal experiment.

Hypothesis C

[b] initially and [b] finally in [bib] do not display any variation in either peak amplitude or duration of the *orbicularis oris* EMG associated with lip closure, and likewise [bit] and [tib].

Results C

[EMG durations for *orbicularis oris* in [b1ib2] could not be measured since there was overlap during the vowel — i.e. the offset of [b1] overlapped the onset of [b2].] Peak amplitudes in arbitrary linear units and durations in csec:

	mean peak amplitude	sd	range	mean duration	sd	range
[b1ib]	62.18	9.44	42			
[bib2]	57.77	10.18	65			
[b1it]	67.91	9.12	35	27.98	3.59	18
[tib2]	59.34	8.94	42	24.84	7.87	11

In [b1ib2], the EMG peak amplitude of [bl] is greater than that o£ [b2] (*t*-test does not support the null hypothesis).

Comparing the bilabial stops in [b1it] and [tib2] the EMG peak amplitude of [bl] is greater than that of [b2] (*t*-test does not support the null hypothesis).

The durations of the EMG associated with [bl] in [blit] and [b2] in [tib] are similar (*t*-test, P=0.1) .

Inference C

The hypothesis is not confirmed with respect to the peak amplitude measurements of EMG from *orbicularis oris* in either pair of utterances ([bib] *vs.* [bib] and [bit] *vs.* [tib]). On the only measurable pair ([bit] *vs.* [tib]) it is strongly confirmed with respect to the durational measurements.

DISCUSSION

Consider the following rules which might be written to summarise the amplitude data:

1. Voiced bilabial stops exhibit less lip contraction when following a rounded vowel.
2. Rounded vowels exhibit more lip contraction when following a voiced bilabial stop.
3. Elsewhere lip contraction is as in a neutral frame.

These must be expressed formally (using, in this case, the notation we are used to in Transformational Phonology). Firstly it will be necessary to derive phonetic segments from phonological segments:

4a. [+bilabial, +stop] → [+x *orbicularis oris* m., +y Y m., +z Z m.]

4b. [+round, +vowel] → [+a *orbicularis oris* m., +b Bm., +c C m.]

where {x, y, z, a, b, c} are appropriate neutral environment integers indicating degree of contraction of muscles {*orbicularis oris*, Y, Z, B, C} respectively, and where the set of muscles {*orbicularis oris*, Y, Z} is recruited for lip closure and the set {*orbicularis oris*, A, B} recruited for lip-rounding.

Notice, in 4a, the features [voice] and [nasal] remain unspecified as a prediction that these two parameters are irrelevant in the rules — i.e. precisely similar results will be obtained for [p] and [m].

The following generalisations can be set up concerning redundancy in the specification of the phonetic segment:

5a. If: [x orbicularis oris m., y Y m.] then: [- , - , z Z m.]

5b. If: [a orbicularis oris m., b B m.] then: [- , - , c C m.]

These conditions state (where Y,Z,B,C are each one or several muscles involved in the appropriate gestures) that a particular set of muscles is associated uniquely with a particular gesture when specified to contract a certain degree. Henceforth:

5a(i). [x oo. m., y Y m.]

will mean: 'set of muscles recruited for bilabial stops and specified with respect to amplitude of contraction in neutral environment (i.e. environment not involving a segment using this set).'

5b(i). [a oo. m., b B m.]

will mean 'set of muscles recruited for round vowels (lips only considered0, etc.'

Then: (where p and q are integers)

6a. [x oo. m., y1 Y m.] → [x-p oo. m., y2 Y m.] / [a oo. m., b B m.] —

where y1 does not necessarily equal y2 and where Y and B may have identity.

6b. [a oo. m.. b1 B m.] → [a+q oo. m., b2 B m.] / [x oo. m., y Y m.] —

where b1 does not necessarily equal b2 and where B and Y may have identity.

Rules 6a and 6b are of course general rules. Using the numerical data from the experiments we are in a position to give tentative values to the integers p and q.

It is the case that the bilabial stop [b] exhibits a 10% reduction in EMG peak amplitude following the rounded vowel [u], and that [u] exhibits a 20% increase in EMG peak amplitude following [b]. Then

7a. [x oo. m., y1 Y m., -voiced, -nasal] → [90%x oo. m., y2 Y m.] / [a oo. m., b B m.] —

7b. [a oo. m., b1 B m., +high, +back] → [120%x oo. m., b2 B m.] / [x oo. m., y Y m.] —

Rules 7a and 7b are more specific to the data given by the experiment.

EXPERIMENT D

A measure which has not hitherto been taken on EMG tracings of speech muscles is that of angles of onset and offset. The rectified/integrated tracing is roughly triangular in shape, but is not in a normal utterance an isosceles triangle where we are considering a bilabial plosive. The data of the previous experiments enables us to measure onset angles for initial plosives and offset angles for final plosives for all items in the set, and in addition, some, but not all, offset angles of the initial plosive and onset angles of the final plosive. Angles were measure by joining the point of onset or offset of the EMG signal in time and the point of maximum amplitude with a straight line and measuring the angle between this line and the baseline.

Hypothesis D

Onset and offset angles for the initial plosive will be the same between [bit] and [but] but different between [tib] and [tub] — in this last pair [b] in [tib] will have greater angles than in [tub]. We are considering still of course *m. orbicularis oris*.

Results D

Angles in degrees:

	mean onset	sd	range	Mean offset	sd	range
[bit]	80.09	2.48	12	75.31	3.46	19
[but]	75.59	5.56	26			
[tib]	79.65	2.52	12	71.89	3.64	25
[tub]				69.19	4.64	22

1. Onset angles [bit] *vs*. [but] are different ([b+i] > [b+u]).
2. Offset angles [tib] *vs*. [tub] are just different ([i+b] > [u+b]) (same at P=0.02).

Inference D

The hypothesis is not confirmed. There was a greater difference between onset angles for [b] in [bit] and [but] than between offset angles for [b] in [tib] and [tub] — the opposite of the hypothesis.

Conclusion D

Since amplitude and durational measurements were strongly confirmed as equal in initial [b+i] and [b+u] position and different in [i+b] and [u+b] position, we have to explain the difference in angles. Tentatively we can suggest that the centre point of the duration has shifted with respect to the timing of the peak amplitude — this can easily be confirmed by measurement (data not included in these notes).

Analysis of the onset and offset angles in [bib] and [bub] showed the following results:

	mean onset	sd	range	Mean offset	sd	range
[b1ib]	72.41	5.9	23	73.44	3.63	29
[bib2]	80.03	2.11	12	70.63	4.76	25
[b1ub]	74.55	5.44	25			
[bub2]				70.35	5.86	30

[bl] and [b2] onsets in [bib] are different — [bl] > [b2].

[bl] and [b2] offsets in[bib] are different — [b1] > [b2].

[bl] onsets in [bib] and [bub] are the same.

[b2] offsets in [bib] and [bub] are the same.

The onsets and offsets of [bl] and [b2] in [bib] were not expected to be different; and similarly the offsets of [bl] and [b2] in [bub] were expected to be different — but the data shows that there are the same. Measurements of angles of onset and offset do not confirm directly the left-to-right effect observed in simple measurement of amplitude peak and duration.

DISCUSSION

Composite use of durational, amplitude and onset and offset angle data permits us to form the following hypothesis:

[bub] along the *m. orbicularis oris* feature) consists of a simple summation of the muscle behaviour of these segments in neutral frames. (The alternatives, of course, are i. that they are superimpositions with no summation or with a weighted summation, or ii. that behaviour is completely different when these segments are juxtaposed by the phonology.)

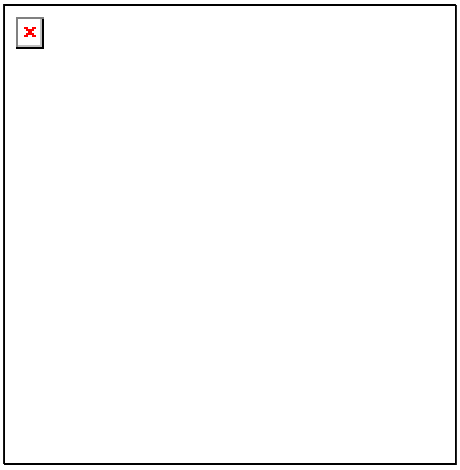

Fig. 2 Stylised version of [bub] in the frame [he—e].

Fig. 2 presents a stylised tracing of [bub] based on averaged measurements of

a. EMG peak amplitude,
b. EMG duration,
c. onset and offset angles,
d. timing of onset, offset and peak occurrences

all related to the plosive bursts along the audio time scale.

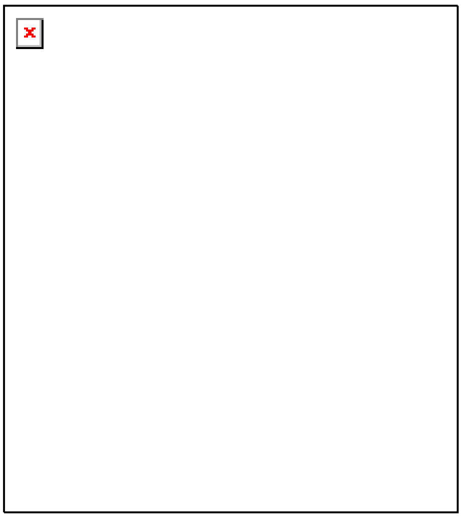

Fig. 3 Simple superimposition of [b]+[u]+[b], obtained from [bit], [tut] and [tub].

Fig. 3 presents a simple superimposition of the same average data obtained from the three utterances [bit], [tut] and [tub].

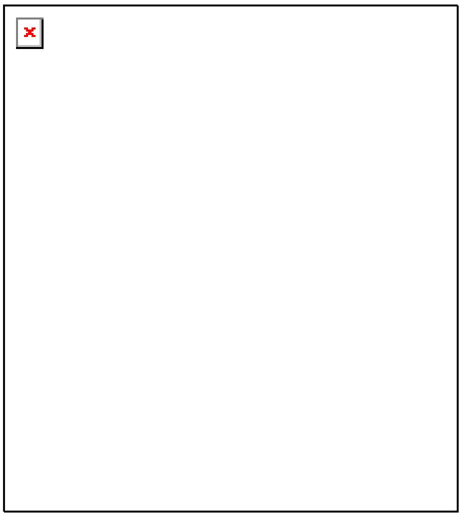

Fig. 4 Simple superimposition of [b]+[u]+[b] obtained from [bit], [tut] and [tub], but with amplitude summation. The dotted line indicates where summation has occurred.

Fig. 4 presents a simple superimposition of the average data from Fig. 3, but with simple amplitude summation where there is segment overlap on this feature.

Notice that visually the synthesised Fig. 3 (no summation) looks more like Fig.2 (actual [bub]) than does Fig. 4 (summation included). We are at least able to conclude, therefore, that along this one parameter we cannot regard segments as simply isolated in time with their overlaps summed. Nor is it simply the case that the overlaps are 'cancelled' in some way — since clearly the superimposed version is providing a peak amplitude for [u] which is lower than the actual non-synthesised amplitude of [u] in [bub]. In other words [bub] is not [b]+[u]+[b] with or without summation in the amplitude domain.

We can only conclude that a left-to-right effect is operating in [u] of [bub] that demands greater amplitude for *m. orbicularis oris* contraction following lip-closure ([b]) than following a segment without lip contraction ([t]).

References

MacNeilage, P. F. and J. L. Declerk (1968) 'On the Motor Control of Coarticulation in CVC Monosyllables'. Paper read at the 1967 Conference on Speech Communication and Processing, Boston, Mass., Nov. 1967; and *I.E.E.E. Transactions on Audio and Electroacoustics*

Mansell, P. (1972) 'Theoretical and Experimental Approaches to a Model of Speech Production'. University of Essex Language Centre, *Occasional Papers* No. l4 (forthcoming)

Tatham, M. and Katherine Morton (1968a and 1970) 'Some electromyography data towards a model of speech production'. University of Essex Language Centre, *Occasional Papers* No. 1; and *Language and Speech* 12:39

Tatham, M. and Katherine Morton (1968b) 'Further electromyography data towards a model of speech production'. University of Essex Language Centre, *Occasional Papers* No. 1

Tatham, M. and Katherine Morton (1970) 'Explaining some apparently context sensitive effects in speech'. University of Essex Language Centre, *Occasional Papers* No. 9

Wickelgren, W. A. (1969) 'Context sensitive coding, associative memory, and serial order in (speech) behavior'. *Psychological Review* 76:1