Explaining Some Apparently Context Sensitive Effects in Speech

Mark Tatham
Kate Morton

Reproduced from *Occasional Papers* 9 (1970) Language Centre, University of Essex.

This work was supported by Award B/SR/6733 from the Science Research Council.

In a paper (Wickelgren [1969]) published in 1969 Wickelgren suggests that an associative-chain model is possible as an explanation of speech behaviour. He proposes that the basic units of speech production are to a certain extent context sensitive, taking apparently the form of allophones indexed in some way as to their immediately preceding and immediately following unit context. Such a model runs counter to the majority of current work on speech production models (Fromkin [1968] MacNeilage and Declerk [1968], Tatham [1970], etc.) and has been seriously criticised by MacKay [1970], MacNeilage [1970] and Whitaker [1970], who believe that any apparent surface context sensitivity in speech is the result of processes lower in the system than implied by Wickelgren.

According to our understanding of Wickelgren's proposal,* the [k] in a sequence such as [uka] is that allophone of [k] contingent upon its immediate context: [u–a] — i.e. a [k] embodying attributes associated with a preceding [u] and a following [a]: perhaps best represented using Wickelgren's own [1969] notation: [uka]- Similarly [k] in the sequence [aku] will be [aku] — i.e. although possessing identity of basic '[k]-ness' in some sense with [uka] there will be detail variations attributable to the differing contexts. The basic units [k] used in these distinct utterances will- be different. Similarly we have the pair: [uki] and [iku].

[*footnote:* *We have discussed the work reported here with Wickelgren and, if we interpret his remarks correctly, we have taken only a very simple interpretation of his suggestions as our starting point.]

One of the articulatory features observed to differ during the [k] of, say, [ku] and [ki] is that of lip rounding/spreading. This is, of course, a secondary feature anyway of [k], but nevertheless one which Wickelgren [personal communication] would originally have included in his context sensitive indexing of the basic [k]. Lip-rounding can be observed before and during [k] in [ku], and lip spreading before and during [k] in [ki] — where [k] is taken to be that portion of the utterance corresponding to the velar stop. Similarly lip rounding and lip spreading can be observed respectively in [uk] and [ik] during [k].

If we now make the utterances symmetrical by using the vowel [a] which is neutral with respect to lip **gesture**, we have the following pairs: [uka] vs. [uki] and [aku] vs. [iku] with [uka] vs. [uki] and [aku] vs. [iku] respectively. In the first utterance of each pair lip rounding is 'free' to occur during [k] and may spill into [a], but in the second of each pair lip rounding is temporally constrained by the lip spreading **gesture** required for [i] — assuming that as lip rounding for [u], 80 lip spreading for [i] is an essential feature of the vowel. In [uki] and [iku] rounding and spreading can be observed to meet during the [k] as a rapid transition from one state to the other. We would expect therefore that motor control for lip rounding associated with [u] and spilling into [k] would be severely temporally limited when the other vowel is [i] as opposed to when it is [a] — where no opposite lip **gesture** is required.

Hence we can establish the hypothesis according to Wickelgren's model:

H1: Muscle contraction associated with lip rounding in the utterances [uka] and [aku] will be longer in duration than during the utterances [uki] and [iku] paired respectively.

Or, alternatively according to the opposing models:

H0: Muscle contraction associated with lip rounding the utterances [uka] and [aku] will be no longer than during the utterances [uki] and [iku] paired respectively.

EXPERIMENT

The following test utterances were used in the experiment:

[úka] [uká] [akú] [áku]

[úki] [ukí] [ikú] [íku]

with stress variations as marked, in case the positioning of stress affects **parameters**. A frame: [hek-ke] was used to stabilise rhythm.

The utterances were recorded by one subject 50 times (not all traces were measurable) each (randomised), monitoring the electromyographic signals from *m. orbicularis oris* (associated with lip rounding) and *m. zygomaticus major* (associated with lip spreading) (henceforth neglected in this paper).

A pair of small (4 mm. diameter) bi-polar silver surface electrodes were affixed with Blenderm tape to the upper lip, spaced about 5 mm. and as near as possible to the midline. Electrode jelly was used on a slightly abraded skin surface to improve contact. The differentially amplified EMG signal was recorded on an Ampex SP300 tape-recorder in the FN mode running at a speed of 15 ips. The audio was recorded simultaneously on another channel of the recorder correctly equalised direct record mode).

Measurements of the duration of EMG activity were made on Mingograf traces of the signal after high-pass filtering at 25 Hz (to remove electrode movement artifacts), rectifying and smoothing by low-pass filtering at 16 Hz. The measured traces therefore were of a processed version of a raw signal essentially flat from 25 Hz (the setting of the high-pass filter) to 2.5 kHz (the upper frequency limit, for a flat response, of the tape-recorder). Filters had a cut-off characteristic of 18 dB/octave. The Mingograf had a fairly level response from DC to 800Hz and was run at a speed of 100 mm./sec.; measurements were accurate to the nearest lcsec.

The EMG signal was taken to begin when it rose above a level determined visually on the traces as being the maximum noise level, and to end when it fell below this same level (Fig. 1).

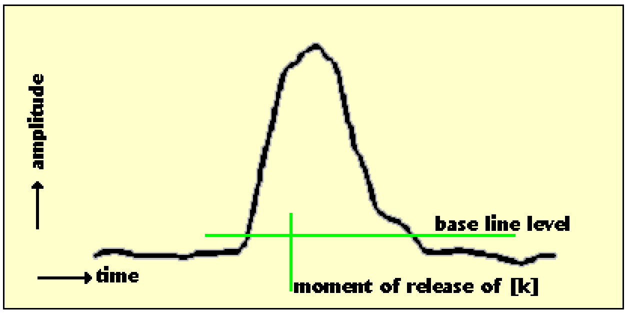

Fig. 1 Typical rectified/smoothed EMG trace from *m. orbicularis oris* during the utterance [iku]. Duration measurements were taken from the point where the trace rose above the base line level to where it fell again below the level.

RESULTS

Duration of EMG associated with [u] taken from *m. orbicularis oris*:

	mean duration in csec.	s	range	number of tokens
[úka]	40.22	8.81	37	49
[úki]	44.18	10.34	33	39
[uká]	39.18	9.91	36	49
[ukí]	40.64	10.18	39	47
[akú]	37.31	5.24	25	48
[ikú]	39.37	4.18	28	46
[áku]	28.47	5.63	20	49
[íku]	30.32	5.36	19	47

1. [úka]/[úki]: t = 1.92 (not significant beyond 0.05)
2. [uká]/[ukí]: t = 0.71 (not significant beyond 0.05%)
3. [akú]/[ikú]: t = 2.14 (just significant beyond 0.05 but not beyond 0.04)
4. [áku]/[íku]: t = 1.42 (not significant beyond 0.054)

CONCLUSION

Differences were not significant beyond the 0.05% level (except for [akú]/[ikú] and this not beyond the 0.04 level) and this was taken as an indication that Ho is the correct hypothesis. This means that there is no significant change in the controlled muscle contraction for lip rounding associated with [u] between items in the pairs [uka] vs. [uki] and [aku] vs. [iku] although the change from [a] to [i] requires a change from neutral to spread for the lips at least during the [i] .

DISCUSSION

Of course, as pointed out earlier, the lip **parameter** for [k] is of secondary importance — or indeed of no importance for the correct articulatory realisation of that segment: but it is exactly for that reason, we can guess, that it assumes the approximate polarity (round vs. spread) or an adjacent vowel segment. This is why we can observe visually a definite difference in the lip states between [aku] and [iku]: during the latter there is a rapid change from spread to round during [k].

What we wish to point out, though, is that the lip **parameter** does seem at first glance an obvious candidate for context sensitive effects according to Wickelgren's model. What may be wrong with this observation and similar cases is that at first glance it is not obvious that the spatial configuration of the lips is a result of two muscle groups (at least) operating somewhat antagonistically. There is contraction of *m. orbicularis oris* for lip rounding and of *m. zygomaticus major* (among others) for lip spreading — but both can and do contract together in just cases such as the utterances used here, producing an overlap between **gestures** — the one increasing, the other decreasing — resulting in the observed transition in the spatial configuration of the lips themselves.

It is interesting to note, though, that the muscles appear to behave the same regardless of the possible occurrence of this antagonistic effect. In [aku] and [iku], for example, *m. orbicularis oris* contracts for similar times: it does not seem to be the case that the prior occurrence (in the second) of [i] delays the onset, or shortens the time given to contraction for rounding.

Now, on the level of merely observed spatial configuration of the lips during [k], then indeed [aku] is different from [iku], but at the level of controlled muscle contraction for at least this **parameter** [aku] is not different from [iku]. And, once again, at least here there is an obvious very peripheral context sensitivity associated with a coarticulatory effect which appears to be unrelated to any basic unit of speech production. That is, [aku] and [iku] exist after coarticulation, but on this **parameter** only [k] exists before coarticulation, or something context sensitive in a much less gross way.

References

Fromkin, Victoria [1968] Speculations on Performance Models, *Journal of Linguistics* 4

MacKay, D. G. [1970] Spoonerisms: the structure of errors in the serial order of speech, *Neuropsychologia* 8

MacNeilage, P. F. [1970] Motor control of serial ordering of speech, *Psychological Review*

MacNeilage, P. F. and DeClerk, J. L. [19681 On the motor control of coarticulation in CVC monosyllables, Haskins Labs. *SR-12*

Tatham M. [1970] A speech production model for synthesis-by-rule, *Working Papers in Linguistics* 6, Computer and Information Science Research Center, The Ohio State University

Whitaker, H. A. [1970] Some constraints on speech production models, University of Essex Language Centre, *Occasional Papers* #9

Wickelgren. W. A. [1969] Context sensitive coding, associative memory, and serial order in (speech) behavior, *Psychological Review* 76.1