Proceedings of The Institute of Acoustics

VARIATION IN GLOTTAL OPEN AND CLOSED PHASES FOR SPEAKERS OF ENGLISH

Peter Davies, Geoffrey Lindsey, Hilary Fuller(1) and Adrian Fourcin

UCL and (1)NPL, SPAR and SPAC Groups

INTRODUCTION

The experiment described here forms part of an investigation into variation in the acoustic excitation produced by the larynx. A previous pilot investigation using a limited amount of data [1] had indicated:

a) that there appeared to exist a degree of relationship between fundamental frequency and the ratio of the durations of the open and closed phases of the larynx cycle

b) that a considerable amount of inter-speaker variation could occur.

The purpose of the present experiment was to process larger quantities of data in order to obtain a more reliable estimate of the more salient features of open and closed phase variation. Both the present paper and reference [1] used an electrolaryngograph as a simple method of monitoring these aspects of vocal fold activity. This technique is useful in investigating vocal fold activity because the waveform it provides is generated directly by the the vocal folds; some of its intrinsic limitations are discussed below. Note the following notation, to be used subsequently in this paper: Lx, laryngograph signal; Tx, duration of larynx cycle (derived from Lx); OP, duration of open phase of larynx cycle (derived from Lx); CP, duration of closed phase of larynx cycle (derived from Lx).

EXPERIMENTAL METHOD

Subjects read fixed English texts, a laryngograph being used to monitor the activity of the vocal folds (the acoustic pressure wave was also sampled using a condenser microphone, but this data played no part in the experiment reported here). The laryngograph signal was sampled (either directly or from a recording made on videotape using a Sony Betamax PCM recorder) at 5000 Hz by a Masscomp MC-500 computer equipped with a Masscomp AD12F 16-bit analogue-to-digital converter plus Masscomp CK10 clock, and the signal was processed in real time to extract the time of vocal fold closure and vocal fold opening for each larynx cycle. In normal phonation the vocal folds close together rapidly, and this is a well-defined event which corresponds to the point of maximum positive time derivative of the Lx signal (which represents the electrical conductivity across the larynx). There is however less agreement concerning the opening of the vocal folds: the vocal folds "peel" apart from below, and it is more difficult to identify a unique event in the Lx waveform which corresponds to the opening of the vocal folds. The view taken here is that there is no method which can in the present state of our knowledge be regarded as giving a unique "right" answer: the opening point was calculated by three different methods, and it is necessary to consider the extent to which results derived by the three different methods were consistent. The three different ways of defining opening were:

a) opening corresponds to a position where Lx has reached a value that divides the range from peak to trough of that cycle in a fixed ratio. Previous experimental work by one of the present authors (A. J. Fourcin) indicated a value of approximately 7:3 for the ratio (peak Lx - opening point Lx):(opening point Lx - trough Lx)

b) opening corresponds to the point of maximum negative time derivative of Lx, i.e. opening is an event that involves relatively rapid changes (though not as rapid as closure) in the state of the larynx. This too is supported by results such as [2] p. 318-9, and [3] p. 117-8, particularly in cases where the Lx signal has a "knee" effect on the closing side of the peak

c) opening occurs when the Lx signal has the same value as it had at the previous closure. This is clearly a reasonable assumption, though empirical evidence is slight (reference [2] p.318-319 might support this assumption also; reference [3] p.117-118 seems inconclusive with respect to this point).

Figure 1 illustrates the difference between these three measures of opening position.

Totals were accumulated of the number of larynx cycles having a given Tx and a given OP, and from these tables scattergrams as shown in figures 2 - 12 were plotted. In these scattergrams, the density at a given point represents the number of occurrences of a larynx cycle with Tx given on the horizontal axis and OP, CP or $\log_e$ (OP/CP) plotted on the vertical axis.

Subjects (all native speakers of English) comprised 8 normal female speakers, one pathological female speaker and 4 normal male speakers.

Texts used comprised:

a) a variety of passages in continuous English of various lengths from about 1 minute to about 10 minutes, used to investigate whether the distributions converged on stable values; these texts were not read by all the speakers

b) a standard text ("Arthur the Rat") of some 5 minutes duration, read by all the speakers

c) sequences of "nonsense" syllables of the form CV where C is an English consonant phoneme and V is an English vowel phoneme, each syllable being said with high falling intonation; these were said by two phonetically-trained normal female speakers.

RESULTS

The results presented here are tentative, but it seems valid to make the following observations:

a) the distributions converge with increasing sample time, and for samples of non-pathological speech above about 2 minutes, apparently present largely stable results.

b) the three methods of calculating position of glottal opening present consistent results (see figures 2, 3, 4), though some are clearly better than others. In general, method (c) gives somewhat shorter closed phases, and this is reasonable in view of figure 1, which is based on a typical shape for an Lx waveform. Opening points in figures 5 to 14 were all calculated by method (a).

Note that, in the case of log(OP/CP), the curious effect of points lying along families of curves is an artefact.

c) Earlier research based primarily on nonsense VCV sequences with high falling intonation (reference [1]) suggested that speakers tend to maintain a relatively constant OP duration while varying

CP with Tx. Distributional analysis of a text consisting entirely of CV syllables with high falling intonation (figures 5, 6) for speaker A (woman) is consistent with this suggestion: the distribution of CP versus Tx forms a sharp diagonal pattern showing that they vary proportionally, while the distribution of OP versus Tx shows a much flatter pattern. This phenomenon is weaker, though still discernible, for a text of natural continuous speech by the same speaker (figures 7, 8). Moreover, the phenomenon appears to be speaker-dependent: for some speakers such as speaker B (man) (figures 9, 10) both OP and CP vary equally with Tx, while for others such as speaker C (woman) (figures 11, 12) it seems to be CP which is relatively constant while OP varies with Tx. As CP tends to be less than OP, this leads to interesting variations in the ratio OP/CP with Tx; for speakers such as speaker A the ratio falls with increasing Tx (figures 2, 3, 4), while for speakers such as B and C the ratio remains approximately constant (figure 13).

d) Another interesting phenomenon is the tendency in the case of certain speakers such as speaker A (figures 7, 8) for the distributions of OP and CP versus Tx to divide into two approximately linear regions.

It is too early to reach firm conclusions regarding these phenomenona, but it is hoped that further distributional analysis by programs which will selectively examine certain parts of speech data, such as vocalic or accented portions, may yield broader generalisations regarding OP or CP constancy.

e) The distributions of the pathological speaker (speaker D, woman, figure 14) differ in a number of respects from those of the other speakers, and the unusual features of these distributions occur consistently with all three methods of calculating the opening point; note that the main part of the distribution shown corresponds to an exceptionally short range of Tx, but two separate, weaker areas of the distribution occur for Tx values exceptionally large for a female speaker. The polypoid condition of this speaker's vocal folds would ordinarily be expected to interfere with CP or OP regularity, and this is clearly shown here in a way which lends itself to clinical interpretation and quantitative assessment.

DISCUSSION

These first findings indicate that the phenomena described in this paper are a possibly important aspect of voiced excitation which can contribute to both inter-speaker and intra-speaker variation.

VARIATION IN GLOTTAL OPEN AND CLOSED PHASES

Our parallel task is concerned both to extend the analyses to obtain a statistically more rigorous description of the phenomena, and to assess their importance in both acoustic analysis, and synthesis.

REFERENCES

[1] Lindsey, G. A., Davies P. J. and Fourcin, A. J. 1986 "Laryngeal Coarticulatory Effects in English VCV Sequences", in Proceedings of the IEE Conference on Speech Input/Output, March 1986.

[2] Fourcin, A. J. 1974 "Laryngographic Examination of Vocal Fold Vibration", in Ventilatory and Phonatory Control Systems, ed. Barry Wyke (O.U.P. London)

[3] Fourcin, A. J. 1982 "Laryngographic Assessment of Phonatory Function", in Assessment of Vocal Pathology, ed. C. L. Ludlow (ASHA Report no. 11)

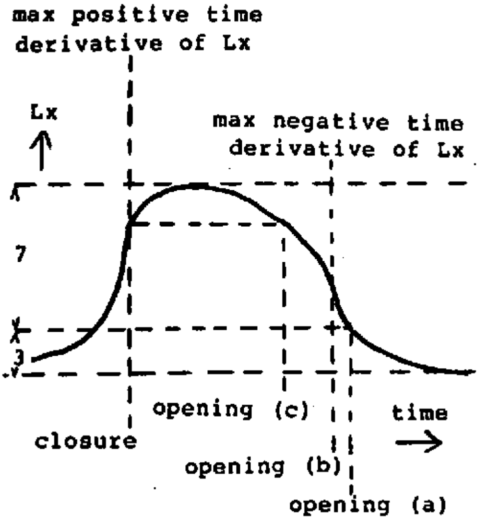

Figure 1
Lx opening and closing points

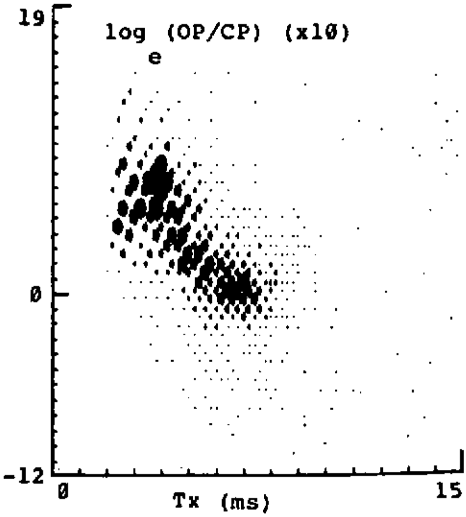

Figure 2 Speaker A
natural continuous speech
opening points by method (a)

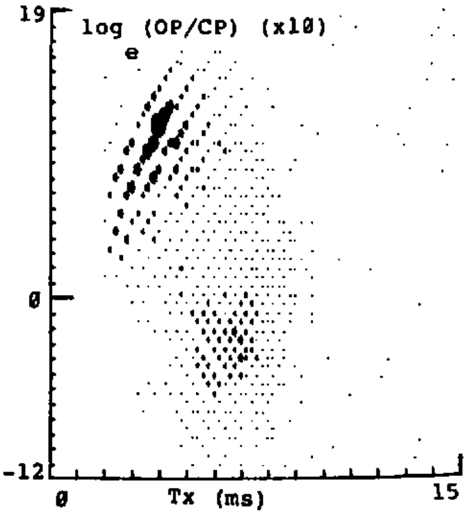

Figure 3 Speaker A
natural continuous speech
opening points by method (b)

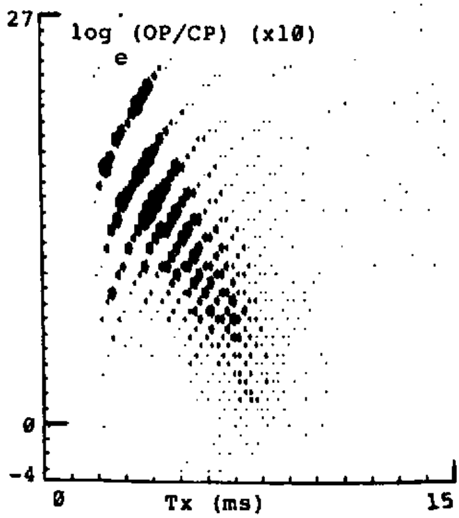

Figure 4 Speaker A
natural continuous speech
opening points by method (c)

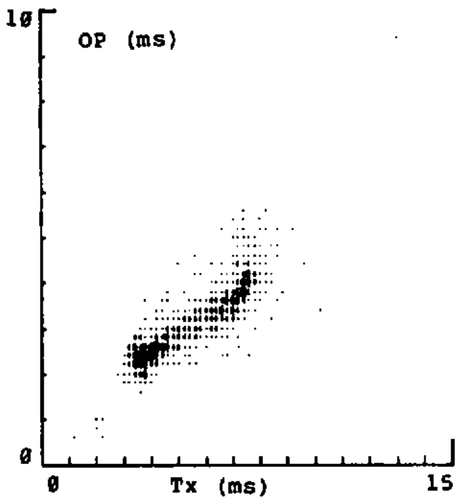

Figure 5 Speaker A
CV's, high falling intonation

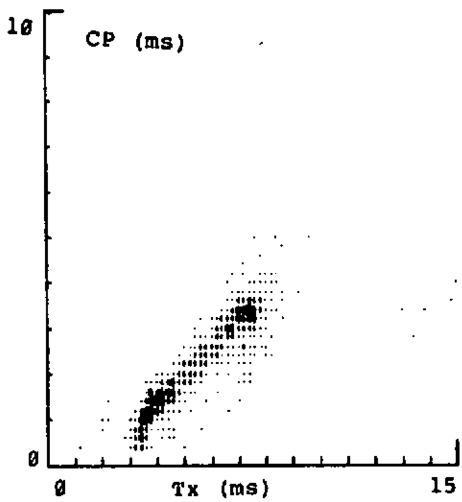

Figure 6 Speaker A
CV's, high falling intonation

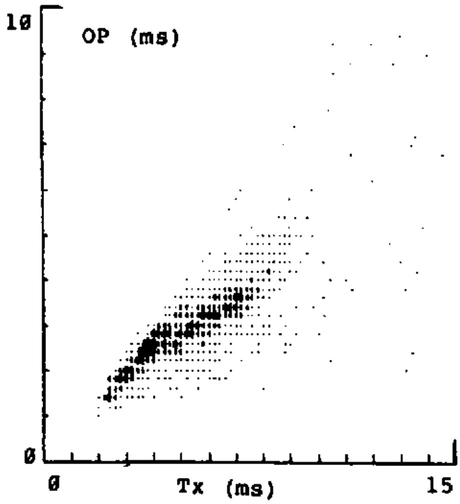

Figure 7 Speaker A
natural continuous speech

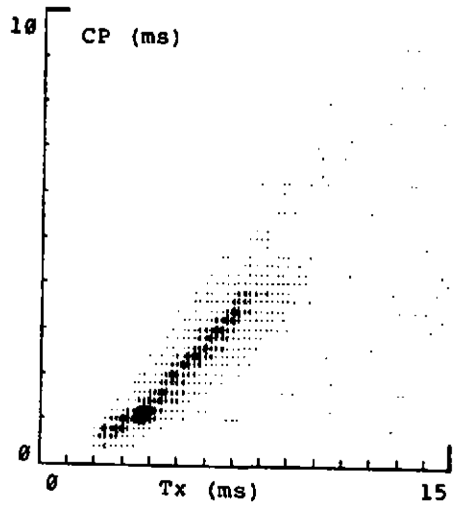

Figure 8 Speaker A
natural continuous speech

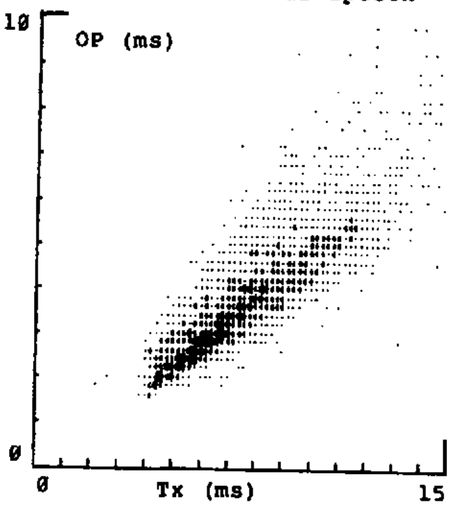

Figure 9 Speaker B
natural continuous speech

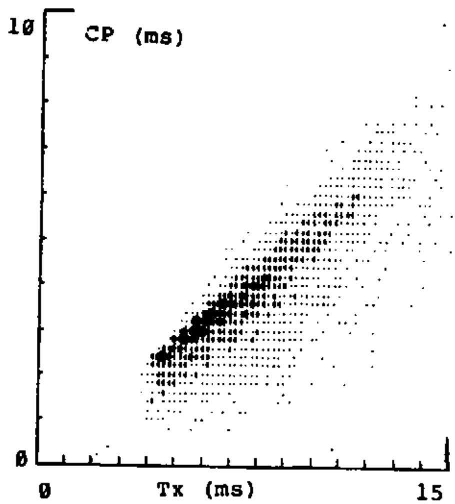

Figure 10 Speaker B
natural continuous speech

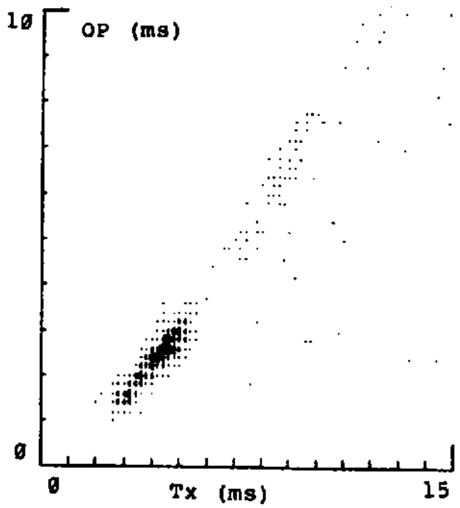

Figure 11 Speaker C
natural continuous speech

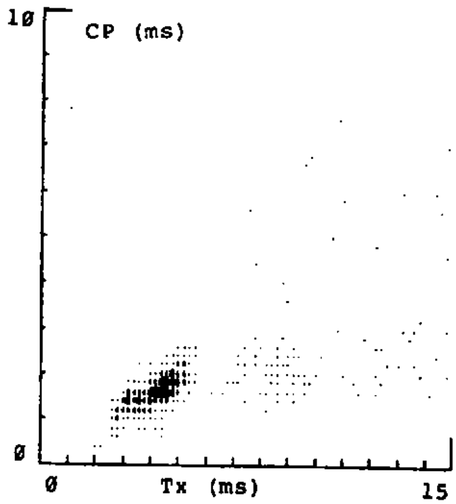

Figure 12 Speaker C
natural continuous speech

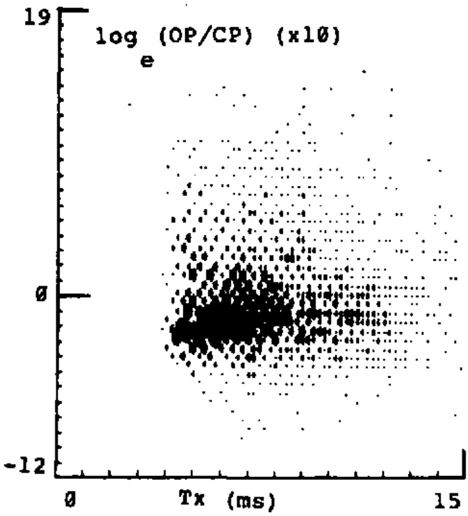

Figure 13 Speaker B
natural continuous speech

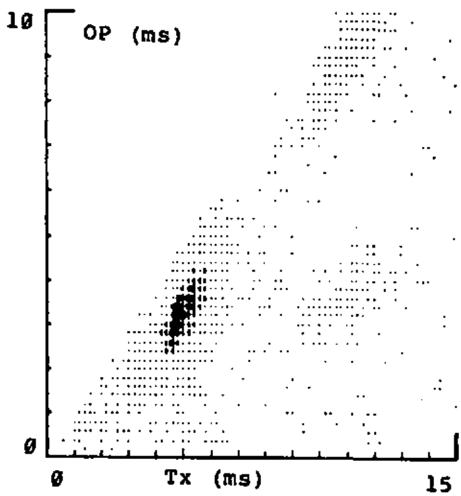

Figure 14 Speaker D
natural continuous speech