GLOTTAL CONTROL OF ASPIRATION AND OF VOICELESSNESS

AILBHE NI CHASAIDE

CENTRE FOR LANGUAGE AND COMMUNICATION STUDIES,
TRINITY COLLEGE, DUBLIN 2, EIRE

ABSTRACT

There are two different, though potentially complementary explanations of how glottal control of voicelessness and aspiration is effected; a) by directly controlling the degree of glottal aperture during stop closure and specifically at oral release; b) by controlling the precise timing of the abduction/adduction gesture.

Photoelectric glottographic data for aspirated (pre- and post-) and unaspirated stops in Icelandic and Irish uttered in differing stress conditions suggest that ; glottal aperture is directly controlled; this and other "strategies" observed may be necessary to maintain voice offset and onset targets under differing aerodynamic conditions; such strategies might be best regarded, not as correlates of stress as such, but rather as evidence of a more general laryngeal response to changes in aerodynamic conditions.

INTRODUCTION

Two rather different proposals have been made regarding the laryngeal mechanism which determines the presence and duration of aspiration. Kim /1/ has hypothesised that the degree of glottal opening during stop closure, and crucially at stop release determines the duration of aspiration. The wider the glottal aperture at the instant of release, the longer the aspiration that ensues; if the vocal folds are already adducted at that point, there will be no aspiration. Working from Kim's data, Catford /2/ gives an approximate graph showing the relationship between glottal aperture during stop closure and aspiration duration.

More recently, Löfqvist et al. /3/ have questioned an underlying assumption of the earlier work, namely, that speakers directly control the degree of glottal opening as a means of determining the duration of aspiration. These authors suggest that "voluntary control of the size of glottal opening is rather poor and that subjects are unable to make very fine graded adjustments along this dimension." Löfqvist /4/ also points out that the glottal gesture is a relatively fixed ballistic opening/closing cycle; once peak glottal opening has been attained, the closing gesture tends to start immediately rather than maintain a static open position. Thus, rather than direct aperture control, Löfqvist proposes that control is exerted on the timing of the laryngeal gesture, to which Peak glottal opening (PGO) is an important index. The later the PGO, the greater the glottal abduction at stop release, and the longer the aspiration. The converse should be true for preaspirated stops. It follows from Löfqvist's account that, all else being equal, one might expect greater peak glottal opening for aspirated than for an unaspirated stop: such a difference would however be a secondary consequence rather than the primary control parameter.

It was felt that voiceless stops in Icelandic and Irish (yielding pre-, post- and unasapirated types) across differing stress conditions might provide a testing ground for these two models of aspiration control. The durations of preaspiration /5/ and postaspiration /6/ can be much shorter in unstressed than in stressed syllables, and it would be of interest to consider which of the above two models might best account for those differences.

MATERIALS

Recordings of four short data sets in Icelandic and Irish (a single subject in either case) included the following signals: photo-electric glottograph (PEG), oral airflow, and audio. Further recordings of data sets 1,2, and 3 were made, where subglottal pressure (strictly speaking, oesophageal pressure) was substituted for PEG. For details on equipment used see /5/ and references therein. The first data set (Icelandic, 66 tokens) contained the three possible bilabial stops in VCV, as exemplified in the words [laʰpa], [la:pa] and [ska:pa]. Each of these was inserted into carrier frames so that the word in alternate sentences did and did not receive the main sentence stress: "Hann sagði -- wið mig." and "Hann sagði ekki -- wið mig." The three further sets involved Irish utterances containing the voiceless dental stop in VCV and #CV. For set two (40 tokens), the word [baʰtʰa] was simply inserted into the frame: "Dúirt sé -- liom." which was repeated so that in alternate repetitions the word received either normal sentence stress or emphatic stress. In set three (48 tokens), a further frame was added: "Dúirt sé -- beag liom." As sentence stress in this last frame falls on beag, the word [baʰtʰa] is in the relatively unstressed (prenuclear) position. The intention here was to elicit three stress levels; emphatic, normal and a (relative) lack of stress. In the fourth set (24 tokens), the word [tʰa:] was repeated with alternating emphatic and normal sentence stress in the frame: "Tá, adúirt sé."

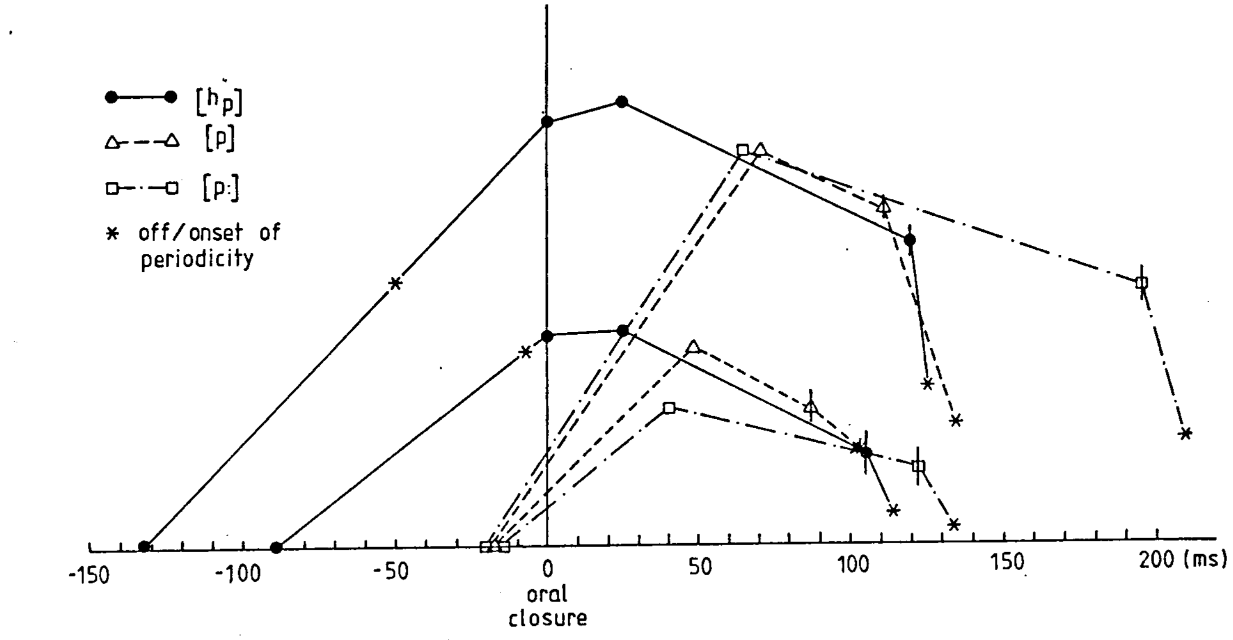

Fig 1. Averaged PEG traces for Icelandic medial stops for stressed (higher peaks) and unstressed (lower peaks) environments.

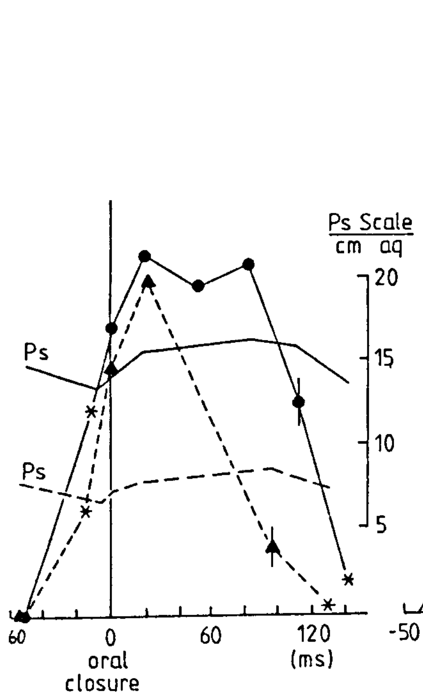

Fig 2. Averaged PEG for -ʰtʰ- in Irish, for emphatic v. normal stress. Ps values referred to in text.

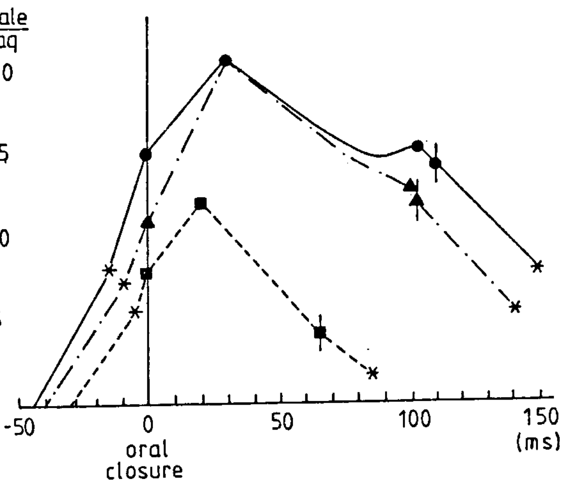

Fig 3. Averaged PEG for Irish -ʰtʰ- in emphatic, normal and normal stress environments.

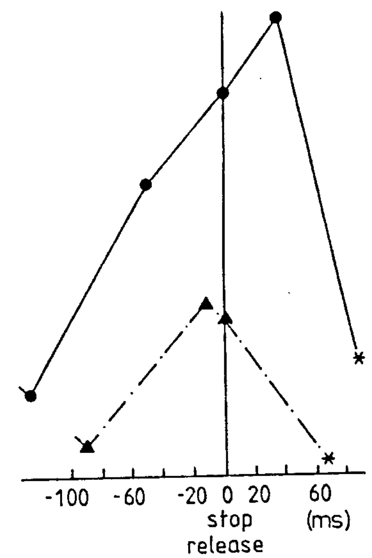

Fig 4. Averaged PEG for Irish #tʰ- in emphatic v. normal stress.

The comparative amplitude of the PEG waveform was measured at the following points: during the preceding vowel (except for stops in #CV) a point taken as the baseline from which the PEG amplitude was measured; at stop closure (for the preaspirated stops); at peak glottal opening; at stop release, and at voice onset. The time intervals between these points were measured, as were the durations of post- and preaspiration. Also measured for the latter, was the time taken to effect devoicing from the start of vocal fold abduction.

RESULTS

Figures 1,2,3 and 4 display the averaged PEG results. In Fig 1, where stressed and unstressed tokens of the three types of voiceless stops in Icelandic are compared, the amplitude of the peak glottal opening (PGO) would appear to be considerably greater for the stressed. And although the duration of preaspiration is a good deal shorter in unstressed than in stressed [ʰp], this is not reflected in the timing of PGO. Postaspiration values do not differ in the stressed and unstressed conditions, a fact which might appear surprising at first glance. Values are uniformly very short (ranging from 10ms to 35ms, and so, postaspiration is likely to be barely if at all perceptible). Also note that offset and onset of periodicity would seem to occur at greater amplitude of glottal opening for the stressed cases. In Fig 3, where Irish unstressed tokens are compared with normal and emphatic, PGO amplitude seems again much lower in the unstressed.

The difference between emphatically and normally stressed tokens (Irish) is not uniformly reflected by a single change in laryngeal behaviour. In #CV, there is not only a large differencein the amplitude of PGO but also in its timing. The peak occurs either before or after stop release, depending respectively on whether the stop occurs in the normally or emphatically stressed syllable. The difference in postaspiration duration is likely to be due to both timing and amplitude effects. The medial stops are rather different (note that these have both some pre- and postaspiration). When only emphatic and normal were contrasted (data set 2, Fig 2) there would appear to be little difference in the amplitude of glottal opening, but a striking difference in the glottal gesture itself. Emphatic tokens are characterized by a second glottal opening peak, the first of which corresponds in timing to that of the normally stressed tokens. The durations of pre- and postaspiration are virtually the same for both stress conditions. Note that the time taken to effect voicelessness is about the same also, but that the amplitude of glottal opening at which voice offsets is considerably greater with emphatic stress; the rate of glottal opening may therefore be quicker. As a consequence of the second opening peak, the amplitude of glottal opening at stop release also appears to be greater for these stops. In Fig 3 (data set 3, containing additionally unstressed tokens) the difference between the emphatic and normal is less striking, although there is again clear evidence of double peaking in the emphatic. It is likely that speakers were differentiating less consistantly between the emphatically and normally stressed tokens, when this was not the only thing they had to attend to.

INTERPRETATION

As the photoelectric glottograph is not calibratable there remains the possibility that results could be due to some artefact, e.g., a shift of the catheder with the light sensor in the pharynx might affect the amount of light picked up, and hence, the amplitude of the waveform. Long term catheder shifting could not however explain the systematic differences noted, given that the stress-varying utterances were read as alternating sentences and that recordings were short in any case. However, if there were to be some stress-related articulatory difference, e.g., laryngeal movement in the vertical dimension, it could conceivably yield the differences in PEG amplitude (though hardly the double peaks of the Irish emphatic stops). There is nevertheless some corroborative evidence for our interpretation. Andersen /7/ reports similar variation in PGO amplitude for voiceless stops spoken at different loudness levels. These findings were based on PEG data, but were to some extent backed up by fibre-optic and EMG data. More recently, on the basis of inverse filtered data, Fant /8/ and Gobl /9/ have reported for [h], a wider glottal opening in the stressed than in the unstressed syllable.

Even if our interpretation is correct, one must further ask whether these differences are actively controlled, or might simply be a passive consequence of some other stress correlate such as increased stop duration or higher subglottal pressure (Ps).

The Icelandic data shows that the degree of glottal opening can not be just a function of duration. Closure duration for the unstressed geminate is as long as for the stressed single stops, but PGO amplitude is much lower. Furthermore, within either stress condition, PGO is not geater for the geminate than for the single stops. Therefore, it can't be the case that the vocal folds will simply deflect more widely given extra time in which to do so.

A second possibility is that the vocal folds are blown wider apart as a consequence of increased Ps (a likely correlate of stress /10/,though not necessarily to be expected in every language: see,for example Welsh /11/). Ps values in our data were higher with increased stress. Peak Ps values in the vowel preceding the stops were on average 6cm Aq higher in stressed than in unstressed tokens for Icelandic (data set 1), and about 4cm Aq in Irish (data set 3). Emphatic tokens were 8 cm Aq higher again. Averaged Ps values for data set 2, are indicated in Fig 2. However, it is very unlikely that the differences in glottal opening degree are passive consequence of the Ps level, judging from an experiment by Löfqvist et al./3/. Sudden pressure changes, induced during PEG recordings by unexpected jabs in the subject's chest, made very little difference to the PEG

trace. Furthermore, the double peak glottal opening of the Irish emphatic stops can not be attributed to pressure variation. Although Ps is higher for the emphatic tokens (see Fig 2), there is no additional sudden increase during stop closure which could account for the second peak.

On balance therefore, it would seem that the observed differences in laryngeal behaviour are under active control. It may be the case that increased activity of the laryngeal musculature is a direct correlate of stress (typically perhaps an increase in abductive activity, but with an additional possibility of initiating a second abductive gesture when necessary). Thus, one could adopt the viewpoint that stress is potentially manifest by increased muscular activity at every level of production; the respiratory /10/, the laryngeal, and frequently, at the level of supralaryngeal articulation /12/.

These differences in laryngeal behaviour might not however be best regarded as correlates of stress as much as evidence of more general laryngeal strategies to ensure maximally equivalent output under differing aerodynamic conditions. To produce a voiceless segment there are potentially (depending on context) two crucial targets:

1) Sufficient glottal opening to ensure voicelessness. (Note: this excludes from present consideration glottalised stops, as occur in certain dialects of English.) The transition from voice -> voicelessness is not instantaneous (Westbury /13/ describes a "voice tail" of up to 40ms for voiceless stops in American English) and it can be very slow indeed when the vocal tract is not occluded, as for preaspirated stops /5/. If a greater degree of glottal abduction were not used at higher Ps levels, attainment of voicelessness might be delayed or prevented. Note that in the data presented, the higher the Ps, the greater appears to be the amplitude at which voice offsets.

2) The second target is the resumption of vocal fold vibration at the appropriate point in time subsequent to closure release. At voice onset, the initiation of vocal fold vibration results from glottal adduction and the Bernoulli effect; at a given stage of glottal narrowing, the vocal folds get sucked together. The point at which this happens depends on two factors working in an inverse relationship: the air flow rate through the glottis, and the degree of glottal narrowing. This may explain why, for the data in Figs 1 and 2, the actual duration of postaspiration is the same across different stress conditions, even though glottal aperture at stop release would seem to be quite different. Glottal closure in the higher stress tokens may simply be "stealing a ride", as it were, on the higer airflow and Bernoulli effect. Thus, at higer Ps and airflow rates, wider glottal opening may be not just tolerated, but actually necessary if VOT is to remain constant. At lower stress levels, too much glottal opening may be counter indicated as it would lead to unwanted aspiration.

Emphatic stress for the Irish medial stops may represent a particularly demanding articulation, given that they have both some pre- and postaspiration. Peak glottal opening occurs early during stop closure, as with preaspirating stops generally. If the ballistic opening/closing gesture were to proceed uninterrupted, the degree of glottal opening at stop release might not be sufficient to ensure the appropriate duration of postaspiration; hence the double opening. In initial postpausal position the situation is rather different, as only one of the above mentioned targets is relevant: appropriate voice onset. This may leave more freedom to use the additional strategy of delaying the peak glottal opening to prevent overshort aspiration at higher respiratory levels.

CONCLUSIONS

To conclude therefore, I would suggest that the degree of glottal opening is varied as a means of controlling voice onset and offset times across differing aerodynamic conditions. When necessary or possible, the additional strategies of a second opening peak or of a change in peak timing can also be brought into play. In other words, the larynx uses more than one control parameter to ensure that the crucial targets of voice offset and voice onset are maintained.

The fine interplay between laryngeal behavior and aerodynamic conditions suggests that there may be active monitoring of Ps at the laryngeal level.

REFERENCES

/1/ Kim, C., A theory of aspiration. Phonetica,21,107-116, 1970.
/2/ Catford, J.C., Fundamental Problems in Phonetics Edinburgh University Press, 1977.
/3/ Löfqvist, A., Baer, T., and Yoshioka, H., Scaling of glottal opening. Phonetica,38, 1981.
/4/ Löfqvist, A., Interarticulator programming in stop production. J. of Phonetics,8,475-490, 1980.
/5/ Ni Chasaide, A., Preaspiration in phonological stop contrasts. Unpublished Ph. D. thesis, University College of North Wales, Bangor, 1985.
/6/ Keating, P., Linker, W., and Huffman, M., Patterns in allophone distinction for voiced and voiceless stops, J. of Phonetics,11,277-290, 1983.
/7/ Andersen, P., The effect of increased speaking rate and (intended) loudness on glottal behaviour in stop consonant production as exemplified by Danish p. Ann. Rep. of the Inst. of Phonetics, Univ. of Copenhagen,15,103-146, 1981.
/8/ Fant, G., Interactive phenomena in speech production. Forthcoming in Proc. 11th Int. Cong. of Phonetic Sciences, Tallinn, 1987.
/9/ Gobl, C., Voice source dynamics in connected speech . Forthcoming in STL-QPSR, 1987.
/10/ Ladefoged, P., Three Areas of Experimental Phonetics. Oxford University Press, 1967.
/11/ Williams, B., Some problems in the description of stress in modern Welsh. Cambridge Papers in Phonetics and Experimental Linguistics,1, 1982.
/12/ Tuller, B., Harris, K.S., and Scott Kelso, J.A., Stress and rate: differential transformations of articulation. J.Ac.Soc.Am.,71,1534-1543,1982.
/13/ Westbury, J.F., Aspects of the Temporal Control of Voicing in Consonant Clusters in English. Texas Linguistic Forum, Texas, 1979.