SYMPOSIUM 3: Models of the Larynx

Chairman: *C. Scully, Leeds, United Kingdom*

Panel members: *G. Fant, I.R. Titze, M. Hirano, F. MacCurtain*

Discussant: *J. Sundberg, K. Stevens*

1. Introduction

Modelling of the larynx can be directed at different goals. One is the provision of better synthetic voices with increased naturalness. The emphasis here has been on the representation of one speaker of one language. A more distant - because it is more ambitious - goal, is the deeper understanding of the regularities and constancies of real speech for many different languages and for different speaker types. Improved quantitative phonetic descriptions should result; and a better understanding of the meshing together of common phonetic properties and individual characteristics. A third major aim is the provision of methods for the diagnosis of pathological states of the vocal folds.

A good model is a simplified representation of a process or system. In the case of speech production these are undoubtedly extremely complicated (see Titze, 1983). The main factors need to be identified, but until we have a better understanding of the fundamentals, our models are likely to contain large numbers of parameters. The modellers need as much quantitative information on natural human speech as their colleagues are able to provide. In return, the modellers may be able to make predictions beyond what can be measured in real speech. It is helpful to identify the larynx as a quasi-independent component of the speech producing system. Its complexity and importance in speech are indicated by the number of scholars investigating larynx activity as seen, for example, in the papers given at the Tenth International Congress of Phonetic Sciences and by the wide variety of techniques employed. Since the symposium on the larynx at the Eighth International Congress of Phonetic Sciences (Fant and Scully, 1977), several conferences and books have been partly or entirely devoted to the larynx in speech and singing (Fink, 1975; Carré, Descout and Wajskop, 1977; Fink and Demarest, 1978; Boë, Descout and Guérin, 1979; Lass, 1979, 1981; Lawrence and Weinberg, 1980; Stevens and Hirano, 1981; Abbs et al., 1983). Modelling is one path towards a greater understanding of the larynx, but it needs to be considered as several systems: (1) neural control mechanisms; (2) anatomical structures, tissue properties and muscle mechanics; (3) articulation; (4) aerodynamics; (5) acoustic sources; (6) acoustic filters. The actions are directed towards the achievement of auditory goals. Not surprisingly, therefore, in the symposium it was necessary to discuss articulation (whether subglottal, laryngeal or supraglottal) with reference to acoustic attributes such as fundamental frequency, voice source characteristics or frequency spectra for different vowel qualities. The larynx does not operate in isolation. To different degrees, in each of the systems listed above, links between the larynx and both subglottal and supraglottal regions need to be considered. Much of the discussion in the symposium concerned the effects of supraglottal changes on the voice source; and the changes in (supraglottal) vocal tract shape associated with different voice qualities the primary control of which is ascribed to the larynx.

In this report, the reviews of the four Panelists are summarised first (see Cohen and Van den Broecke, 1983). A report on the symposium follows. Statements are credited to individual scholars, but any misrepresentation is the responsibility of the Chairman. An attempt is made to relate discussion in the symposium to some current research issues.

2. Panelists' reviews

Hirano reviewed recent advances in research on the structure, mechanical properties and adjustments of the vocal folds.

From a mechanical point of view, the five layers of the vocal folds can be reclassified as: the cover (epithelium and superficial layer of the lamina propria); the transition (intermediate and deep layers of the lamina propria); and the body (vocalis muscle). The layer structure varies along the length of the vocal fold, with gradually increasing stiffness towards the ends. The membranous vocal fold is most pliant at the midportion because of its structure as well as its location. Vocal fold structures change with age: newborns have no vocal ligament; maturation of the layer structure appears to be completed around the end of adolescence; aging changes the layer structures, somewhat differently in women and men.

Vocal fold tissues exhibit non-linear stress-strain curves. Stiffness increases with elongation and is greater in the longitudinal direction (along the length of the vocal fold) than in the transverse direction; saturation effects are found in the longitudinal direction. Other mechanical properties have been measured also. Electrical stimulation of intrinsic muscles demonstrates their effects on the position, shape and structure of the vocal fold (Hirano, 1975).

Vocal fold adjustments in speech have been extensively investigated by Hirose and his co-workers (see Hirose and Sawashima, pp. 137-154 in Stevens and Hirano, 1981). A summary of their findings follows: (1) CT primarily controls F_0. Occasionally, CT contraction in segmental glottal adjustment for enhancing voicelessness is suggested. (2) PCA is active for voiceless consonants. The PCA activity is positively related to the degree of glottal opening. For voiced phonemes, PCA is inactive. (3) IA presents a pattern reciprocal to PCA. (4) The activity level of VOC and LCA is influenced by context. Generally speaking, their activity is decreased for consonants. The suppression of VOC and LCA is occasionally more significant for voiceless consonants than for voiced consonants. The pattern of the activity is different between the two muscles in some specific cases.

Titze reviewed the difficulties in computer stimulation of biomechanical processes of speech production, the progress being made to overcome them, predictions made by modelling, and future expectations.

Although studies on speech production can borrow from research into other biomechanical processes, phonation presents unique problems associated with its rapid tissue and air movements, well into the kHz range, where low frequency measurements of viscoelastic properties cannot be trusted. The structures must be considered as three-dimensional, because their dimensions are comparable to wavelengths. This necessitates a distributed-parameter (continuum) approach and tbe inclusion of displacements transverse to as well as along the tissue fibre direction. Computational problems arise as glottal area approaches zero. Both finite-difference and the current finite-element approaches to representing the vocal folds have computational difficulties, but these should be eased by developments in computing. The major obstacle remains the refinement of theories of pressure and flow development in the glottis, tissue deformation under these pressures, and the effects of laryngeal muscle contractions on the viscoelastic properties.

Successes include a clear demonstration of the flow-induced self-oscillatory nature of vocal fold vibration. Asymmetry in the net aerodynamic driving pressure acting outwards on the medial surface of the vocal folds with respect to opening and closing keeps the vocal folds oscillating. This asymmetry derives from the resistance and inertance of the air in the glottis and is enhanced by an inertive vocal tract load, while it is reduced by a capacitive vocal tract load (Ishizaka and Flanagan, 1972). The most efficient way of maintaining asymmetry is through vertical phase differences, as in two-mass or multiple-mass models.

Models have quantified the effects of vocal fold dimensions and configurations upon glottal area shape, vocal fold contact area, aerodynamic variables and the properties of the acoustic output. They have begun to predict the biomechanical significance of the layered structure of the vocal folds and of tissue anisotropy (Titze and Talkin, 1979). As more data become available modelling will help to explain the effects of variations in size, shape, viscoelastic properties, muscular contraction, lesions, stress relaxation and unsteady flow conditions.

Fant considered recent significant advances in the source-filter concept of voice production.

In current models source and filter are assumed to be short-time invariant. A firmer theoretical basis is being constructed in which acoustic and mechanical interactions between filter and source are included. The major complication for modelling is that, in the glottal open state (a phase of acoustical energy charging), the sub- and supraglottal parts of the system are acoustically coupled, whereas, with the glottis closed, (an energy discharge phase), they each execute approximately free and separate oscillations.

In the most complete speech production model (Ishizaka and Flanagan, 1972), lung air pressure provides the power; the system is self-oscillating and self-adjusting, without a source in the linear network sense. The shape of the glottal area function Ag(t) may be assumed instead, and is combined with lung air pressure to derive glottal volume velocity Ug(t), defined as the voice source (Ananthapadmanabha and Fant, 1982). Alternatively, a 'short circuit' or 'no load' source Us(t) can be derived. These two voice sources Ug(t) and Us(t) differ in two main respects: Ug(t) is skewed to the right, with a slower rising branch and a faster falling branch than Us(t); oscillatory ripples, usually close to F_1 in frequency, are superimposed on Ug(t). Us(t) provides a simple source but a complicated, time-variable filter function while, with Ug(t) as source, the filter function is simple, being linear and only slowly time-variable, but Ug(t) depends on the entire sub-and supraglottal configuration and has to be derived by complex calculations. In a third, approximate, model (Fant, 1979) a smoothed version of Ug(t) applies a constant current source to the vocal tract in parallel with the time-varying glottal plus subglottal impedance. Two stage synthesis has been proposed (Mryati et al., 1976). Here an Ag(t) voltage, with an F_1 and F_2 impedance load, gives a current function which is used as the source in a formant synthesizer. The perceptual importance of the various approximations has not yet been fully investigated.

Parameterisation of the glottal flow or glottal area functions was discussed, and extensions to the Fant (1979) model for better matches with real speech data. Important characteristics of the Ug(t) function include: the area under the curve, related to low frequency acoustic amplitude; a low frequency spectral maximum Fg which is prominent at low voice effort; and the amplitude of the negative spike of the flow derivative at closure, which controls formant amplitudes.

Problems in the use of different kinds of inverse filtering to obtain reliable estimates of glottal flow were reviewed. More research is needed to improve the modelling of women and children's voices, speaker specific factors and speech source dynamics, which are probably carried more by vocal fold actions than by lung pressure variations. Rules for voice sources need to be based on an understanding of production events and underlying physiological parameters.

MacCurtain presented evidence of the pharyngeal gestures and associated larynx adjustments which occur in the formation of different voice qualities.

A 5-parameter descriptive framework is proposed, based on data from seven languages. Phonetic theory as currently formulated fails to account for some voice quality contrasts. The techniques used, xeroradiography (combined with acoustic and electro-laryngographic recordings) show soft tissue in extremely fine detail, even indicating individual muscle groups.

English informants produced different voice qualities for an [i] vowel. Out of a descriptive framework of 12 anatomical points for the larynx and pharynx, 5 points altered significantly between rest position (expiration) and phonation. Large inter-speaker variations in muscle patterning were found for a common auditory target of, for example, 'harsh' voice; it may be misleading to attribute one set of componential elements (Laver, 1980) to one voice quality.

The 5-parameter framework was applied also to minimal pairs of phonological oppositions in which contrastive voice qualities are required. As an example, results for a speaker of Dinka were cited: in moving from 'brassy' voice quality, associated with high tone, to 'hollow' voice quality, associated with low tone, the speaker greatly widened both the laryngopharynx and the lower oropharynx; the thyroid cartilage was lowered to 'resting' height, but the dimensions of the laryngeal airway (from the anterior end of the vocal folds to the supero-posterior cricoid tip) remained unchanged, in spite of the required frequency contrast.

The parameters are being applied to the evaluation of voice disorders (Berry et al., 1981) and to a study of singing.

3. Neurological control

The modelling of the larynx as part of a neural system has been discussed by Muller, Abbs and Kennedy (see pp. 209-227 in Stevens and Hirano, 1981).

Hirano: The role to sensory feedback and reflexes of the PNS has been over-emphasised; more attention should be given to central mechanisms, including the extra-pyramidal system.

Hirose: Part of the cerebellum and the basal ganglia participate in the preprogramming of movement before the motor command is realised at the level of the motor cortex. More significantly, another part of the cerebellum is checking the outflow of the motor command from the cortex to the periphery before the commands reach the peripheral organ itself. Peripheral feedback mechanisms are certainly important, particularly in the course of learning skilled movements, including speech. But after a certain period of learning, the CNS knows precisely what should be done according to the template stored in the brain and updating mechanisms provided by cerebello-cerebral interactions.

4. Biomechanical interactions and aerodynamic forces in the control of F_0 and the voice source

Stevens: It is important to make a distinction between effects seen across languages and those seen across individuals. There are relatively few possibilities for creating contrasts which are phonologically relevant; these should conform to quantal principles. On the other hand, inter-speaker differences are likely to be characterised by fine, complex gradations. Three or four basic parameters are needed to characterise the controlled variables of F_0 and the voice source waveform with its associated spectral emphasis, ranging from 'breathy' to 'pressed' phonation. Because the muscles which alter F_0 and those which control the waveform interact, these two acoustic aspects of speech interact also. A speaker may change the vocal tract shape in order to enhance contrasts created primarily in the larynx; larynx actions, on the other hand may be able to create auditory effects which are assumed to be associated with the vocal tract: an example might be the auditory property of nasalisation created by abduction of the vocal folds.

Fujisaki: His functional model of F_0 control by the cricothyroid and vocalis muscles represents F_0 changes as the response of a second-order linear system to step functions. With suitable values for the damping factor (different for speech and for different kinds of pitch transition in singing) good matches to F_0 contours are generated. With the assumption that the vocal folds vibrate like simple elastic membranes, so that frequency is proportional to the square root of tension, and using the known non-linear relationship between tension and elongation for skeletal muscles, it has been shown that, to a first order of approximation, the logarithm of F_0 is proportional to the elongation of the vocalis muscle. Data from real speech (Honda et al., 1980) support the analysis.

Another implication of the non-linear stress-strain curve of muscle tissue was discussed.

Titze: It should not be assumed that an increase in any aerodynamic variable operating at the larynx will necessarily raise pitch. Indeed, an increase in volume flow rate accompanies decreased longitudinal stiffness of the vocal folds and under these circumstances increased airflow accompanies a decrease in F_0. The empirically observed correlation between increased transglottal pressure drop and higher F_0 is not simple to explain. The effect is indirect. The higher aerodynamic forces on the vocal folds due to increased pressure drop result in increased amplitude of vocal fold vibration. This would not increase F_0 if the vocal folds constituted a linear oscillator. A pendulum, or a child on a swing for example, has the same frequency of oscillation regardless of amplitude (over a limited range at least). But, in the case of vocal fold tissue, increased deformation is associated with greater mechanical stiffness. As the amplitude of vibration of the vocal folds increases the effective stiffness of the vocal folds increases also; for this reason F_0 rises as transglottal pressure drop increases. Control of F_0 is made more complicated by relaxation effects. After vocal fold stiffness has been suddenly increased, it declines gradually.

The declination of F_0 contours in speech can probably be accounted for by an observed gradual decrease in subglottal pressure. It is not necessarily possible to ascribe a fixed change in F_0 to a fixed change in pressure drop. $\delta F_0/\delta \Delta P$ may well vary depending on the biomechanical state of the vocal folds. This is an example of the physiological variables interacting in rather complex wavs to determine acoustic results (Collier, see p. 440 in Cohen and Van den Broecke, 1983, with comments by Titze).

Ohala: Non species-specific biological explanations may be offered concerning pitch control. Some survival value may attach to the possession of rapidly operating mechanisms for both F_0 raising and also F_0 lowering, to exaggerate the animal's body size, for instance.

Löfavist: Caution is needed in assigning functions to particular muscles. For example three kinds of articulatory actions may be used to make voicing cease. The vocal folds can be widely abducted, or strongly adducted, or abruptly stiffened. Interactions between vocalis and posterior cricoarythenoid muscles may be just as important as the reciprocal activity observed for interarytenoid and posterior cricoarytenoid muscles in the control of the acoustic feature ± voice.

Published research for natural speech (see, for example, Rothenberg, 1973; Gauffin and Sundberg, 1980) suggests that, within a 'plateau' region for phonation, the form of the total volume flow rate through the glottis varies. At the edge of the plateau, voicing dies away, for a variety of inimical articulatory and aerodynamic reasons. Covariation in the a.c. (acoustic) and d.c. components of air flow through the glottis implies that voice and aspiration sources covary. Published data for natural speech have been incorporated in a functional model for the voice source, with three controlling variables: pressure drop, average glottal area and a vocal fold stiffness factor (Scully and Allwood, 1983).

5. Larynx - vocal tract physiological interactions

With models, individual factors may be identified and manipulated as they cannot be in real speech. For example, acoustic coupling and anatomical links between larynx and vocal tract can be dissociated. Modelling of acoustic coupling alone gives variations of fundamental frequency with vowel type in the opposite sense to that of real speech so physiological links presumably give overriding effects (Guérin, Degyrse and Boë, pp. 263-277 in Carré et al., 1977).

The hyoid bone has been observed in continuous speech (see, for example, Botherel, 1975) and for isolated vowels (see, for example Rossi and Autesserre, 1981). The forces acting on the hyoid bone were considered at the symposium. Eight pairs of muscles are attached to the hyoid bone, and five pairs of ligaments or membranes: an astonishing number for this small bone. Through the geniohyoid, mylohyoid and digastricus muscles, the larynx is likely to be affected by jaw movements; and the larynx is connected to the tongue through the hyoglossus muscle pair. Even the soft palate is connected to the larynx, via the palatopharyngeus muscles, which attach to the laminae of the thyroid cartilage.

Whereas it is generally accepted that longitudinal stretching of the vocal folds increases the stiffness and raises F_0, the effects of vertical stretching are not so clear. Interestingly, the association between high F_0 and advanced tongue root with wide pharynx discussed by Honda, below, is in contrast to the low F_0-wide pharynx association found by MacCurtain for Dinka (MacCurtain, 1983).

Honda: Observed geniohyoid muscle activity associated with pitch raising and correlations between high F_0 and horizontal position of the hyoid bone suggest additional mechanisms for pitch raising besides intrinsic larynx muscle activity. The intrinsic pitch of vowels can be explained by the same mechanism of the laryngeal external frame function. The hyoid bone is pulled forward and downward by the geniohyoid and genioglossus (posterior fibres) muscle forces, the latter being the more important in reducing the anterior-posterior length of the tongue in high vowels. As a result of the hyoid bone shift, the thyroid cartilage is rotated forward about the cricothyroid joint, thus stretching the vocal folds and raising F_0. Tongue movements can influence vocal fold state indirectly (via the epiglottis) also. When the epiglottis is pulled upright as the tongue root moves forward, the aryepiglottic fold is stretched. Increased tension here may raise F_0 through increased stabilisation of the arytenoid cartilages. But the (vertically) wide laryngeal ventricle seen for high vowels may perhaps not stiffen the vocal folds for two reasons: first, the vocal folds are not tightly coupled to the false vocal folds; secondly, the vocal folds are more compliant in the vertical direction than in the horizontal direction.

MacCurtain: Because of the large number of muscle attachments, there is much redundancy in the positional control of the hyoid. For some speakers the cornu shifts horizontally. Different angles of tilt of the hyoid are seen for different voice qualities.

Liberman: Emg activities for geniohyoid and other extrinsic laryngeal muscles show a positive correlation with F_0, especially in the upper quartile of the range.

Fischer-Jørgensen: Data published by Alfonso show that F_0 variations for different vowels correlate better with the position of the jaw than with that of the tongue.

6. Aerodynamics

If the slowly changing (d.c.) components of air pressure and volume flow rate of air are considered separately from the acoustic (d.c.) components, these former are variables in an irreducible aerodynamic system comprising the whole respiratory tract. Since aerodynamics links actions to sounds, a single articulatory change results in multiple acoustic changes. With this approach, the average flow rate through the glottis reflects the articulation at the larynx and elsewhere. A less simplified approach would need to incorporate the effects of aerodynamic forces on glottal area directly. Average glottal area is not necessarily a good representation of vocal fold abduction-adduction. Modelling shows that where longitudinal stiffness of the vocal folds is increased at constant vocal fold setting and subglottal pressure, average airflow decrease; so that average glottal area derived from the aerodynamic orifice equation would decrease (Titze and Talkin, 1979). This kind of effect needs to be taken into account if a model maps from average airflow onto the airflow waveform for the voice source. A sophisticated model of respiratory control needs to be developed. Should average subglottal pressure and glottal area be considered as two independent controlling factors or should average subglottal pressure, on the contrary, be treated as a dependent variable, to be derived from average glottal area combined with a lung volume control model (in which lung volume decrement or pressure in the alveoli of the lungs or nett expiratory force could be the independent variable)? Models need to be able to explain why subglottal pressure falls during an utterance.

7. The voice source and acoustic coupling

The quasi-periodic waveform of volume flow rate at the glottis can be recovered for vowels in natural speech by inverse filtering. The complex shapes and structures of the vocal folds have been represented in a number of different ways in models.

Flanagan described the two-mass model of the vocal folds. When each vocal fold was represented as a single mass the system was self-oscillating but was unduly sensitive to vocal tract loading. Synthetic speech, obtained by coupling the two-mass model to a transmission line model of the vocal tract, was demonstrated. This modelling was for speech synthesis, without the intention of duplicating fine physiological detail. With the vocal folds set to a suitable operating glottal area and with subglottal pressure applied, dominant acoustic effects of natural speech were produced (see Ishizaka and Flanagan, 1972).

As articulatory synthesis of speech this research has remained unsurpassed. Other approaches to modelling have sought insight into the wide range of normal and abnormal vibration modes found in real speech. The vocal folds have been represented as many masses and springs, and as continuous viscoelastic media (see Titze, 1983). The vocal folds have been represented as a pair of beams (Perrier, 1982). It is hoped that this research may resolve some of the problems of commands for two-mass models, while providing a closer match to the complexities of the natural physiology (Guérin, pers. comm., 1982). The forces acting upon the parallel structures of the body and the cover of the vocal folds have been analysed by Fujimura (see pp. 271-288 in Stevens and Hirano, 1981).

Fant considered ripple effects seen in voice waveforms and skewing effects. The flow rate waveform is skewed relative to the glottal area waveform, with sharpening of the flow cut-off. Cut-off steepness is modified by the inclusion of frictional losses in the analysis. Since skewing of the airflow waveform is associated with inertance of air above as well as in the glottis, the skewing effect is different for different vowels. The effects have been modelled by Rothenberg (see pp. 305-328 in Stevens and Hirano, 1981).

Fujimura: It is interesting to consider whether the different pharynx shapes found for contrasting voice qualities (see MacCurtain, 1983) may influence the acoustic output in several ways which could be separated by modelling. Besides the different filter shapes, the voice source could perhaps be affected in two ways: through the physiological links of the external frame function of the larynx and in a direct, acoustic, way.

Titze: Yes, acoustic coupling effects would be expected: a large cross-section area in the lower pharynx, just above the vocal folds, reduces the skewing effect, while a constricted pharynx enhances it.

Fant: In addition. differences of shape in the vestibule and lower pharynx may influence acoustic sources because of their effects upon the total pattern of airflow just above the glottis.

8. Singing versus speech

Sundberg: Trained singers are valuable subjects for research into larynx behaviour. The normal and natural way for speech and untrained singing is with a physiological emphasis. Covariations are allowed. For example, loudness and pitch generally rising and falling together gives a 'natural' effect. Singing, on the other hand, is oriented towards the side of perception, with more stringently controlled auditory goals. Singing demonstrates phenomena not found in speech, with more demands made on physiological control. For example, gastric and oesophageal pressure measurements show that the diaphragm muscle, found to be inactive in previous studies on speech, is active during professional singing. Muscle forces controlling movement act faster in singing, giving more rapid changes in subglottal pressure than those found in speech (von Euler et al., 1983). The automatic results of physiological and acoustic coupling between the larynx and other structures may be unacceptable in professional singing. Singers often experience vowel-dependent difficulties in singing on a given pitch; these and other voice source effects which are revealed in singing need to be investigated further. In the course of their training singers probably learn to break up the habitual relationships between the voice source on the one hand and pitch and loudness on the other. Control of the voice source spectrum independently of loudness and pitch may not be required in speech but probably is important for opera singers.

9. Inter-speaker variations

Models ought to conform to physical constraints yet be capable of simulating different speaker types, whether the inter-speaker differences arise from anatomical and physiological differences or from the options selected. Apart from contrasts between speech and professional singing, two other aspects were discussed.

10. Voice source waveforms for women speakers

Karlsson (Stockholm) showed voice waveforms obtained by inverse filtering, for four Swedish women speakers using a normal F_0 range with normal and loud voice. There is a wide range of waveshape parameter values across speakers; greater than any inter-sex differences.

Price: By contrast with Karlsson's data, American women speakers gave voice waveshapes with a 40% shorter 'closed' fraction than those for men. The 'closing' time was three times as long for women, with a much more rounded 'corner'.

These differing results could possibly be ascribed to the techniques of inverse filtering used (from mouth flow rate with a Rothenberg mask for the Swedish data; from acoustic pressure for the American English). However, it may be that the differences are genuine cross-language ones, possibly reflecting different social attitudes and conditioning; or even simply inter-speaker differences. Much more emphasis on the analysis of women's speech is called for.

11. Pathological conditions

Hirano gave an account of pathological conditions of the vocal folds, including paralysis, polyps and carcinoma. Each type of pathology creates its own kind of abnormal mechanical state for the pair of vocal folds with implications for their modes of vibration.

As acoustic and other accessible measures are related to specific pathologies, empirically based non-invasive assessment techniques will be developed. Laryngograph waveforms differ for a small single area of vocal fold contact and for a small total area associated with numerous very small contact areas for rough surfaces (Fourcin, pers. comm., 1983). Acoustic analyses of speech outputs are being matched to histological examinations, so that changes in vocal fold cell structure may be inferred from acoustic data (MacKenzie et al., 1983). Within the descriptive framework of MacCurtain, some of the vocal tract parameters are likely to prove particularly relevant for the diagnosis of disorders of the larynx.

Hirano raised the question of the effect on source-filter acoustic interaction when there is no closed phase, as in some laryngeal pathologies.

Fant: The increase in acoustic interaction between the supraglottal and subglottal systems affects resonance frequencies and bandwidths. The problems of obtaining true glottal flow are increased, since frequencies and bandwidths must represent closed glottis conditions if inverse filtering is to be an exact inverse of the transfer from glottal flow to the speech wave.

12. Quantitative phonetic descriptions

In the time available, questions of appropriate phonetic features for the description of larynx behaviour in models and natural speech could not be considered in detail.

Ladefoged reviewed the variety of approaches needed for larynx descriptions, from both physiological and acoustic bases. The vocal folds can be investigated as vibrating structures or as the voice waveshape, related to phonetic properties of the output speech. Material from a wide range of languages needs to be analysed. In some languages the phonetic properties of breathy voice and high F_0, creaky voice and low F_0 hang together; in other languages they 'hang apart' and are not found to be correlated.

13. Conclusions

The symposium offered a rapid scan and overview of a number of areas of research relevant for improved larynx modelling. Some of these areas merit, and are already the subjects of, more specialised discussion meetings.

Acoustic and physiological data can be correlated with auditory phonetic descriptions, preferably for many different languages. The growing interest in the analysis and synthesis of women's voices should be applauded and encouraged. Physical modelling of abnormal conditions of the vocal folds should assist acoustic differentiation of pathological states of the larynx. The valuable new tool of inverse filtering (whether from volume flow rate or from acoustic pressure) will provide voice source data. These real speech waveforms can provide a means of assessment of outputs from vocal fold models and more natural source inputs for acoustic models of speech, whether for terminal-analog or for line-analog synthesis. Acoustically significant features of the glottal flow waveshape have been identified. These should be described quantitatively in future analyses, whatever additional parameters may be of interest to the experimenter. Future modelling can contribute to an increased understanding of turbulence noise source properties and of their interactions with the voice source. Acoustic modelling of interactions between the voice source and the vocal tract filter is well advanced and the main factors seem to have been identified; the contribution of the subglottal acoustic filter is becoming better understood. There is a clear need for models of anatomical and physiological interactions between the larynx and supraglottal structures, supported by physiological experiments on real speech in which acoustic effects such as F_0 changes are recorded. Fink and Demarest (1978) have provided a possible appraoch to this kind of modelling.

Unification of terminology must probably await increased understanding. Registers in speech, for example, may not correspond to those of singing (see Hollien, p. 487 in Cohen and Van den Broecke, 1983), but different voice qualities, phonation types and registers in speech need to be labelled auditorily and related to vibratory patterns for the vocal folds (see Stevens and Hirano, 1981, p. 168).

References

Abbs, J. et al. (1983). *Vocal Fold Physiology Conference*, Madison, Wisc. (in press).

Ananthapadmanabha, T.V. and Fant, G. (1982). Calculation of true glottal flow and its components. *Royal Inst. of Technol. Dept. of Speech Communication and Music Acoustics, Stockholm, QPSR 1/1982*, 1-30.

Berry, R.J., Epstein, R.L., Fourcin, A.J., Freeman, M., MacCurtain, F. and Noscoe, N. (1982). An objective analysis of voice disorder, parts 1 and 2, *Brit. J. of Disorders of Comm.* **17**, 67-85.

Boë, L.J., Descout, R. and Guérin, B., eds. (1979). *Larynx et Parole*. Proceedings of a GALF Seminar, Grenoble, February 8-9, 1979. Grenoble: Institut de Phonétique de Grenoble.

Bothorel, A. (1975). Positions et mouvements de l'os hyoïde dans la chaîne parlée, *Travaux de l'Institut de Phonétique de Strasbourg* **7**, 80-132.

Carré R., Descout, R. and Wajskop, M., eds. (1977). *Articulatory Modelling and Phonetics*. Proceedings of a Symposium, Grenoble, July 10-12, 1977. Brussels: GALF Groupe de la Communication Parlée.

Cohen, A. Van den Broecke, M., eds. (1983). *Abstracts of the Tenth Intl. Congr. of Phonetic Sciences*, Foris Publications, Dordrecht, Holland.

Euler, C. von, Lagercranz, H., Leanderson, R. and Sundberg, J. (1983). Diaphragmatic activity during singing. *Proc. of the Stockholm Music Acoustics Conf. July-Aug. 1983*, to appear.

Fant, G. (1979). Glottal source and excitation analysis. *Royal Inst. of Technol. Dept. of Speech Communication and Music Acoustics, Stockholm QPSR 1/1979*, 85-107.

Fant, G. (1983). The voice source - acoustic modelling. In: Cohen and Van den Broecke, 1983, 151-177.

Fant, G. and Scully C., eds. (1977). *The Larynx and Language. Proceedings of a Discussion Seminar at the 8th International Congress of Phonetic Sciences Leeds August 17-23, 1975 Phonetica* **34**, 4.

Fink, B.R. (1975). *The Human Larynx, A Functional Study*. New York: Raven Press.

Fink, B.R. and Demarest, R.J. (1978). *Laryngeal Biomechanics*. Cambridge, MA: Harvard University Press.

Gauffin, J. and Sundberg, J. (1980). Data on the glottal voice source behaviour in vowel production. *Royal Inst. of Technol. Dept. of Speech Communication and Music Acoustics, Stockholm, QPSR 2-3/1980*, 61-70.

Hirano, M. (1975). Phonosurgery - Basic and clinical investigations. *Otologia Fukoka* **21**, 239-440.

Hirano, M. (1983). Structure, mechanical properties and adjustments of the vocal fold. In: Cohen en Van den Broecke, 1983, 187-193.

Honda, K., Hibi S.R., Kiritani, S., Niimi, S. and Hirose, H. 980). Measurement of the laryngeal structures during phonation by use of a stereoendoscope. *Ann. Bull. RILP, Univ. of Tokyo* **14**, 73-78.

Ishizaka K. and Flanagan, J. (1972). Synthesis of voice sounds from a two-mass model of the vocal cords. *Bell Syst. Tech. J.* **51**, 1233-1268.

Lass, N.J., ed. (1979, 1981). *Vols. 2 & 5, Speech and Language: Advances in Basic Research and Practice*. New York: Academic Press.

Laver, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge: C.U.P.

Lawrence, V.L. and Weinberg, B., eds. (1980). *Transcripts of the Eighth Symposium Care of the Professional Voice, Parts I, II, and III*, June 11-15, 1979, The Juilliard School, NYC. New York: The Voice Foundation.

MacCurtain, F. (1983). Pharyngeal gestures in contrastive voice qualities. In: Cohen and Van den Broecke, 1983, 199-213.

Mackenzie, J., Laver, J. and Hiller, S.M. (1983). Structural pathologies of the vocal folds and phonation. *Edinburgh Univ. Dept. of Linguistic Work in Progress* **16**, 80-116.

Mryati, M. Guérin, B. and Boë, L.J. (1976). Etude de l'impédance d'entrée du conduit vocal. Couplage source-conduit vocal. *Acustica* **35**, 330-340.

Perrier, P. (1982). Etude d'un modèle continu des cordes vocales sous forme de deux poutres bi-articulées: premières simulations, unpublished Diss. Doct. Ingén., ENSERG l'Inst. Natl. Polytechn. de Grenoble.

Rossi, M. and Autesserre, D. (1981). Movements of the hyoid and the larynx and the intrinsic frequency of vowels. *J. of Phonetics* **9**, 233-249.

Rothenberg, M. (1973). A new inverse-filtering technique or deriving the glottal air flow waveform during voicing. *J. Acoust. Soc. Amer.* **53**, 1632-1645.

Scully, C. and Allwood, E. (1983). Simulation of singing with a composite model of speech production. *Proc. Stockholm Music Acoustics Congress 1983*, to appear.

Stevens, K.N. and Hirano, M., eds. (1981). *Vocal Fold Physiology.* Proceedings of a Conference, Kurume. January 15-19, 1980. Tokyo: University of Tokyo Press.

Titze, R. (1983). Approaches to computational modelling of laryngeal function: successes and prevailing difficulties. In: Cohen and van den Broecke, 1983, 179-186.

Titze, I.R. and Talkin, D.T. (1979). A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. *J. Acoust. Soc. Amer.* **66**, 60-74.