A SUGGESTED MEANS TO STUDY THE FORCES ON THE BASILAR MEMBRANE DURING THE AUDITION OF A SPEECH SOUND*

K.H. KRISHNAMURTHY

The ascertainable responses of the basilar membrane to the impinging sound are recognised to be too broad for the actual, known, auditory sharpness. In pitch for e.g., the travelling wave is too broad for the known acuity (Lawrence 1966). Though the latter is held to be the result for a contrast increase further up the acoustic pathway (Bekesy 1960), we may enquire if more minute mechanical responses take place at the membrane itself, and if so, how to assess them. As these cannot be observed directly, especially as they might be occuring *in vivo* for a specific sound, e.g., speech, the next best means is to study the micro-pressure pattern as given by the oscillogram of the concerned speech ensemble in more detail.

Though strictly speaking the oscillogram represents pressure changes on the tympanum, we can regard it as roughly representing the pressure pattern on the basilar membrane across the cochlear fluids itself (Figure 1), since it is known that the transmission in the middle ear is remarkably linear (Sweetman and Dallas 1969) and subject to very negligible distortion (Zalin 1961). This is a measure of expediency no

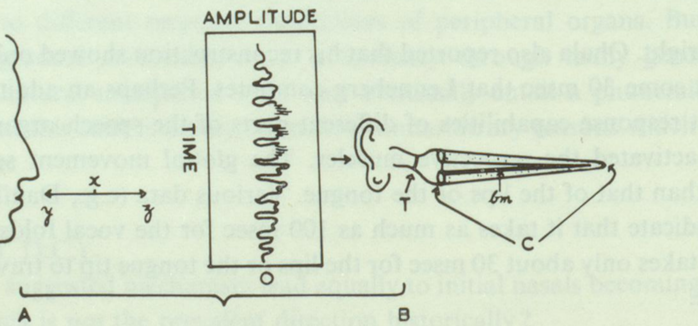

Fig. 1. The Net Process of Speech. A-B, the extent of the transmission medium between the speaker and the hearer, *x*, any particle in this medium which describes on the impact of a speech sound, a to and fro movement between *y* and *z*, whose extent and details constitute the pattern of 0, the oscillogram. T., tympanum, C. Cochlea and *bm*, the basilar membrane.

* Grateful thanks are due to Prof. C.R. Sankaran and Dr. W. Bethge for the gift of the oscillograms and the suggestions in the preparation of the manuscript.

doubt now; its experimental validation with a model is under separate study. Oscillogram is well utilised from Fletcher onwards to assess the physical characteristics of speech sound. What is new here is the stress on the fact that every feature of its track is directly related to a specific pressure change, so that the full analysis of its form can be taken to provide a minute account of the pressure forces available for *all* the mechanical reactions of the membrane.

1. INDIVIDUAL TRACKS

Oscillograms of the words 'val', 'pushkal', 'pali' (Figure 2) spoken by a single subject were studied with magnification where needed. Figure 3 is an idealised diagram of 'val'.

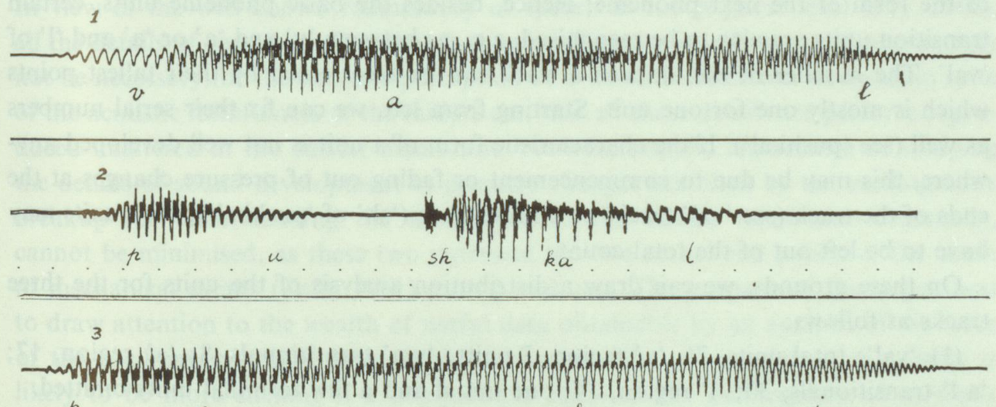

Fig. 2. Oscillogram of three Speech Ensembles: 1. 'Val', 2. 'Pushkal' & 3. 'Pali'.

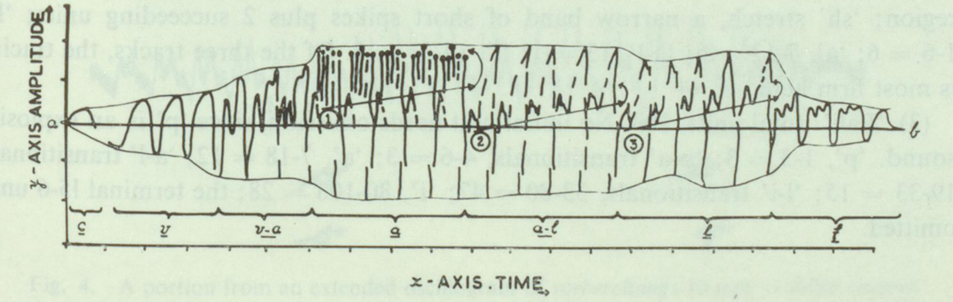

Fig. 3. Idealised diagram of the oscillogram track of 'val' showing the base-line a-b and 26 of the total 76 units extending on its either side linearly. The regions indicated by these selected units and their number in each region are as follows: *c* (commencement, 1), *v* (v-region, 3), *v-a* (transition *v-a*, 3), *a* (a-region, 5), a-1 (a-1 transition, 5), 1 (1-region, 5), *f* (fading, 4). The (1), (2) and (3) bigger arrows along with their terminal smaller curved ones indicate the direction and the diminishing extent of the three 'melody' traces' in *a*, *a*-1 and 1 regions. The envelope of the whole track in two light lines shows the range of amplitude; note that this is steady in *a* and more or less steady in a-1 while in *l*, it shows a median greatest width and a diminishing on either side.

All the tracks show horizontal succession of sharp repetitive units, as is common to any sound. Each unit is an elongated triangular tracing extending on either side of a baseline *ab* but varying in its detailed form at different regions. Assuming, as customary, that any major change in form here indicates a corresponding change in the phonetics of the oscillogram itself, i.e., the generation of a new phoneme, we can resolve each track into its constituent phonemes, which equal in number to these major changes. They are three for 'val', seven for 'pushkal', four for 'pali'. Depending on the detailed shape, these units can also be classified as simple ('v' of 'val', 'p' of 'pali') indented ('l' of 'pushkal'), peaked ('p' of 'pushkal') and spiked ('a' of 'val'). The EEG-like form of the spikes results because of the very rapid and numerous overtones not resolvable into just one or two peaks. For each phoneme, there is more than one unit, whose unique form usually changes gradually but very systematically to yield to the form of the next phoneme. Hence, besides the basic phoneme units, certain transition units are also to be recognised, e.g., as between 'v' and 'a' or 'a' and 'l' of 'val'. The number of the units in a track can be ascertained by their tallest points which is mostly one for one unit. Starting from left, we can fix their serial numbers as well (see 'pushkal'). If the characteristic form of a unit is not well developed anywhere, this may be due to commencement or fading out of pressure changes at the ends of the track (see 'val' track) or a noise trace ('sh' of 'pushkal'); such units may have to be left out of the total count.

On these grounds, we can draw a distribution analysis of the units for the three tracks as follows:

(1) 'val': total units, 76. 'v' region, 7 units; 'v-a' transitionals, 8; 'a' region, 17; 'a-l' transitionals, 30; 'l' region, 14; one initial and a few terminal units omitted.

(2) 'Pushkal': due to a prominent internal smooth region, i.e., a silence zone, the units fall into 2 sets, one, from 1-21, before this zone, the other, 1-21 again, after 'sh' region. 'p', 1-10 = 10; 'p-u' transitionals, 11-13 = 3; 'u', 14-21 = 8; silence region; 'sh' stretch, a narrow band of short spikes plus 2 succeeding units; 'k', 1-6 = 6; 'a', 7-12 = 6; 'a-l', 13 = 1; 'l', 14-21 = 8. Of the three tracks, the tracing is most firm here.

(3) 'Pali': total units, 108. No initial unit needs omission, since 'p' is an explosive sound. 'p', 1-3 = 3; 'p-a' transitionals, 4-6 = 3; 'a', 7-18 = 12; 'a-l' transitionals, 19-33 = 15; 'l-i' transitionals, 33-80 = 47; 'i', 80-108 = 28; the terminal 15-6 units omitted.

2. OVERALL FEATURES

A study of the overall form of the track also suggests significant conclusions (e.g., existence of internal 'melody traces') regarding the general development of the oscillogram as a whole. In 'val', within 'a' region and towards the vertices of the units, a dense packing of many minor spikes occurs in all the units successively (Figure 2).

In addition, these units themselves appear to form together a single laterally sustained tracing gradually diminishing upward, towards the left of the base line; this means that at this particular level of amplitude, a large number of individual sounds were produced in a *graded manner*. viz. to form a melody. Three such tracings are seen in 'val' (Figure 2 and Figure 3) in the initial region of 'a', in the transition region between 'a' and 'l' and in the middle and end region of 'l'. The melody trace in 'pali' is much simpler. It is a single trace continuous from 'a' to 'l' region. Its course however is very prominent over the entire track which itself looks as if twisted steadily upward.

3. DISCUSSION

In view of the well-known redundancy of speech at the physical level (Fry 1960), all the details of sound development that an oscillogram reveals as noted above, may not be necessary for the hearer's perception of a word. Neither can we hold, in spite of the acoustic faithfulness of the middle ear, that all these details are perforce reproduced unaltered at the basilar membrane. Nonetheless, the importance of studying the details of sound development as pressure changes exhibited by the oscillogram's breakup for understanding the mechanism at the basilar membrane eventually, cannot be minimised, as these two represent the two actual end points of the chain of mechanical events (Zalin 1961) during audition. The purpose of this paper was to draw attention to the wealth of useful data obtainable by an analysis of the form of the oscillogram tracks towards this objective. The importance of this study is likely to be more striking if a comparative study of the oscillograms of carefully selected contrastive speech ensembles is carried out on the same lines as here. For this, however, recording of more expanded nature as in Figure 4, are necessary.

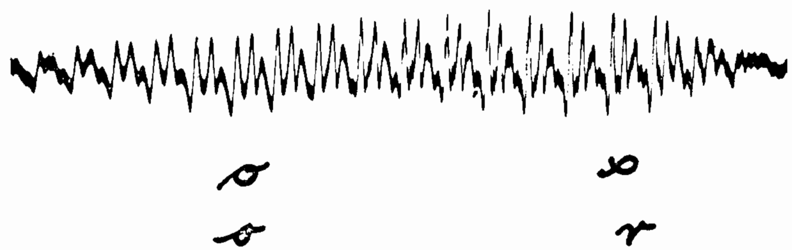

Fig. 4. A portion from an extended oscillogram of *vorbereitung*; 10 mm = 0.001 second.

Department of Biology
Jawaharlal Institute of Postgraduate
Medical Education and Research
Pondicherry, India

REFERENCES

Bekesy, G.V.
1960 *Experiments in Hearing* (New York).

Fry, D.B.
1960 "Linguistic Theory and Experimental Research", *Transactions of the Philological Society.*

Lawrence, M.
1966 "Comment on Mygind's Pressure Theory for Stimulation of the Labyrinthine Epithelium", *Archives of Oto-Laryngology* 83:10-11.

Sweetman, R.H. and P. Dallos
1969 *Journal of the Acoustical Society of America* 45:58-71.

Wever, E.G. and H. Lawrence
1954 *Physiological Acoustics* (Princeton University Press), p. 400.

Zalin, A.
1961 "On the Function of the Kinocilia and Stereocilia with Special Reference to the Phenomenon of Directional Preponderance", *Journal of Laryngology and Otology* 81:119-135.