TONGUE RETRACTION IS NOT SO USEFUL AFTER ALL

Sidney Wood

According to a century-old tradition in phonetics, small adjustments of tongue fronting or retraction provide an active and useful means of modifying vowel quality. However, acoustic theory does not support this tradition in every case. Stevens (1972) has demonstrated that spectral sensitivity to constriction location perturbations is not continuous along the vocal tract. His experiments with simple tubes indicated that when the anterior (palatal) portion, the mid (velar) portion (with lip-rounding) or the posterior (pharyngeal) portion are constricted vowel spectra are hardly affected by moderate displacement of the constriction location within those regions.

This can also be seen by studying three-parameter model nomograms (Fant 1960, Stevens and House 1955) which show that there are in fact four such regions where vowel spectra are relatively insensitive to location perturbations - along the hard palate and in the lower pharynx for spread-lip [i-ɪ]-like and [ɑ-æ]-like vowels, and along the soft palate and in the upper pharynx for rounded [u-ʊ]-like and [o-ɔ]-like vowels. The same four regions were deduced from a spectrographic study of eskimo vowels (Wood 1971) and X-rayed vowel articulations confirm that these regions are exclusively used for vowels in speech (Wood 1977).

Advancing and retracting palatal and velar vowels means that the constriction is displaced along the hard and soft palates. Theoretically, this should yield but little spectral advantage since these are two of the four regions mentioned above. Three model experiments were designed to repeat Steven's experiments, this time to test the sensitivity of vowel spectra to perturbations of the constriction location in natural human vocal tract configurations. Mid-sagittal vocal tract profiles were systematically modified by retracting the tongue body from palatal [i] and [ɛ]-like configurations and by advancing the tongue body from a velar [u]-like configuration. The resonances of each new configuration were found by sweeping a line electric analogue (LEA at the Royal Institute of Technology, Stockholm, by courtesy of Gunnar Fant), Fig. 1.

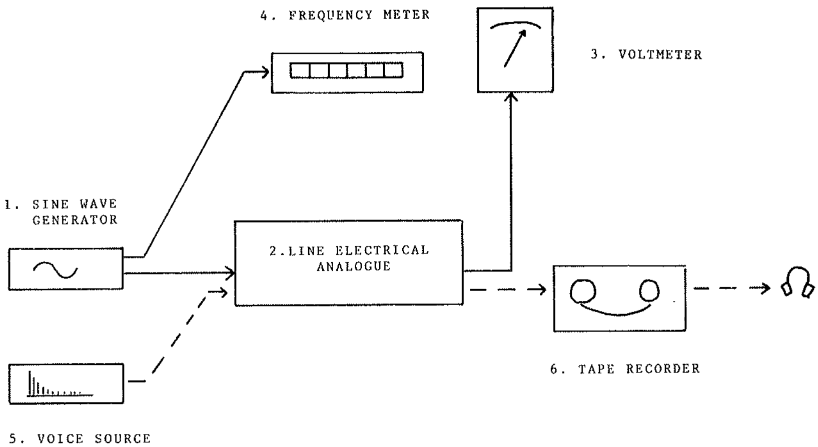

Fig. 1. The configuration to be measured is set upon the analogue (2). A sine wave from the generator (1) is passed through the analogue to a voltmeter (3). Voltage maxima occur at resonance frequencies which can be read off from the frequency meter (4). For monitoring and recording synthetic vowel qualities, a voice spectrum from a voice source (5) passes through the analogue to a tape recorder (6).

The tongue body was retracted 5 mm in 1 mm steps from palatal [i]-like and [ɛ]-like configurations. This amounted to 20 mm retraction of the constriction along the domed roof of the mouth. Finally, the tongue was advanced 6 mm in 2 mm steps from a velar [u]-like configuration. This amounted to 20 mm advancement of the constriction along the soft palate. The degree of constriction was kept constant in each experiment (cross-section area at the constriction 0.65 cm^2 for [i] and [u], 2.6 cm^2 for [ɛ] so the only variable was the constriction location with consequent modifications to the front and back cavities.

The modification of the [i] profile is illustrated to the left in Fig. 2. A similar modification was made for the [ɛ] configuration. The range of movement represents retraction from a prepalatal constriction through midpalatal to postpalatal. X-rayed vowel articulations confirm universal language-specific preferences for either the prepalatal or midpalatal posi-

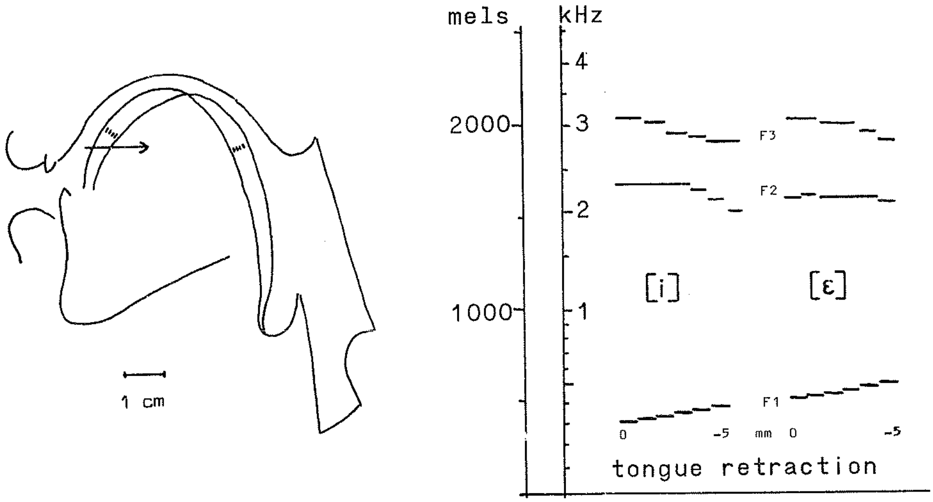

Fig. 2. Retraction of the tongue along the hard palate from palatal [i] (left). A similar modification was made for [ε]. Spectral consequences of these modifications (right).

tions in speech (Wood 1977). The spectral consequences of the modifications are given to the right in Fig. 2.

F_3 fell continuously in both [i] and [ε] for each retracted step. This is why the prepalatal [i] of say Swedish or Russian sounds sharper than the midpalatal [i] of English (Fant 1960) and why Swedish [e] and English [i] sound alike to Swedes. F_2 was hardly affected by retraction through the prepalatal and midpalatal locations and did not begin to fall until the end of each series. It fell appreciably in the last two (postpalatal) steps from [i] and had just started to fall at the last step from [ε]. F_1 rose gradually. The trend of these results is predictable from the pressure and volume velocity standing waves and from the energy distributions (Chiba and Kajiyama 1941, Fant and Pauli 1975). The consequences of widening the prepalatal part and narrowing the postpalatal part are (i) contrary and largely self-cancelling for F1, (ii) negligible for F2 and (iii) cumulatively negative for F3 (on account of the latter's prepalatal volume velocity node and velar pressure node).

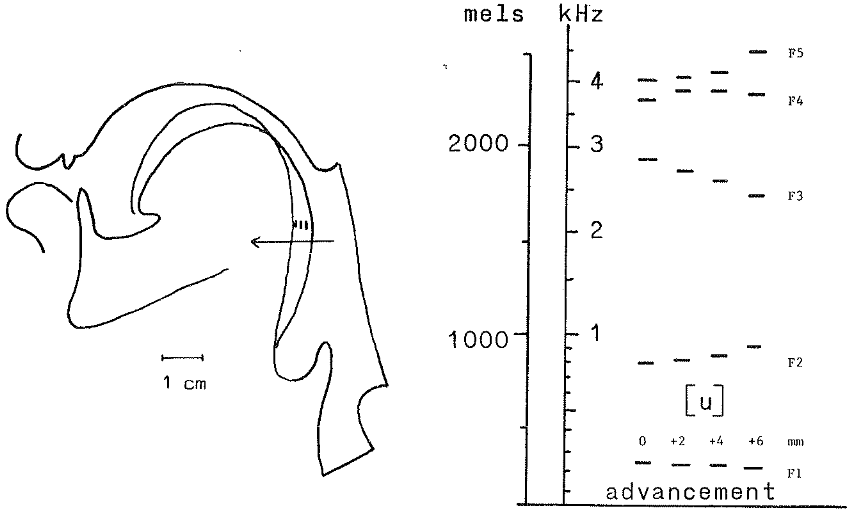

Fig. 3. Advancement of the tongue along the soft palate from velar [u] (left) and the spectral consequences (right).

The lack of change in F_2 from the prepalatal to midpalatal locations is expected from Stevens's finding. The postpalatal location is apparently beyond the region where formant frequencies are relatively insensitive to variation of the constriction location. F_2 changes rapidly when the constriction is retracted beyond the midpalatal position and would require considerable articulatory control in a monophthong. In actual speech vowel constrictions occur only fleetingly in this region, either during transitions to or from adjacent consonants or during diphthongs where the main information is in the gliding formant.

The modification of the [u] constriction and the spectral consequences are given at Fig. 3.

The main result of advancing the tongue body along the soft palate was that F_3 fell sharply, a cumulative consequence of widening the upper pharynx where F3 of an [u]-like configuration has its largest pressure maximum and a considerable excess of potential over kinetic energy and simultaneously narrowing the postpalatal region. Where F3 has a pressure node and an excess of kinetic energy. F_2 was hardly affected by advancing the constriction at the posterior end of the soft palate, which agrees with Stevens's finding. F_2 began to rise when the constriction was near the middle of the soft palate and it had risen about 120 Hz when the constriction was advanced to the front end. This is within the range of F_2 variation found in natural speech for [u]-like vowels.

Although it has been known for many years that the conceptual basis and assumptions of the traditional model were largely false, the implications have not yet been fully drawn. At first attention turned away from articulation and towards the composition of the acoustical signal and to the reactions of listeners to acoustical cues. But it is still highly relevant to ask what the speaker is doing when he produces those acoustical cues that we now know the listener needs. It is especially important in view of the growing interest in articulatory programming and the motor control of speech. Analysis of X-ray films (Wood 1977) showed that for the articulation of vowels the tongue aims to narrow the vocal tract at one of the four regions mentioned above, a simpler task than had hithertoo been envisaged. The tongue musculature was found to be admirably situated for creating the four constrictions and the sphincter function of the palatoglossi and the pharyngeal constrictors ensure the accuracy of all but the palatal manoeuvres. The experiments reported above confirm that vowel spectra are relatively insensitive to location perturbations in those regions. Not only is voluntary displacement of the constriction location physiologically unlikely, there is little spectral advantage to be gained from doing it anyway.

References

Chiba, T. and Kajiyama, M. (1941), The vowel, its nature and structure. Tokyo

Fant, C.G.M. (1960), The acoustic theory of speech production. The Hague: Mouton

Fant, C.G.M. and Pauli, S. (1975), Spatial characteristics of vocal tract resonance modes. Speech Communication 2, 121-132. Stockholm: Almqvist & Wiksell

Stevens, K.N. (1972), The quantal nature of Speech: evidence from articulatory-acoustic data. In David, E.E. and Denes, P.B. (eds), Human Communication, a Unified View, 51-66

Stevens, K.N. and House, A.S. (1955), Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America 27: 464-495

Wood, S. (1971), A Spectrographic Study of allophonic variation and vowel reduction in West Greenlandic Eskimo. Working Papers 4: 58-94, Department of Linguistics, Lund University

Wood, S. (1977), A radiographic analysis of constriction locations for vowels. Working Papers 15: 101-131, Department of Linguistics, Lund University.