MEASUREMENTS OF FORMANT BANDWIDTHS IN THE TIME DOMAIN

Lennart Nord[1]

Abstract: Two methods of estimating formant bandwidth in the time domain were tried. The vocal tract with the glottis closed was excited, either by pulses or by a gated sine wave signal, with its frequency adjusted to one of the resonances of the vocal tract. Bandwidth values were obtained from the decay curves. Measurements on Swedish vowels, articulated by two male and two female speakers were performed and the results compared to sweep-tone measurement data.

Introduction

Acoustic data such as bandwidth values are needed for a thorough modelling of the vocal tract, including losses. Sources of error when measuring bandwidths are glottal and nasal coupling which will broaden the formant peaks. A movement of the articulators, shifting the resonance frequency will also result in too large values.

Two techniques of measuring bandwidth in the time domain will be described. The experiments are part of a series of experiments, investigating the acoustical properties of the vocal tract. (For some aspects of the wall impedance of the vocal tract, see Fant, Nord and Branderud (1976).)

Theory

A resonance can be described in the frequency or in the time domain, see fig. 1. The bandwidth B in the frequency plane

---

1) Department of Speech Communication, Royal Institute of Technology (KTH), S-100 44 Stockholm 70, Sweden.

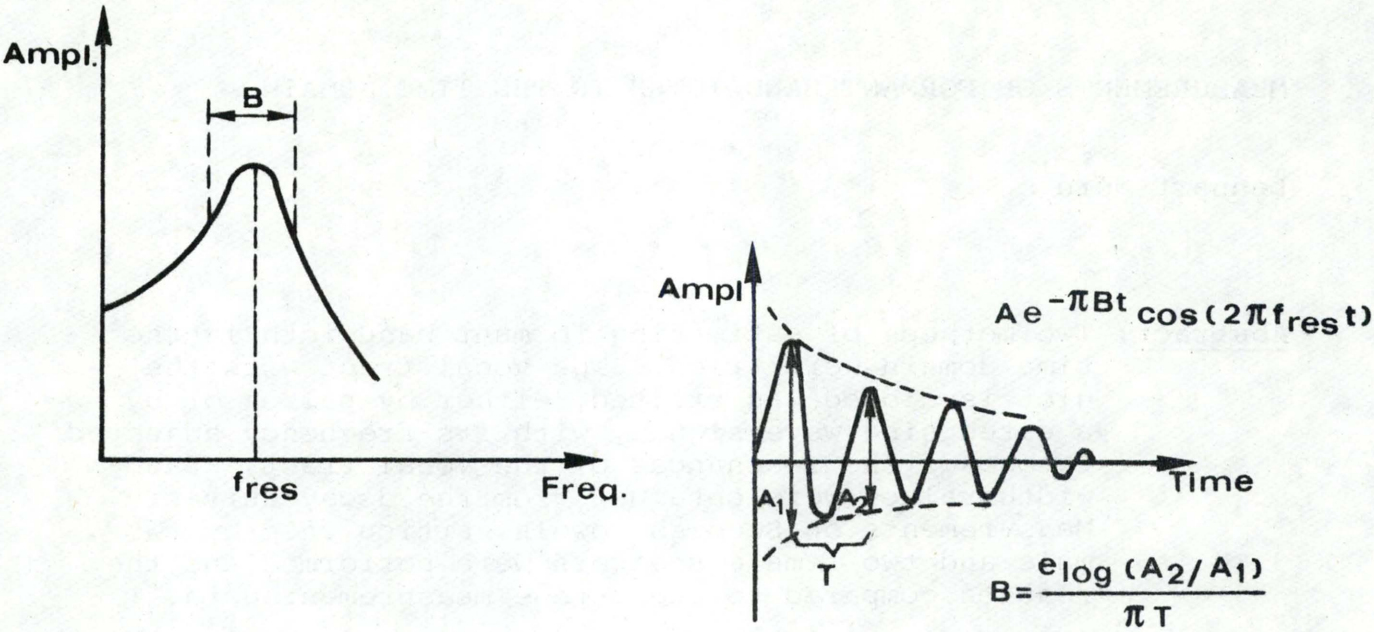


Figure 1

a) Frequency response, and b) impulse response for a resonator.

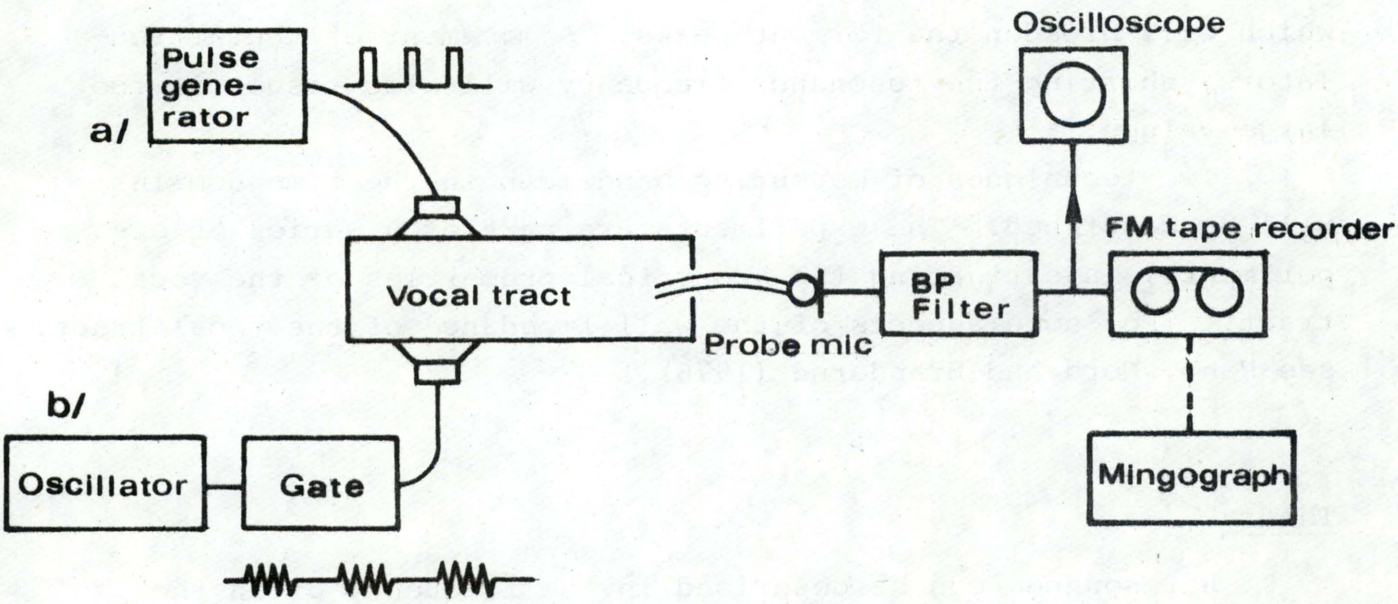


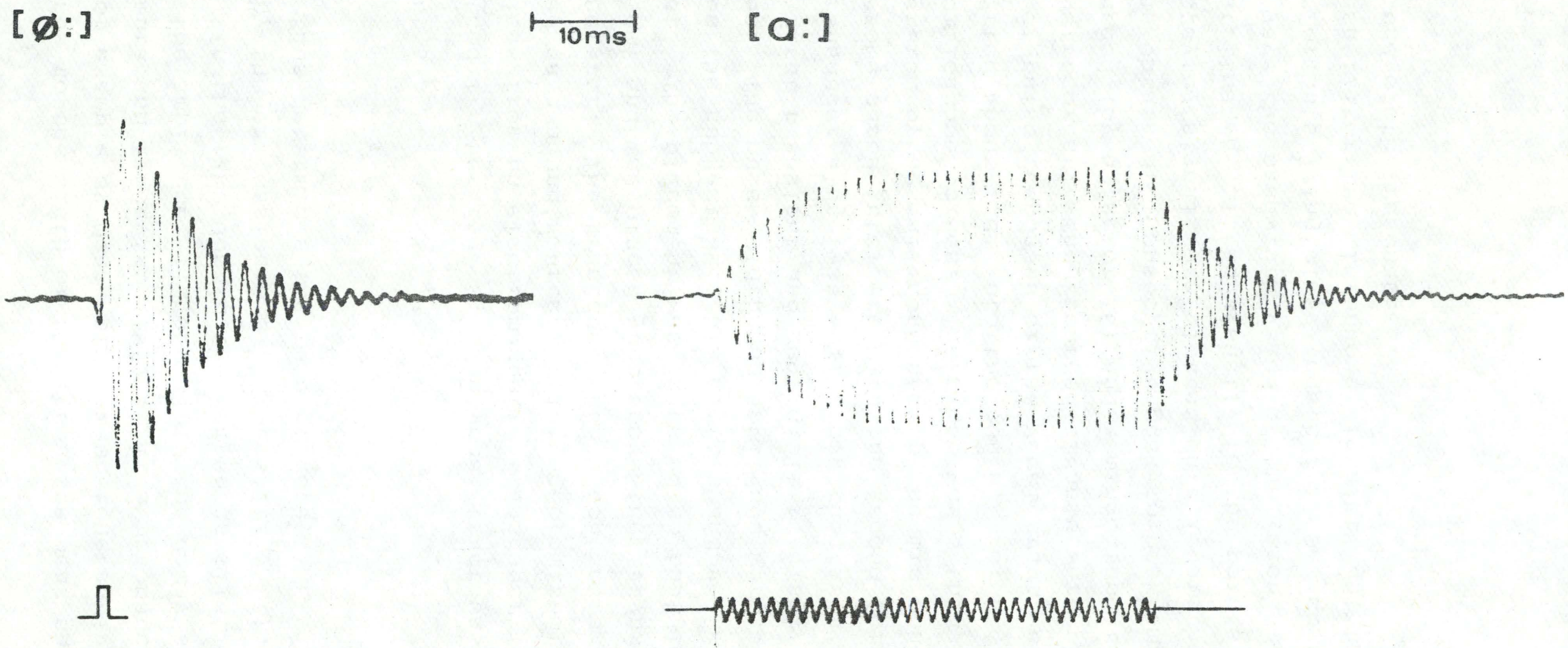
Figure 2

Two methods of bandwidth measurement.
a) Pulse excitation.
b) Gated sine wave excitation.

corresponds to the width of the resonance peak and can be measured, for example by means of a sweep-tone analysis technique, see Fujimura and Lindqvist (1971). If a resonance is excited by a pulse, the amplitude will oscillate and decay according to fig. 1b. In the time domain the damped oscillation can be described by the expression $A \cdot e^{-\pi Bt} \cdot \cos(2\pi f_{res} \cdot t)$, where the factor B thus determines the rate of decay.

Experiment

Method 1 See fig. 2. The vocal tract was excited by short pulses from an electromagnetic transducer (of a type normally used as a sound source for laryngectomized) driven by a pulse generator. The transducer had a contact area of approximately 7 cm^2 and was placed just above the larynx on one side of the throat. The placement was adjusted so that loud clicks were heard when the subject held his glottis and velum closed and his mouth open. The pulse width and amplitude were adjusted to give a strong first transient and negligible ringing. (The resonance of the transducer alone was well above the first formant but could in some cases interfere with the second formant.) The sound was registered by a small condenser probe microphone and displayed on an oscilloscope screen. The probe was a thin damped plastic tube inserted about 1 cm into the mouth. For a given articulation a number of pulse responses of the vocal tract were recorded on an FM tape recorder. The recorded signals were played back at a reduced speed (by a factor 16) and a mingograph tracing of the signal was obtained. Consecutive peak-to-peak amplitude values of the decaying signal were plotted on a logarithmic scale and a straight line was fitted visually to the points. The slope of the line gave the bandwidth value of the resonance. For vowels with two formants lying close together it was sometimes necessary to bandpass filter the microphone signal to get the bandwidth value of one of the resonances.



<u>Figure 3</u>

Mingograms of pulse excitation response and gated sine wave excitation response signals. The pulse and sine wave are indicated.

Method 2 The exciting pulse did not correspond to an ideal pulse, which meant that only little energy was transmitted through the vocal tract wall at higher frequencies. This made it difficult to measure anything more than the bandwidth of the first formant reliably.

Another way of exciting the tract was therefore tried. The transducer was fed with a sine wave signal of the resonance frequency of interest. By gating this signal, the rise and decay pattern of the recorded response signal could be treated in the same way as the pulse response. In fig. 3 both types of response signals are shown together with the indicated pulse and sine wave excitation signals. The latter method demanded less energy from the transducer as all the energy was used; however, the mechanical constraints of the system, even with this method, made it difficult to get bandwidth data from higher formants.

Measurements

Both methods were tried on a number of male and female subjects. Some practice was usually needed to make the subjects aware of the type of articulation needed with the glottis closed and no nasal coupling. Method two also required that the articulation was held fixed while the first resonance frequency of the tract was found with the oscillator. This could be avoided by some automated frequency scanning.

Results and discussion

Results for two male and two female speakers are plotted in fig. 4. No consistent difference was found between the measurement values for the two methods that were tried. Therefore, no distinction is made in fig. 4.

Included in fig. 4 are also average values from sweep-tone measurements made by Fujimura and Lindqvist (1971). The similari-

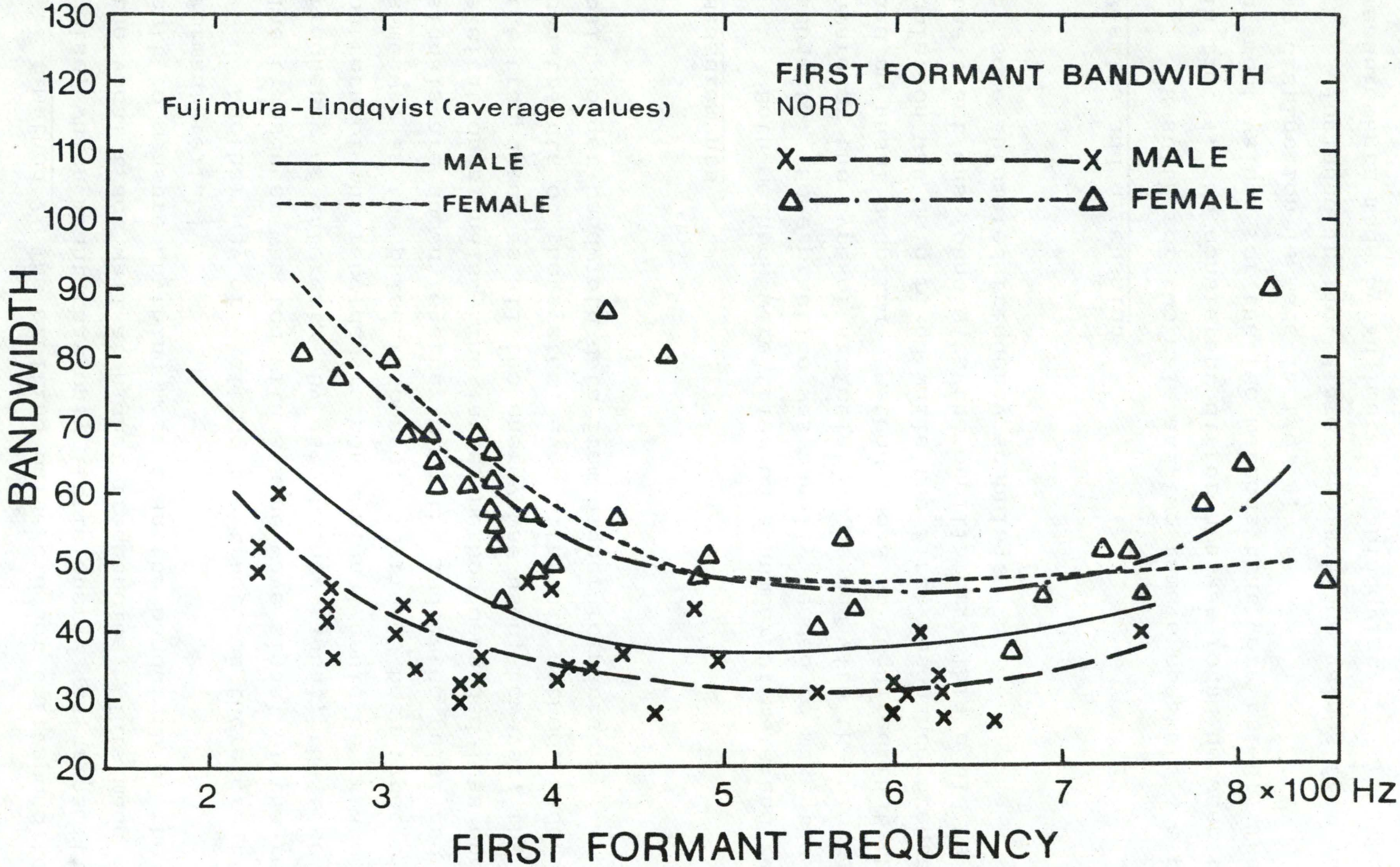


Figure 4

Bandwidth versus first formant frequency for two male and two female speakers. Average values from sweep-tone measurements by Fujimura and Lindqvist (1971) are included as a comparison.

ties are quite good with the same tendencies concerning male-female differences and high bandwidth values when F_1 is low, due to increased wall losses. For discussions on bandwidths, see articles by Fujimura and Lindqvist (1971) and by Fant (1972).

The time domain data in fig. 4 appear to have somewhat lower values than the sweep-tone measurements. This might be explained by the fact that measurements in the frequency domain take longer time and will give broader peaks if the resonance frequency shifts during the frequency sweep. Time domain analysis is generally faster and a changing of the articulation will not affect the bandwidth to the same degree.

If two formants are very close, such as F_1 and F_2 for high back vowels, a satisfactory filtering of F_1 cannot be made. The impulse response will not decay exponentially and B_1 cannot be evaluated in the manner described above. (Compare the problem of extracting bandwidths in the frequency plane when two formant peaks lie close together.)

An alternative method of pulse excitation is to use an electric spark as an acoustic source inside the mouth cavity. This technique has been described by House and Stevens (1958).

Summary

The two methods of estimating bandwidths in the time domain of the vocal tract, closed at the glottal end, gave promising results for the measurements of B_1. Values for some vowels articulated by two male and two female speakers were consistent with sweep-tone measurement data obtained by Fujimura and Lindqvist (1971).

References

Fant, G. (1972): "Vocal tract wall effects, losses and resonance bandwidths", STL-QPSR 2-3, p. 28-52

Fant, G., L. Nord and P. Branderud (1976): "A note on the vocal tract wall impedance", STL-QPSR 4, p. 13-20

Fujimura, O. and J. Lindqvist (1971): "Sweep-tone measurements of vocal tract characteristics", JASA 49, p. 541-558

House, A.S. and K.N. Stevens (1958): "Estimation of formant band widths from measurement of transient response of the vocal tract", JSHR 1/4, p. 309-315