Dept. for Speech, Music and Hearing
Quarterly Progress and Status Report

Indirect studies of glottal cycles by synchronous inverse filtering and photo-electrical glottography

Fant, G. and Sonesson, B.

KTH Computer Science and Communication

journal: STL-QPSR
volume: 3
number: 4
year: 1962
pages: 001-003

http://www.speech.kth.se/qpsr

I. SPEECH ANALYSIS

A. INDIRECT STUDIES OF GLOTTAL CYCLES BY SYNCHRONOUS INVERSE FILTERING AND PHOTO-ELECTRICAL GLOTTOGRAPHY

Two effective methods for indirect studies of the vibration of the vocal cords have developed in recent years. One method, the inverse filtering, has come to frequent use, see Miller [1], Lawrence [2], Holmes [3], Fant [4][5], and Cederlund et al [6]. The inverse filtering, also referred to as antifiltering, is a process whereby the formant structure is removed from the electrical speech wave resulting in a periodic function with a waveform which within certain limits reflects the volume velocity of air passing the vibrating vocal cords. The other indirect method is the photo-electric glottography developed by B. Sonesson [8]. A powerful light source is applied to the skin of the neck just below the larynx. Residual light transmitted through the flesh into the subglottal system is scattered through the vibrating glottis and picked up by a lightprobe in connection with a photocell. An oscillographic record of the photocell output is a periodic curve proportional to the amount of light emitted through the vibrating glottis and should accordingly display the time-varying area of the glottis slit.

The time varying glottal area and the volume velocity flow are approximately proportional [7][9] and the glottogram may thus be expected to resemble the inverse filtered speech wave. A direct test of the degree of similarity of the two methods was accordingly attempted. The experiment took place in a sound insulated room at the Speech Transmission Laboratory. A 28 years old male with baritone voice served as a subject. In order to suppress the gag reflex the subject's mouth and throat was sprayed with a lightly anesthetizing solvent. The condenser microphone for the inverse filtering was placed 5 cm in front of and at the side of the speaker's lips. The microphone and associated amplifier had a frequency response down to less than 20 c/s.

The inverse filtering comprised two units of the type described in ref. [5] and a low-pass filter of third order adjusted for a 1400 c/s cutoff. The vowel sustained by the subject had a

first formant of 750 c/s and a second formant of 1600 c/s corresponding to a phonetic quality of an open [æ]. A pitch of close to 100 c/s was maintained. The glottograph curve and the inverse filter output were displayed on a pitch synchronized dual-beam cathode ray oscilloscope and photographs were taken with a polaroid camera, see Fig. I - 1, for typical examples.

The traces show striking similarities with regard to the general waveshape. In phonation A the agreement is very good and in B the glottograph had a larger scale factor. The peak of the wave is somewhat sharper for the glottogram than for the inverse filter curve which is not at present explainable on theoretical grounds. A promising feature is the similar trend in the closed and opening phase of the glottal period. Differences in these respects from one phonation to another reflect differences in voice quality, and possibly the relative degree of leakage in the closed interval. Non-horizontal curves in the closed phase are presumably not an artifact due to faculty compensation with the antifilter as previously suspected. The general agreement between the glottographic curves and the inverse filter curves, at least in the overall waveshape, adds extra faith to the reliability of either method. However, several tests of more detailed nature are needed before one can recommend inverse filtering as a clinical routine method for indirect studies of voice mechanism and as a partial correlate of voice quality. The same is also true of the glottogram. The inverse filtering has its merits in less strain on the patient. A system with combined formant tracking and antifiltering of F1 and F2 would in addition allow the study of the main features of glottal waveshape in connected speech.

G. Fant and B. Sonesson

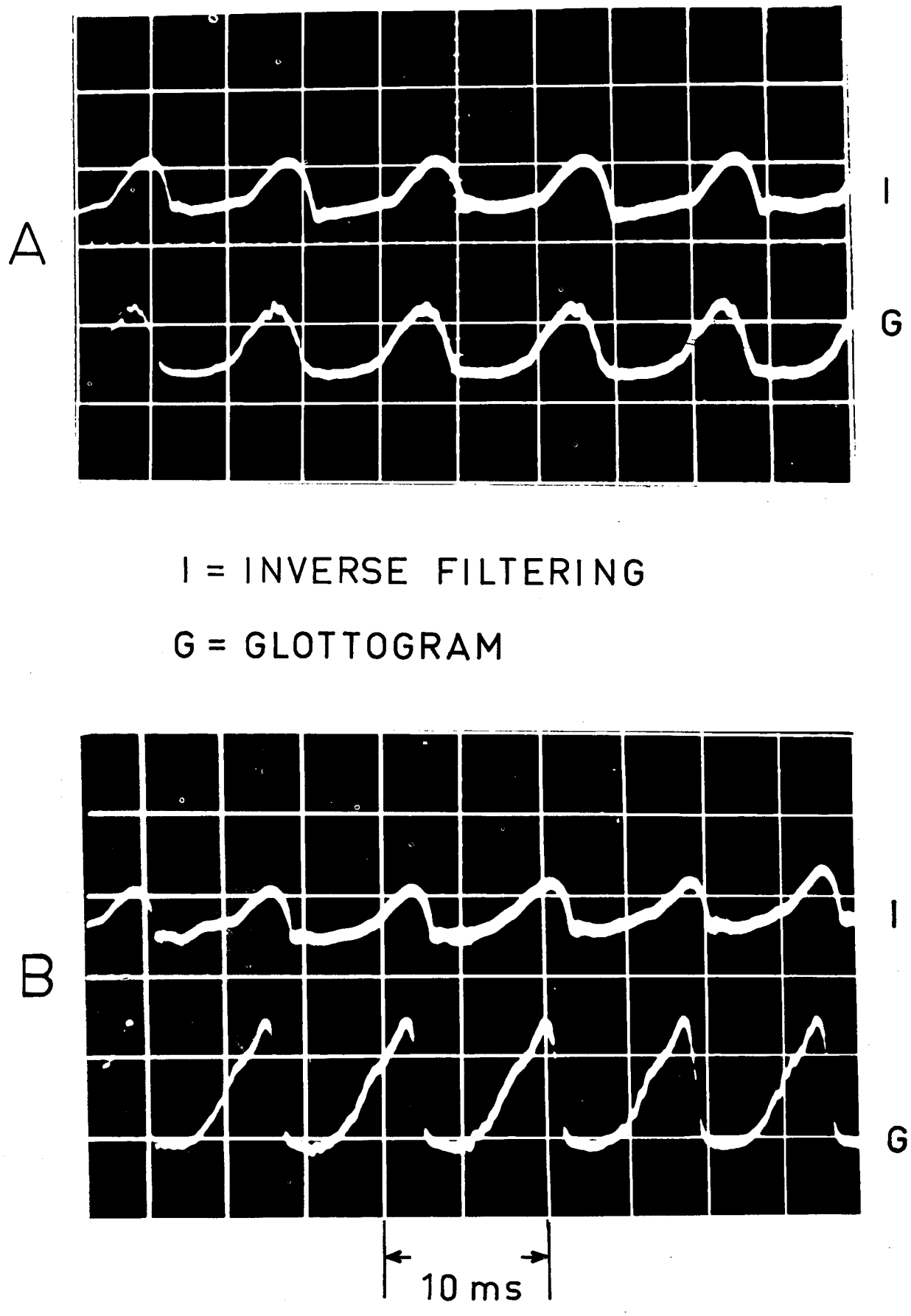

Fig. I - 1. Synchronous records of glottal pulsations by means of inverse filtering and glottograph. A and B are two separate phonations. F_0 = 105 c/s.

References:

(1) Miller, R.L.: "Nature of the vocal cord wave", J.Acoust.Soc.Am. 31 (1959), pp. 667-677

(2) Lawrence, W.: "Formant tracking by self adjusting inverse filtering", Paper C5 presented at the Speech Communication Seminar, Stockholm 1962, to be publ. in the Proc. from this Seminar

(3) Holmes, J.N.: "The effect of simulating natural larynx behaviour on the quality of synthetic speech", Paper F6 presented at the Speech Communication Seminar, Stockholm 1962, to be publ. in the Proc. from this seminar

(4) Fant, G.: "The acoustics of speech", Proceedings of the Third International Congress on Acoustics Stuttgart, 1959, Ed. L. Cremer, Amsterdam 1961, pp. 188-201

(5) Fant, G.: "A new anti-resonance circuit for inverse filtering", STL-QPSR 4/1961, pp. 1-6

(6) Cederlund, C., Krokstad, A., and Kringlebotn, M.: "Voice source studies", STL-QPSR 1/1960, pp. 1-2

(7) Flanagan, J.L.: "Some properties of the glottal sound source", J. of Speech and Hearing Research 1 (1958), pp. 99-116

(8) Sonesson, B.: "On the anatomy and vibratory pattern of the human vocal folds", (Thesis work) Acta Oto-Laryngologica Suppl. 156, Lund 1960

(9) Fant, G.: Acoustic theory of speech production, 's-Gravenhage 1960