FOURIER ANALYSES OF PHOTO-ELECTRIC GLOTTOGRAMS

Børge Frøkjær-Jensen and Jørgen Rischel

1. Glottis area versus volume velocity curve.

According to Flanagan (2), (3) there seems to be a good agreement between the glottal area and the volume velocity wave when operating in a constant and moderate subglottal pressure range of about 4 cm H_2O, - the effects of inertance neglected. It should therefore (within the mentioned restrictions) be possible to calculate the acoustic spectrum of the primary voice source based upon photo-electric glottograms (but not upon the Fabre glottograms)*) by means of a digital computer.

According to Fant and Sonesson (1) simultaneous recordings of photo-electric glottograms and inverse filterings seems to give nearly the same waveforms, which supports the theoretical calculations of Flanagan and Fant.

2. Expected differences in the glottal waveforms of vowels and consonants.

A problem of great current interest is the possibility of interaction between poles and zeros of the vocal tract transfer function and the glottal volume-velocity waveform.

The supraglottal impedance relationships are analysed by Flanagan (reference 3, pp. 46-47) on the basis of transmission line theory. He finds that the driving point impedance of the uniform vocal tract (set in position for the neutral reference vowel) reaches maxima at those frequencies where the imaginary term of the propagation constant **) $x\ell$ equals $\pi/2$, $3\pi/2$, etc. The impedance at these maxima, which coincide with F_1, F_2, etc. is determined as the square of the characteristic impedance of the tract divided by the oral radiation impedance. It is clear that the driving point impedance reaches its highest maximum when the radiation impedance is low, i.e. at the F_1 frequency,

---

*) This has actually been done by C. Bordone-Sacerdote & G. Righini, "Glottal wave detected as a high frequency modulation", 5th Int. Congr. of Acoustics (Liège 1965), Paper A 34.

**) These formulae are based on the simplifying assumption that the real term is zero, i.e. the vocal tract is lossless.

where the driving point impedance is so high that tract-source interaction occurs (reference 3, p. 47).

Halle and Stevens (reference 4, p. 268) calculate the conditions at different first formant frequencies. They take the tract as a resonant circuit tuned to F1 and find by considering the impulse response of this supraglottal filter that it will react to the glottal wave with a pressure peak which at low F_1 values (F_1 below 250 cps) is of the order of the steady subglottal pressure. Thus the glottal vibrations will clearly be impeded at very low F_1 frequencies.

The authors presume that this effect is counteracted by adjusting the vocal cords towards a more open state so that they can vibrate with a lower pressure drop across them. This adjustment is assumed to condition the longer duration of vowels before voiced than before voiceless obstruents (in English). *)

In order to test whether the Fourier analysis of photoelectric glottograms constitutes a usable method for empirical investigation of these problems we have made a few analyses of typical glottograms. These are, of course, to be taken as pilot experiments only.

3. Experimental procedure.

The two sustained sounds [æ:] and [z:] spoken several times by the authors have been recorded by means of the photoelectric glottograph. Because of lack of a high-speed analog/digital converter at our laboratory we have photographed the oscilloscope screen and by hand measured 281 ordinates per cycle from large scale photographic enlargements of two typical glottograms. The ordinate values have been used as input data for a Fourier analysis made by computer. The computer outputs, parts of which are shown in Fig. 1 and 2, contain the sine and cosine coefficients, the phase differences, the absolute amplitude, and the logarithmic amplitude (in dB). Based on these outputs the line spectrums in Fig. 3 have been drawn. These compare the two primary voice source functions: that of the vowel and that of the consonant.

---

*) For discussion of this problem: see the article in this report by Eli Fischer-Jørgensen.

HARMONICS NO.	FREQUENCY IN CPS	AMPLITUDE IN MVOLTS	AMPLITUDES DECIBELS
DC	0	1090.11833	40.75
1	120	796.87701	38.03
2	240	306.75773	29.74
3	360	40.70412	12.19
4	480	33.45657	10.49
5	600	14.34248	3.13
6	720	15.20397	3.64
7	840	3.03825	-10.35
8	960	4.10174	-7.74
9	1080	1.71341	-15.32
10	1200	4.02491	-7.90
11	1320	1.63957	-15.71
12	1440	0.84083	-21.51
13	1560	0.78087	-22.15
14	1680	1.79724	-14.91
15	1800	0.59086	-24.57
16	1920	0.98869	-20.10
17	2040	0.59233	-24.55
18	2160	0.67564	-23.41
19	2280	0.58004	-24.73
20	2400	0.37273	-28.57
21	2520	0.55248	-25.15
22	2640	0.12864	-37.81
23	2760	0.65609	-23.66
24	2880	0.35398	-29.02
25	3000	0.51560	-25.75

Fig. 1.

HARMONICS NO.	FREQUENCY IN CPS	AMPLITUDE IN MVOLTS	AMPLITUDES DECIBELS
DC	0	1314.15202	42.37
1	120	805.75991	38.12
2	240	145.58764	23.26
3	360	86.47672	18.74
4	480	14.93569	3.48
5	600	9.15732	-0.76
6	720	5.18039	-5.71
7	840	3.07824	-10.23
8	960	2.38402	-12.45
9	1080	3.46416	-9.21
10	1200	0.91052	-20.81
11	1320	1.27400	-17.90
12	1440	1.15195	-18.77
13	1560	1.31342	-17.63
14	1680	0.99831	-20.01
15	1800	1.06320	-19.47
16	1920	0.36361	-28.79
17	2040	1.28072	-17.85
18	2160	0.89437	-20.97
19	2280	0.45730	-26.80
20	2400	0.65082	-23.73
21	2520	0.69761	-23.13
22	2640	0.44947	-26.95
23	2760	0.50149	-25.99
24	2880	0.33686	-29.45
25	3000	0.74337	-22.58

Fig. 2.

4. Results.

4.1. The glottograms.

The photographs in Fig. 3 show typical differences in the waveform of the glottal pulse. The vowel has a very symmetrical pulse (this is not always the case) with steep slopes, whereas the consonant pulse has a less steep slope in the closing phase. If, arbitrarily, we replace the lowest 7 per cent of the amplitude of each glottogram by a straight line (some of the slight bottom curvature is probably associated with transillumination and nonlinearity phenomena, compare reference 3, p.44), the "open" phase (triangular portion) of the vowel pulse is 4.2 ms = 51 per cent of the total duration of the cycle, whereas the corresponding part of the consonant pulse is 5.6 ms = 68 per cent of the total duration of the cycle. The fundamental frequency is the same in both analyses: 120 cps, i.e. the total duration of the cycle is 8.3 ms (1.4 ms per division on the oscilloscope screen). - The curves shown in Fig. 3 do not give any measure of the absolute degree of closure obtained in the two types of sounds, so we are not making any deductions concerning the degree of leakage through the glottis from these curves (cp., however, the curves of Danish voiced fricatives on p.13 of this report).*)

The differences observed may condition a weaker excitation and a greater damping of formants in the consonant.**)

4.2. The line spectra.

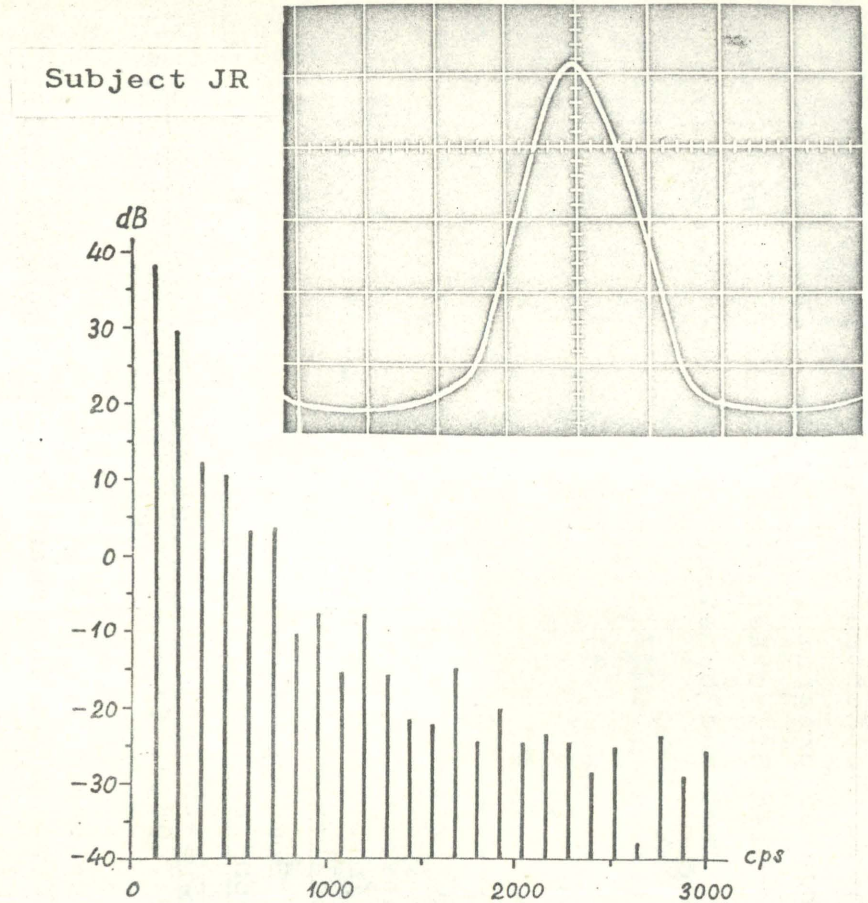
The line spectra for the two sounds are shown below the photographs in Fig. 3.

Up to about 1000 cps (the 9th harmonic) the slope of the primary voice source rolls off at a rate of 12-15 dB/octave, which agrees well with other measurements. Above 1000 cps the line spectrum is not reliable because the dynamic range is limited to about 60 dB. In future experiments we are going to use some kind of high-shaping network, which may extend the usable frequency range considerably.

---

*) See also Lisker, Abramson, Cooper, and Schvey in Status Report (SR-5/6), Haskins Laboratories, 1966, pp. 4.1-8.

**) The absolute amplitudes for the two sounds in Fig. 3 cannot be directly compared, since the recordings have no calibration.



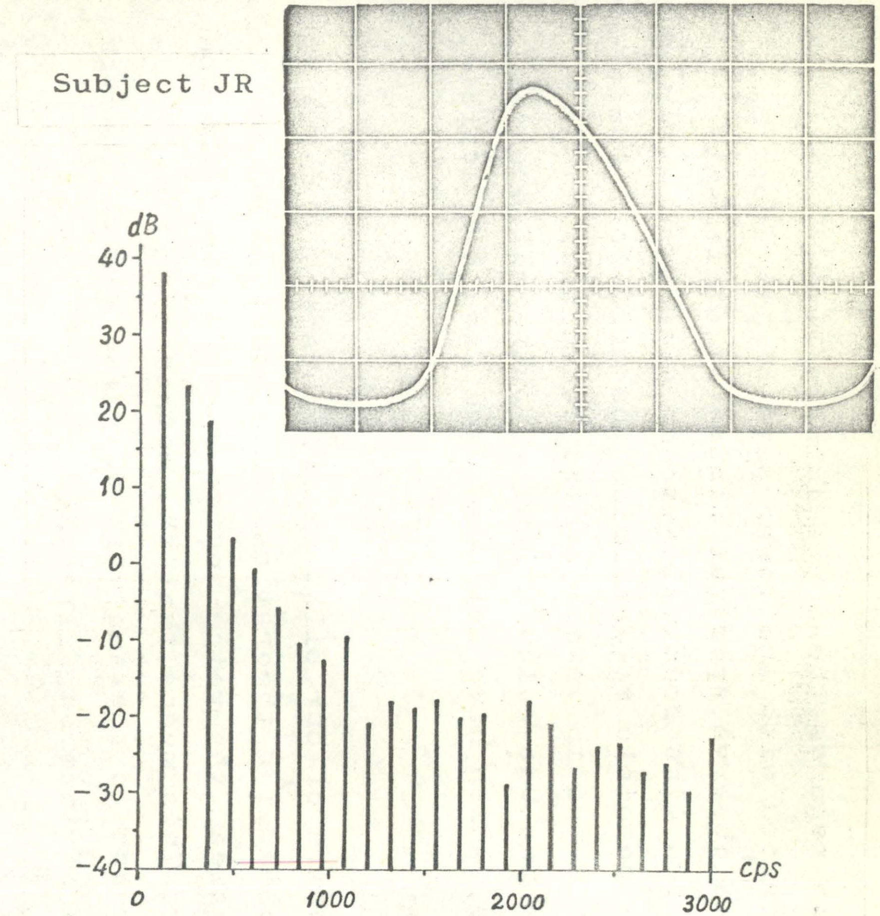


Fig. 3.

Glottogram and line spectrum of [ɛ:]

Glottogram and line spectrum of [z:]

Results of Fourier analysis of the glottograms are shown in the photocopies of the computer analysis.

Therefore, as regards these preliminary Fourier analyses we can only be concerned with the lower part of the spectrum. In this range we can observe spectral minima at different frequencies in the two line spectra. With an extensive material it will probably be possible to determine the different locations of the voice source zeros in vowels contra consonants.

Our analyses include calculations of phase differences among the harmonics, which have not been used for this article.

5. Conclusions.

The line spectra of our pilot experiments seem to be in fair agreement with earlier theoretical data and to confirm that differences between the voice source spectra of vowels and consonants do occur. This means that the method used here seems to be a usable approach to the analysis of voice source spectra parallel to the use of inverse filtering.

Acknowledgement.

We want to thank the Northern Europe University Computing Center (NEUCC) for run of programs, and "Statens Almindelige Videnskabsfond" (the Danish State Research Foundation), which has supported the development and construction of the glottograph.

References.

(1) G. Fant, B. Sonesson, "Indirect Studies of Glottal Cycles by Synchronous Inverse Filtering and Photo-electrical Glottography", STL-QPSR 4(1962), pp.1-3.

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

(3) J.L. Flanagan, Speech Analysis, Synthesis and Perception. Springer Verlag (1965). Chapter 3.5, The Source for Voiced Sounds.

(4) M. Halle, K.N. Stevens, "On the Mechanism of Glottal Vibrations for Vowels and Consonants", QPR 85(1968), MIT.

(5) R.L. Miller, "Nature of the Vocal Cord Wave", Journal of the Acoustical Society of America 31(1959), pp. 667-677.