Data on the glottal voice source behavior in vowel production

Gauffin, J. and Sundberg, J.

journal: STL-QPSR
volume: 21
number: 2-3
year: 1980
pages: 061-070

http://www.speech.kth.se/qpsr
IV. MUSIC ACOUSTICS

A. DATA ON THE GLOTTAL VOICE SOURCE BEHAVIOR IN VOWEL PRODUCTION*

J. Gauffin and J. Sundberg

Abstract

This paper investigates the relationships between certain characteristics of the glottal volume velocity waveform, i.e., the glottogram, and the SPL. Data were collected from a singer and a non-singer. They sustained a vowel at constant fundamental frequencies and continuously changing degrees of vocal effort. The results show that the maximum rate of change of the transglottal airflow during the glottal vibratory cycle, i.e., the maximum amplitude of the differentiated glottogram, predicts the SPL of the resulting vowel with an accuracy of ± 3 dB, approximately, given the fundamental and formant frequencies. This seems to apply to sustained vowels phonated at different fundamental frequencies, in different registers and in differing types of phonation. Our results agree with predictions based on the Fant (1979) mathematical model of glottogram waveforms.

Introduction

Let us consider Fig. IV-A-1 showing an idealized glottal volume velocity waveform during phonation, i.e., a glottogram. The glottogram can be described as a waveform where each cycle consists of three phases: (1) the horizontal phase corresponding to the closed phase of the glottal cycle; (2) the rising phase corresponding to the opening; and (3) the falling phase, which often is the steepest, corresponding to the closing phase. The variability of a glottogram is restricted for physiological reasons. For instance, the closed phase may occupy between zero and 70% of the total cycle time, approximately. Corresponding values for the opening phase are 30 and 50%. The closing phase is rarely much shorter than 0.5 msec. Because of this limited variability in the glottogram shape, we may expect reasonably simple interrelationships between glottograms and the associated spectra, i.e., the source spectra.

Recently, the present authors demonstrated how the glottogram amplitude relates to the amplitude of the voice source spectrum fundamental (Sundberg & Gauffin, 1979). The relationship is shown by the data points in Fig. IV-A-2. The solid line refers to the case of

Fig. IV-A-1. Idealized glottal volume velocity waveform, i.e., a glottogram and the differentiated waveform.

Fig. IV-A-2. Peak amplitude of the fundamental as function of glottogram (peak-to-peak) amplitude in four voices phonating at various pitches, degrees of vocal effort, and in different types of phonation.
a purely sinusoidal glottogram. The datapoints pertain to different subjects, phonating at different pitches and vocal efforts and in differing modes of phonation. We see that the scatter of the datapoints is greater at low glottogram amplitudes and that the dependence approaches a one-to-one relationship for greater glottogram amplitudes. The figure also shows that the amplitude of the fundamental increases more rapidly than the glottogram amplitude. In the same investigation we also showed that the glottogram amplitude is strongly dependent on the mode of phonation. If a high subglottic pressure is combined with a high degree of adduction, i.e., approximation of the vocal folds, the glottal amplitude is low. We call this type of phonation "pressed". If subglottic pressure as well as adduction activity are reduced, phonation turns into what we call "flow phonation". It is characterized by a high glottogram amplitude.

The sound pressure level, or the SPL, depends on the amplitude of the harmonics closest to the first formant in normal phonation. The main excitation of the first formant takes place at the instant of glottal closure in the glottal cycle. Given the general shape of a glottogram, we realize that the maximum rate of change in the volume velocity during a glottal cycle would be related to the amplitudes of the spectral overtones. In other words, the maximum amplitude of the differentiated glottogram (cf. Fig. IV-A-1) must be of importance to the source spectrum overtones. These relationships have recently been described in mathematical terms by Fant (1979). The present investigation is a pilot study of the voice source variability in a couple of subjects during vowel phonation. Glottogram amplitude, peak amplitude of the differentiated glottogram, and, particularly, the relationship between the sound spectrum and the amplitude of this derivative are examined.

Experiment

In one experiment glottograms were collected from male singers and nonsingers by means of an inverse-filtering equipment ad modum Rothenberg (1973). The subjects phonated sustained [ae]-vowels with
varied vocal effort at constant pitches. The resulting glottograms were recorded by means of a computer.

Since these measurements suggested a more detailed study of the relationships between vocal effort, spectrum, and the peak amplitude of the differentiated glottogram, a second experiment was carried out. In this experiment a singer and a nonsinger phonated an [æ]-vowel at constant pitch and with continuously changed vocal effort. The sound was recorded on one track of a tape recorder in an anechoic room. On the second track, the differentiated glottogram was recorded. By using the signal from an ordinary microphone having flat frequency characteristics and no phase distortion in the frequency range of interest as the input to the inverse-filter, the differentiated glottogram was obtained directly as the inverse-filter output.

The subsequent analysis was straight-forward. Using a level recorder, we determined the amplitudes in various spectral bands from the output of a calibrated multifilter. In addition, we also measured the overall SPL and the peak amplitude of the differentiated glottogram. Because of the characteristics of the filter and of the level recorder, the accuracy of measurement was approximately ±1.5 dB.

Results

Fig. IV-A-3 shows data collected from the first mentioned experiment, namely the glottogram amplitude as function of the SPL in a singer and a nonsinger. The symbols represent pitches. A marked difference between these subjects is that glottogram amplitude increases almost at the same rate as the SPL in the singer, while in the nonsinger a 10 dB-increase in SPL leads to only a 4 dB-increase of the glottogram peak amplitude, approximately. This appeared to be a typical difference between singers and nonsingers.

Fig. IV-A-4 compares a singer and a nonsinger with respect to the peak amplitude of the differentiated glottogram as function of pitch. The parameter is vocal effort. The derivative is seen to reach higher values and varies within a wider range at extreme fundamental frequencies in the case of the singer.
Fig. IV-A-3. Glottogram amplitude as function of SPL for a singer (left) and a non-singer (right). The solid line represents the case that the glottogram amplitude increases at the same rate as the SPL. The slope of the dotted line in the right graph shows the case that a 10 dB increase in SPL is accomplished by a 4 dB increase in the level of the fundamental.

Fig. IV-A-4. Peak amplitude of the differentiated glottogram, \( \frac{dU_0}{dt} \), as function of phonation frequency for a singer and a non-singer.
Let us now examine the spectral consequences of these glotto-gram parameter variations. Fig. IV-A-5 shows the levels in different spectrum bands as function of their center frequencies for two vowel sounds. The bands were chosen such that the levels could be determined for the fundamental, the first formant, the second formant, and higher formants, respectively. The parameter is SPL. The data pertain to the same two voices as in the last figure. At SPL:s over 70 dB, only the singer increases the amplitude of his fundamental at almost the same rate as the SPL, cf Fig. IV-A-3, while at lower SPL:s, the fundamental changes much less than the SPL in both voices. We also notice, that the higher overtones increase in amplitude more rapidly than the lower overtones, as vocal effort is raised. Finally, the singer is seen to have a higher spectral level in the highest spectrum band, in which the "singer's formant" is located.

Fig. IV-A-6 shows these same data plotted as function of the SPL. Because SPL is generally determined by the spectral components around the first formant, the corresponding spectrum band shows values lying close to the identity line in the plot in most cases. However, in very soft phonation the strongest spectral component is the fundamental.

Let us now relate the SPL to the peak amplitude of the differentiated glottogram. In doing this, a compensation should be made taking into account the fact that the spectrum level depends not only on the source spectrum, but also on the frequencies of the formants and of the fundamental. This correction, required in order to make the spectral levels in differing vowels comparable, was determined by synthesizing vowels with the same formant and fundamental frequencies as in the analyzed vowel sounds. Fig. IV-A-7a and IV-A-7b show the results. First we can compare the data pertaining to various fundamental frequencies in the singer and the nonsinger. We see that the data points lie fairly close together. Note that these data points stem from vowel spectra in which the SPL is carried by the first formant in most cases but by the fundamental in some cases.
Fig. IV-A-5. Sound level of the fundamental first formant, second formant, and higher formants for a singer and a nonsinger. The SPL is the parameter.

Fig. IV-A-6. Spectral levels in four different frequency bands as function of the SPL.
Still, there is no marked scatter in the data points even in such phonations, which are intermediate in this respect. This holds for both voices. If we pool the data, the scatter does not increase to any great extent. The same is true even if we add the data points pertaining to pressed and breathy (Fig. IV-A-7c) or modal and falsetto register phonation (Fig. IV-A-7d). We arrive at the conclusion that the SPL can be predicted from the differentiated glottogram within +/- 3 dB for these speakers. The overall trend of the data points is an S-shaped curve having the slope of one, approximately, for medium amplitudes of the differentiated glottogram. The linear regression coefficient for the entire material is 0.972.

Discussion

We want to make three comments on the results. The first is that, basically, our results can be explained by using the mathematical model of glottogram waveforms proposed by Fant (1979).

Second, we would like to stress the following. As long as the closing time in the glottal cycle can be shortened, the amplitude of the derivative may be raised, even if the glottogram amplitude is kept constant. When the closing time has been shortened as much as physiological and acoustical limitations allow, the amplitude of the derivative may still be increased, but only by increasing the glottogram amplitude. A major difference between the singer and the nonsinger was found to be that the dynamics of the glottogram amplitude was much greater in the singer. This suggests that it is only the singer who takes full advantage of the glottogram amplitude in order to increase the amplitude of the derivative and hence SPL. The reason for this would be that the nonsinger tends to use a more pressed phonation as vocal effort is raised beyond a certain point. Under these conditions, the glottogram amplitude may even decrease instead of increase with rising vocal effort. We conclude that the singer seems to adopt a wiser strategy by avoiding pressed phonation even at high degrees of vocal effort: pressed phonation is likely to be harmful to the vocal folds and high SPL-values are needed in singing.
SPL as function of peak amplitude of the differentiated glottogram, \(dU_G/dt\). The SPL value for each data points is compensated for changes in formant and fundamental frequencies.
Third, we want to point out that our investigation is based on a detailed study of the voice source in two subjects only, sustaining one vowel. Still, we believe that our results possess generalizability at certain points. The clear relationship between the amplitude of the differentiated glottogram and the SPL of the radiated sound would depend on basic physiological constraints on the vibrating vocal folds. This is probably the reason for the limited scatter in the plots showing this dependence under widely varied phonatory conditions. We believe that our results have a general validity in this respect. With regard to the level of the high formants (including the "singer's formant") the dependency of the amplitude on the differentiated glottogram should be examined in greater detail.

Conclusions

1. Given the frequencies of the formants and the fundamental, the glottogram amplitude and the peak amplitude of the differentiated glottogram seem to determine with an accuracy of ca +/- 3 dB: (a) the amplitude of the source spectrum fundamental and (b) the SPL of the radiated vowel, respectively.

2. Singers seem to vary their glottogram amplitude within a wider range than nonsingers as vocal effort is changed. The result is that the singers get access to a greater dynamic range, particularly at extreme fundamental frequencies: by avoiding pressed phonation even in loud tones, they can increase the glottogram derivative amplitude even when the glottal closing time cannot be shortened any further.

3. The mathematical description of glottal waveforms proposed by Fant (1979) can be used to explain our results.

Acknowledgments

The authors are indebted to Martin Rothenberg and Gunnar Fant for fruitful discussions. The present work was supported by the following funds: the Swedish Council for Planning and Coordination of Research, the Swedish Council for Research in the Humanities and Social Sciences, the Swedish Natural Science Research Council, the Bank of Sweden Tercentenary Foundation, and the Swedish Medical Research Council (contract no. B80-17X04783-5).

References

