An attempt to improve the clinical usefulness in phonetograms

Gramming, P. and Gauffin, J. and Sundberg, J.

journal: STL-QPSR
volume: 26
number: 2-3
year: 1985
pages: 077-085

http://www.speech.kth.se/qpsr
B. AN ATTEMPT TO IMPROVE THE CLINICAL USEFULNESS OF PHONETOGRAMS*

Patricia Oramund**, Jan Gauffin, and Johan Sundberg

Abstract

Phonetograms, i.e., graphs showing SPL versus fundamental frequency of voice in soft and loud phonation have frequently been used in voice clinics as an aid for describing voice function. However, a phonetogram does not offer an exhaustive picture about the voice status. For instance, the same SPL value can be obtained using different modes of phonation. This suggests the need for complementing phonetograms with some information reflecting mode of phonation.

In previous research it has been demonstrated that the amplitude of the fundamental may vary within wide limits depending on the mode of phonation. This pilot study presents an attempt to explore the information offered by the amplitude of the fundamental (Lp).

The variations in Lp related to changes of phonatory conditions are mapped, and typical patterns are revealed. Two patients suffering from phonasthenic dysphonia strained their voices by 30 minutes of loud reading. Phonetograms displayed decreased SPL values in loud phonation after loud reading but essentially unchanged SPL values in soft phonation, thus suggesting improvement of voice dynamics. However, after loud reading the patients also showed a trend to a smaller range of Lp variation, which suggests a change towards a pressed/strained/tense mode of phonation.

Introduction

The use of so-called phonetograms has been repeatedly advocated in clinical voice analysis (see, e.g., Calvet, 1953; Coleman & al., 1977; Schutte, 1980a; Schutte & Seidner, 1983). The phonetographic method has certain important advantages: (1) phonetograms are easy to establish, particularly using computational methods (Bloothooft, 1962), (2) phonetograms take into account the dynamic aspects of voice, and (3) the information contained in a phonetogram informs about voice source characteristics. In clinical voice research a considerable amount of work has been devoted to the development of the phonetographic method. Several attempts have been made to capture in quantitative terms the information contained in phonetograms (Coleman & al., 1977; Komiyama, & al., 1984; Klingholtz & Martin, 1983; Schutte, 1980b).


**Department of Phoniatrics, ENT Clinic, Malmö General Hospital
On the other hand, phonetograms have important limitations. Some authors (Heinemann & Gebriel; 1982) have observed that an obvious clinical voice improvement is not always reflected in phonetograms, and various ways of complementing phonetograms have been proposed, e.g., with measures of waveform perturbation (Heinemann & Gabriel, op.cit; Pabon, 1984), spectral energy at high frequencies (Seidner & al., 1981), or airflow (Komiyama & al., 1982).

The ultimate goal of a clinical voice analysis is an efficient mapping of the physiology of vocal-fold vibration. However, the relationship between the voice source and the SPL values shown in the phonetogram will be complicated as soon as the patient changes articulation with pitch and/or intensity; such articulatory changes, which may be made quite unconsciously, affect the formant frequencies, and a change of the first formant frequency normally affects the SPL even under conditions of constant mode of phonation.

Another complication is that in soft phonation, the SPL value is mainly dependent on the amplitude of the voice-source fundamental. In loud phonation, on the other hand, voice-source overtones determine the SPL value (Gaffin & Sundberg, 1980). Important information about the voice source is reflected in the amplitude relationships between L0 and the higherpartials. For instance, the peak amplitude of the acoustic glottogram, i.e., the transglottal airflow waveform, is closely related to the amplitude of the source-spectrum fundamental and, hence, L0. L0 is decreased when the mode of phonation is shifted between the extremes of "breathy" and "pressed/strained/tense" phonation. The amplitude of the first formant, on the other hand, normally depends on the steepness of the closing phase of the acoustic glottogram, or, more precisely, the peak amplitude of the differentiated glottogram. Thus, the amplitude of the voice-source fundamental reflects a phonatory dimension which is quite separate from the dimension reflected in the amplitudes of the source-spectrum overtones. This implies that different parts of the phonetogram contour may refer to quite different voice-source characteristics between which phonetograms fail to differentiate.

A question of primary concern is how the SPL value is related to the characteristics of the glottal voice source. Fig. 1 shows this relationship for an ideal case, where the first formant frequency is high (600 Hz) and where all formant frequencies as well as all voice-source parameters remain constant regardless of fundamental frequency. Thus, the figure illustrates the SPL and L0 values that would result if the voice-source waveform was kept constant throughout the fundamental frequency range.

The general trend is that SPL increases with rising fundamental frequency. However, as demonstrated by the model experiment, this increase is not completely smooth; when the frequency of a harmonic partial equals the first formant frequency, the SPL reaches a peak. In the figure, such peaks can be observed at 600/1 Hz, 600/2=300 Hz, 600/3=200 Hz. In natural speech, these variations are somewhat smaller (see Pant & al., 1963). In the same Fig. 1, the corresponding values of the ampli-
tude of the fundamental, $L_0$, are plotted. As expected, the $L_0$ curve rises monotonically with fundamental frequency up to 600 Hz, i.e., to the frequency where the fundamental equals that of the first formant. The figure also illustrates that the difference between the SPL and the $L_0$ decreases with rising fundamental frequency, and become almost identical when the fundamental is close to the first formant.

Summarizing, it is apparent that the information contained in a phonetogram on the voice source is obscured by a number of different factors which normally are not controlled. In other words, the same SPL value can be produced with different modes of phonation. A more qualified description of voice function can be expected if the phonetogram is complemented with data on the amplitude of the fundamental. The present pilot study is a first attempt to explore this parameter as a complement to the phonetogram for the purpose of voice status description.

**Experiment**

Two experiments were carried out. In the first experiment, a male and a female subject were used. Both subjects had normal voices and training in singing. The subjects first sustained the vowel /a/ as softly and loudly as possible at various fundamental frequencies. Then, replicating the same SPL-values, they repeated the experiment, deliberately using a breathy voice in soft phonation and a pressed voice in loud phonation. This experiment was carried out in an anechoic room.

In the second experiment a male and a female patient, which both suffered from phonastenic dysphonia, served as subjects. First, they sustained the vowel /a/ as softly and loudly as they could on different fundamental frequencies. Then, they read a text during 30 min trying to keep a vocal intensity of about 80 dB SPL while exposed to a 70 dB SPL noise presented in ear phones, and thereafter the same experiment was repeated (see Kitzing, 1979). These recordings were made in a sound treated room, about 4x3x2.3 m.

The analysis was straight-forward. The tape recorder output was connected to an Ithaco filter and the resulting signal was recorded by a B&K 2307 level recorder. The filter was set to 10 kHz for the SPL measurements, and at about 1.5xFo for determining the amplitude of the fundamental.

**Results and Discussion**

Fig. 2 shows the SPL and $L_0$ values observed for the two subjects phonating with their normal voices. The dynamics varies between about 15 and 20 dB in the male subject, and in the female subject the corresponding value is 10 and 25 dB.

In loud phonation, the increase of SPL with rising pitch is much greater in the female voice than in the male voice. With rising pitch the amplitude of the fundamental increases slightly more than the SPL in
Fig. 1. Readings of SPL and L₀ obtained from a formant synthesizer excited with a source having a constant triangular waveform corresponding to a spectrum envelope fall of -12 dB/octave throughout the fundamental frequency range.

Fig. 2. Phonetograms of two normal voices, complemented with the associated L₀ values.
Fig. 3. $L_0$ values of two normal voices sustaining vowels loudly with normal and deliberately pressed phonation (upper series) and softly with normal and deliberately breathy phonation (lower series).
both subjects. In the female voice the difference between SPL and \( L_0 \) is much smaller than in the male voice, so that at the top pitch the SPL equals \( L_0 \). This simply reflects the fact that the fundamental comes much closer to the first formant in the female voice.

In soft phonation the difference between SPL and \( L_0 \) is smaller in both voices, and in the female voice the SPL is almost equal to the \( L_0 \) suggesting a rather sinusoidal acoustic glottogram. The increase of \( L_0 \) with rising pitch is almost exactly the same as was observed in the formant synthesizer, which indicates that both subjects kept the amplitude of their source-spectrum fundamental approximately constant in soft phonation.

Fig. 3 compares the \( L_0 \) values in breathy, normal, and pressed phonation. It should be recalled that these comparisons pertain to phonations having identical SPL values. The graphs illustrate that as compared to normal phonation, \( L_0 \) is reduced in pressed phonation and increased in breathy phonation.

Fig. 4 presents the phonetogram observed for the patients before and after the 30 min session of loud reading. The slight but consistent trend is that the level of maximum voice intensity increased somewhat after the exercise, while no clear differences can be seen in soft phonation. Thus, the exercise slightly expanded the area between the phonetogram curves, and this is generally considered as a sign of voice improvement (see, e.g., Schutte, 1980a).

The gain in \( L_0 \) accompanying an increase of vocal intensity has been found to differ between a singer and a nonsinger (Gauffin & Sundberg, 1980). While, according to Fant (1960), a 10-dB increase of voice intensity generally is associated with an \( L_0 \) increase of 4 dB in untrained voices, Gauffin & Sundberg found the \( L_0 \) to increase at approximately the same rate as the SPL in a singer. Presumably this reflects the fact that nonsingers tend to change their mode of phonation towards pressed phonation when they increase voice intensity, while singers would attempt to avoid such pitch-dependent changes of phonation mode. Voice therapy may be expected to bring about a similar difference. In any event the increase of \( L_0 \), associated with an increase of SPL, seems a relevant voice characteristic to analyze.

Fig. 5 shows these data for the two subjects and for the two patients. The data from the normal voices illustrate how a change of phonation mode towards pressed phonation affects the plots. It can be seen that the \( L_0 \) difference, associated with a given SPL change, becomes smaller if soft phonation is breathy and loud phonation is pressed. Both patients show a different distribution of data points before and after the loud-reading session. The female patient in particular displays a trend similar to that of the normal subjects. In other words, the trend is that a given increase of the SPL is accompanied by a smaller increase in \( L_0 \) after than before the loud-reading session. This suggests that the patients' loud phonation became more pressed after the loud-reading session.
Fig. 4. Phonetograms of the voices of two patients suffering from phonasthenic dysphonia recorded before and after (dashed and solid lines, respectively) 30 minutes of loud reading.

Fig. 5. Difference in amplitude of the fundamental as function of the SPL difference observed when loudness of phonation was increased while the fundamental frequency was kept constant. The two upper graphs refer to two normal subjects and the two lower graphs refer to two suffering from phonasthenic dysphonia. In the case of the normal subjects, the open circles show the SPL and $L_0$ differences observed between normal soft and normal loud phonation, while filled circles compare breathy soft phonation and loud pressed phonation. In the case of the patients, open and filled circles show values observed before and after 30 minutes of loud reading.
The above has demonstrated that $L_0$ values represent a valuable complement of the phonetogram, as it, unlike the phonetogram, informs about the mode of phonation. This does not mean that a phonetogram complemented by $L_0$ data would offer an exhaustive description of voice function. Probably, a better description would be obtained by using separate frequency bands for measuring (1) $L_0$, (2) the level of the first formant, and (3) the spectral level in the frequency band 4-8 kHz. These measures would inform about the peak amplitude of the glottogram, the peak amplitude of the differentiated glottogram, and the glottal leakage, respectively (see Gauffin & Sundberg, 1980; Fant, & al., 1985). Measuring these parameters would probably involve minor problems using modern microcomputer technology.

**Conclusions**

The informative power of a phonetogram regarding voice function is limited, as the same SPL value can be produced using differing modes of phonation. Also, there is no simple relationship between the voice-source characteristics and the SPL values. For instance, different parts of the phonetogram contour often refer to quite different voice-source characteristics. Also, pitch and/or intensity-dependent changes of articulation complicate the relationship between SPL and voice source. The $L_0$ seems a worthwhile complement of the phonetogram, as it informs about mode of phonation. In particular the gain in $L_0$ accompanying an increase of SPL seems an interesting measure.

**Acknowledgements**

The authors are indebted to Docent Peter Kitzing for encouraging this research as well as for his valuable suggestions on this manuscript.

**References**


