

# VOCAL REGISTERS REVISITED

*Gerrit Bloothoof and Peter Pabon*

Utrecht Institute of Linguistics - OTS, Utrecht University  
Trans 10, 3512JK Utrecht, The Netherlands  
{Gerrit.Bloothoof,Peter.Pabon}@let.uu.nl

## ABSTRACT

The paper concerns the quest to find acoustical parameters that describe properties of vocal fold functioning from professional singing to voice pathologies. Improved data handling within the framework of phonetogram recordings has been used to arrive at a better understanding of what acoustical parameters could tell us about vocal registers.

Keywords: voice analysis, vocal registers, phonetogram recordings, acoustic parameters

## 1. INTRODUCTION

For over a century the scientific definition of a vocal register has been related to the same mechanical principle of vocal fold vibration, with modal register and falsetto register as the dominant modes. Although there is no reason to dispute this view there remains a slight friction with everyday terminology, where a variety of terms like chest voice, head voice, and middle register seem to be inextinctable. Besides an understanding of these terms from acoustically irrelevant secondary resonance sensations in the singer's body, they may also point to more or less distinct timbre qualities that subdivide the modal and falsetto registers from an acoustical point of view. We have studied this possibility by means of an improved analysis of phonetogram registrations of 12 male non-singer subjects.

## 2. THE PHONETOGRAM

The phonetogram is a display of acoustic voice parameters in a diagram which has fundamental frequency (F0) on the horizontal axis and vocal intensity (I) on the vertical axis. In its simplest form the phonetogram only shows which combinations of F0 and I a subject is able to sing. This normally constitutes an oblique, egg-shaped area. In addition to F0 and I, in computerised recordings it is possible to measure acoustic voice parameters which reveal details of the quality of

phonation. Their values can be shown in the phonetogram by means of a color scale.

The advantage of a phonetogram recording is that it provides a *complete* overview of vocal possibilities. Acoustic voice qualities can be seen in a glance for *all* possible combinations of fundamental frequency and vocal intensity, which facilitates the interpretation of results.

The central issue in phonetogram recording is the choice of additional acoustical parameters, their relation to physiological characteristics of the voice, and the way to display their values. Up to now, the display and storage of data was limited to a single value of an acoustical variable per F0-I combination for technical reasons. This could be the average, the best or the latest value. This serious limitation has been overcome now and in the latest version of phonetogram recording it is possible to store all raw parameter values. The raw values allow for a more sophisticated statistical processing and a more lucid post-recording presentation of important acoustic properties in the phonetogram.

Besides fundamental frequency (F0, represented in semi-tones) and vocal intensity (I, RMS value in decibels), we compute the stability of F0 (jitter, in dB), the difference in maximum peak energy and RMS energy (crest factor, in dB), and the time to reach the maximum peak relative to the period duration (relative rise time, as percentage). These acoustic parameter values are computed and sampled at a rate of 76 Hz.

## 3. MEASUREMENTS

Two phonetogram recordings on the vowel /a/ were made in a sound treated booth for each of 12 male non-singer subjects. A microphone (B&K 2032) was connected to a headset to keep the microphone distance constant (25 cm). Subjects had visual feedback from a screen with F0 on the horizontal and vocal intensity on the vertical axis. A colour scale displayed vocal density at every combination of F0 and I. Each phonetogram took between 20-30 minutes to complete. This resulted

on average in a total of about 120.000 raw samples per subject.

The raw data were processed to exclude (too) unstable phonations and to exclude computation errors such as octave errors in F0 estimations. About 30% of all raw data was rejected in this way.

To exclude the influence of occasional outliers in the data, statistical processing involved the computation of the *median* value of acoustical parameters as a function of F0 and I. The elementary unit for summation of data was one semitone by one decibel (see Fig. 1, upper panel). This implies around 1000 units for the total phonetogram (roughly 36 semitones x 30 decibel). If a unit had more than 0.5% of the total number of samples this was considered to be sufficient for direct median computation. If this was not the case, continuity was assumed in the acoustic characteristics in the phonetogram and data from surrounding units (upto 16 units) were added until the threshold was reached. If even this was not sufficient the threshold was lowered to 0.1%.

#### 4. RESULTS AND DISCUSSION

Figure 1 shows the improvement of data representation by the computation of the median value. In stead of a granular representation (upper panel) a smooth display is the result (lower panel). This makes it much easier to identify regions in the phonetogram with comparable properties.

To get an overall impression of the properties of the three acoustical voice variables we summed all data per F0-I unit across *all* subjects without any normalisation. Figure 2 gives the result for the jitter variable. The most unstable (darkest) region is found for soft phonations in the low F0 range, with less pronounced extensions to soft phonations in the high F0 range. For the crest factor Fig. 3 shows the highest values in the lower half of the F0 range for high intensities (pulse-like waveform with a corresponding flat spectrum), and lower values (close to sine-like waveforms) at the high F0 region and at lower vocal intensities. A quite abrupt change in the crest factor can be seen that roughly divides the phonetogram in a lower and upper half. Finally, Fig. 4 gives the values for the relative rise time. This parameter was less stable, but shows a pattern resembling that of the crest factor. Extremely short rise times are found at maximum vocal intensity, probably indicating a pressed phonation type.

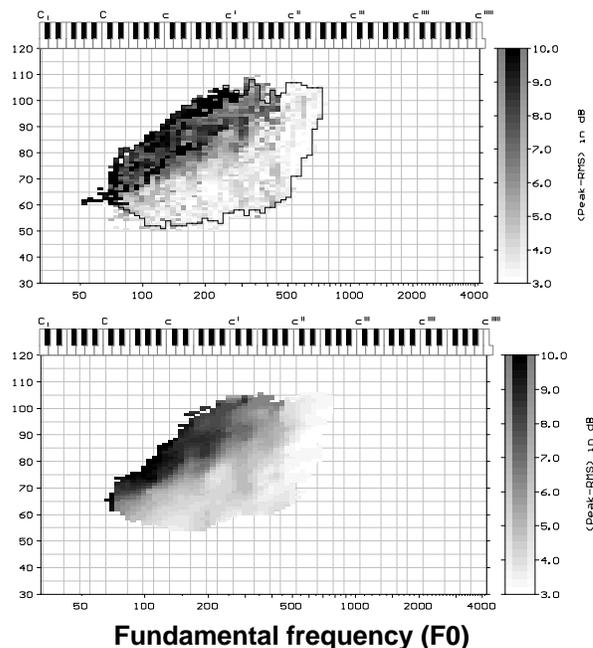


Figure 1. Phonetogram with crest factor display. The upper panel shows a representation with the last measured crest value, the lower panel shows the result after median smoothing.

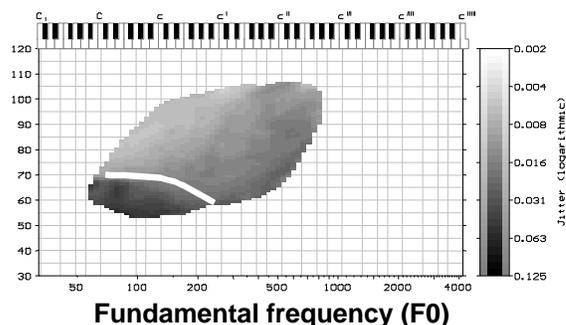


Figure 2. Phonetogram (summed over all subjects) with jitter display. The darker the area the less stable the phonation. The white line marks the most unstable region.

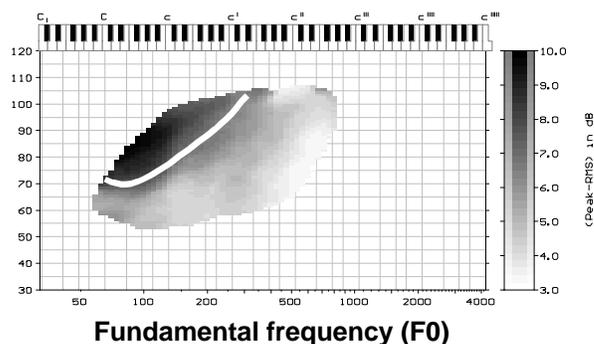
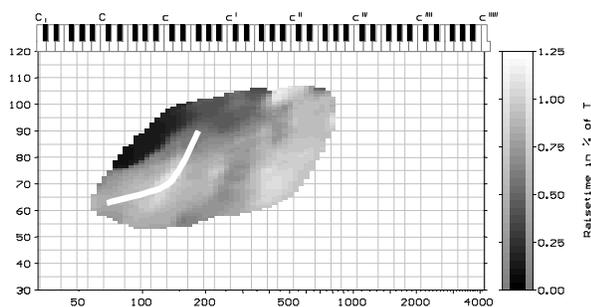


Figure 3. Phonetogram (summed over all subjects) with crest factor display. The darker the area the more prominent the harmonics in the signal. The white line roughly separates sine-like phonations and phonations rich of harmonics.



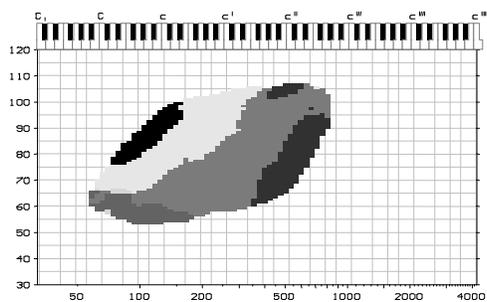
### Fundamental frequency (F0)

Figure 4. Phonetogram (summed over all subjects) with display of the **relative rise time**. The darker the area the more abrupt the beginning of the voice period. The white line roughly marks the close to pressed phonations.

Instead of three phonetograms, each displaying an acoustic variable, we designed a presentation that simultaneously shows the most salient aspects of the three variables. The aspects each get their own grey value, and they are:

- Crest factor in three ranges: (1) 3-4 dB, indicating sine-like phonations (dark grey); (2) 4-6 dB, intermediate (light grey); (3) >6 dB indicating a signal rich of harmonics (white)
- Jitter >3%, indicating very unstable phonations (darkened relative to crest factor grey-value)
- Relative rise time <6%, indicating pressed phonations (black)

Figure 5 gives this presentation for the summed data.

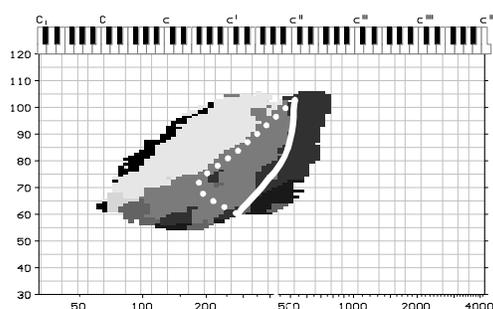


### Fundamental frequency (F0)

Figure 5. Phonetogram (summed over all subjects) with display of acoustically characteristic areas. From upper left to lower right: black indicates near to pressed phonation, white indicates rich harmonics, light grey has intermediate harmonics and dark grey are sine-

Figure 5 shows that the crest factor representation dominates the sub-areas in the phonetogram. The iso-crest line of 6 dB that subdivides the phonetogram in a lower and upper half is most prominent. In addition, Fig. 3 indicated that the change in crest factor around the iso-crest line of 6 dB is relatively large. Since the crest factor is strongly related to spectral slope, this goes together with a considerable perceptual change. In our view this may be an additional explanation for the notion of chest voice. Although chest voice is not a vocal register in the mechanical sense, a sudden change in timbre may well demarcate a perceptual boundary. The physical nature of this phenomenon is still to be found.

Separate phonetograms for the twelve subjects are given in Fig. 6. Apart from inter-individual differences the overall representation is roughly the same. In addition, there seems to be an interesting role for the iso-crest line of 4 dB at high F0. For one subject, Fig. 7 gives the boundaries of the modal and falsetto register as indicated by the subject. The iso-crest line of 4 dB marks the upper boundary of modal register in this case. The lower boundary of the falsetto register does not have its counterpart in the acoustical measures used. Interestingly, the iso-crest lines of 4 dB and 6 dB coincide at the transition of modal to falsetto register at the phonetogram contour. These relations are not systematically found and investigated yet for all subjects. Still, the new data handling in the phonetogram promises to bring new opportunities to bridge the gap between acoustic measurements and physiological findings.



### Fundamental frequency (F0)

Figure 7. Phonetogram of a single subject with register boundaries. The solid white line is the upper boundary of modal register; the dotted white line is the lower boundary of falsetto register.

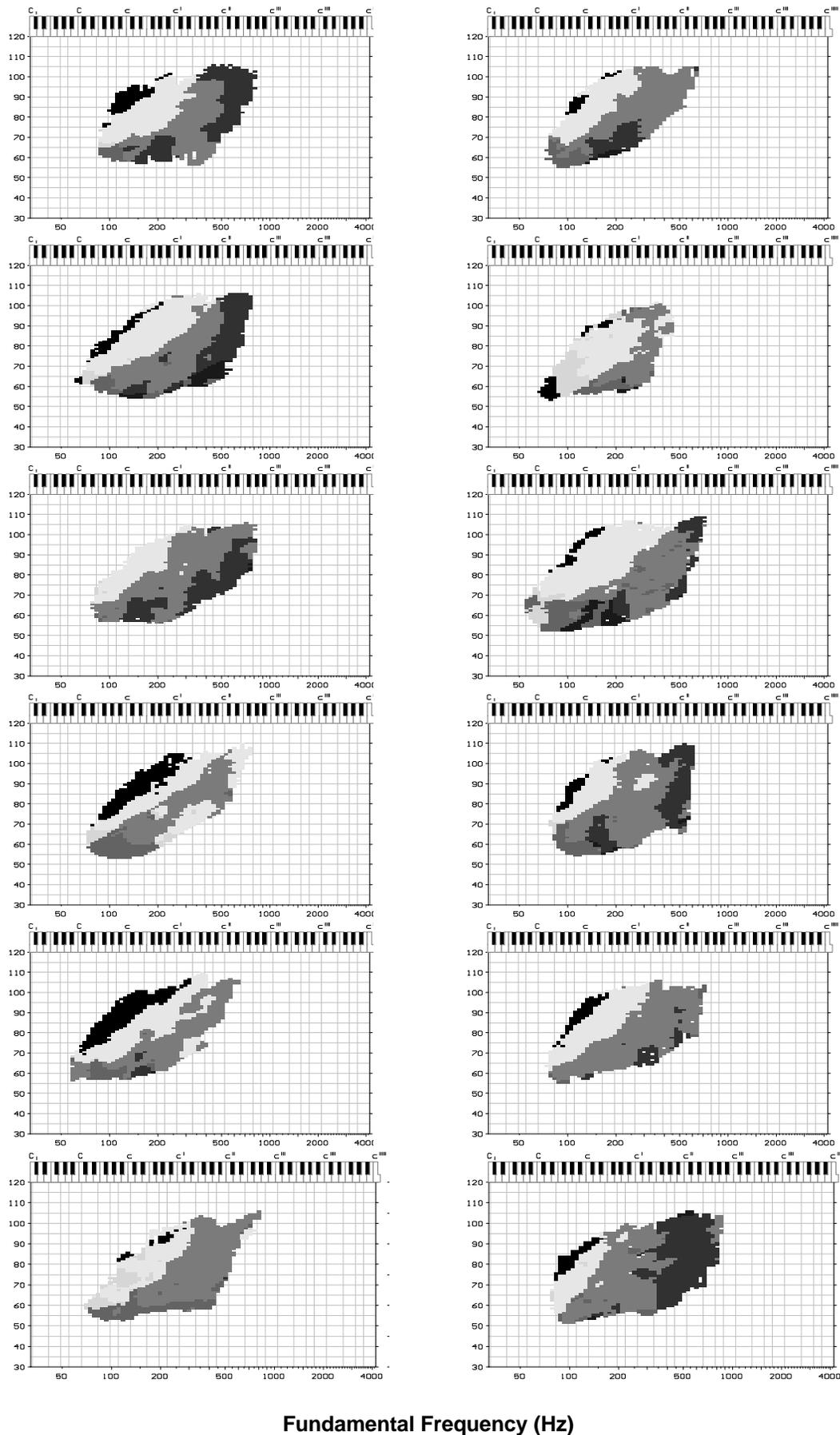


Figure 6. Phonetograms of twelve non-singer subjects displaying major acoustic voice features. From upper left to lower right: black indicates near to pressed phonation, white indicates rich harmonics, light grey has intermediate harmonics and dark grey are sine-like phonations. Relatively dark shades in the lower left corner indicate unstable phonations.

