STANDARDIZING VOICE RANGE PROFILE MEASUREMENTS (Phonetography / Stimmfeldmessung)

B. Frøkjær-Jensen, Danish Voice Institute, Ellebuen 21, DK-2950 Vedbæk, Denmark.
M. F. Pedersen, F. Pabst, T. Hacki, H. K. Schutte, H. Lindskov Hansen.

ABSTRACT

Approch to specifications for standardized phonetography is suggested. The influence from the frequency response of the phonetograph is discussed in details.

GENERAL PRESENTATION

Phonetography is a method used for registration of the total frequency range and the total dynamic range of the voice. The result is shown in a graph called a phonetogram, a voice area, or better a voice range profile (Stimmfeld in German, courbe vocale in French). The fundamental frequencies of several arbitrarily chosen tones are shown along the horizontal axis, and the corresponding sound level pressures are shown along the vertical axis. In some cases phonetograms have been 'modulated' with information about registers and spectral properties [1, 2].

During the past years, phonetography has proved to be a very useful tool for assessment of voices, normal [3] as well as pathological [4, 5], for assessment of the dynamic and tonal range in speech, and useful for testing singers and their potentialities [6, 7].

In 1981 Schutte and Seidner recommended a "Standardization of Voice Area Measurements/Phonetography" which was presented at a meeting in the Union of European Phoniatricians [8]. Most of these proposals are still valid, but new technical possibilities and the practical experience with phonetography during the past 15 years have revealed the fact that some of these specifications have to be reevaluated.

The basic specifications have not changed since 1981: The microphone distance is standardized to 30 cm, or if another distance is used, the measurements are recalculated so that they reflect the levels at a distance of 30 cm. In most cases the dynamic range is fixed at 80 dB from 40 to 120 dB. The frequency range is normally put at five octaves, and normally both the frequencies and the appurtenant tone names are shown. Lowest and highest frequency vary somewhat, but in general a standard pitch range from 50 to 2000 Hz seems to be in use.

RECOMMENDATIONS

One of the main problems in phonetography has been the problem associated with the frequency weighting of the phonetograph. Some researchers use a linear frequency response or a C weighting, as they claim that only by using that sort of frequency weighting will it be possible to make a fair registration of the lowest tones. Other researchers prefer an A-weighted frequency response. In this way, they claim, it is much easier to get a good signal-to-noise ratio in noisy surroundings, e. g. in a laboratory or in a clinic, as the A-weighted response cuts away most low-frequency disturbances. Furthermore, the A-curve results in phonetograms which correspond visually to the audible sound impression of low tones because of the reduced low-frequency sensitivity of the ear.

Thus, it is clear that a phonetogram which is based upon sine tones will depict the frequency response of the phonetograph, which means that a 50 Hz tone will be damped 30 dB relative to a 1000 Hz tone, and this is not acceptable. But the human voice does not produce sine tones, the human voice has a complex spectrum often with very strong harmonics. The question is then: How does the spectral composition influence the contour of a voice range profile ?

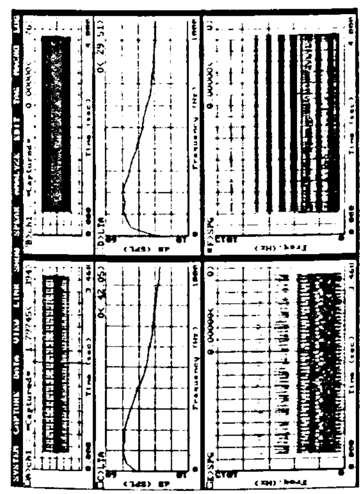

Fig. 1: Acoustic analyses of the 50 Hz tone (left) and the 100 Hz tone (right). Upper curves: Oscillograms, Middle: Power spectra, Lower curve spectrogram.

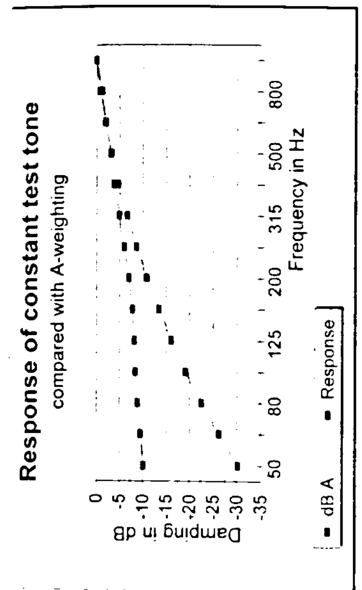

Fig 2: Illustration of the phonetograph's response on a constant input signal with the spectrum shown in Fig. 1 when using A-weighting.

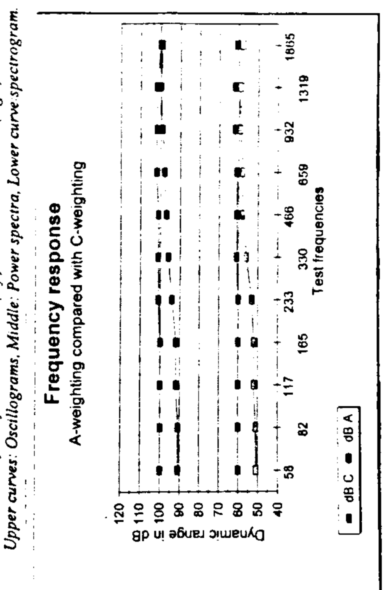

Fig. 3: Two phonetograms made with A-weighting and C-weighting of the tones analysed in Fig. 1. This is a worst case analysis with strong 2nd and 3rd harmonics. A gradual damping of 10 dB is observed in the A-weighting.

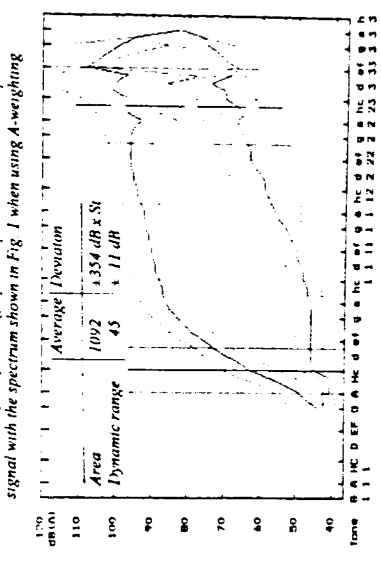

Fig. 4: Phonetograph statistics: Average with standard deviations based upon 18 well-trained logopeds. The vertical lines indicate the lowest and highest average frequencies with standard deviations.

In order to suggest the best recommendation for the frequency weighting in the phonetograph, an investigation of the influence of the A-curve and the C-curve has been carried out:

The instrumental set-up consisted of a phonetograph which was altered so that it was possible to switch between curve A and curve C. The input was generated by a tone generator with tones having a high content of harmonics. A series of tones from 58 to 1865 Hz, with a spectrum which sounded as a soft voice quality, were supplied to a loudspeaker. The sound pressure level at the place of microphone was kept constant at two levels: 100 dB and 60 dB respectively, by means of a simultaneous SPL monitoring with a calibrated Brüel & Kjaer sound pressure level meter type 2236, which had built-in frequency weightings according to curve A, C (IEC 651 type 1), and linear frequency response.

Fig. 1 shows the acoustic analyses of two low-pitched tones at approx. 50 Hz and 100 Hz where the discrepancies are most pronounced. It is noticed that the spectral energy is concentrated in the 4-6 first harmonics (soft voice quality), and that the fundamental frequency and the second (and third) harmonics are of equal level.

Fig. 2 shows the frequency weighting of curve A (the lower line) and the output from the phonetograph (the upper line) with the above-mentioned input signal. It is seen that the signal is not damped 30 dB at 50 Hz relative to 1000 Hz, but only 10 dB. The reason is that it is not the level of the fundamental frequency alone which determines the output level, but the sum of all the harmonics (which are not so strongly damped as is the fundamental frequency).

The influence on a phonetogram of a type A weighting relative to a type C weighting is shown in Fig. 3, which is drawn on the basis of 11 generated tones with 100 dB and 60 dB levels. It is seen that the type C weightings (the black lines) are quite near the 100 dB and 60 dB lines, whereas the A weighting (the grey lines) are reduced with 10 dB at the lowest frequency.

The greatest discrepancy between the real and the measured levels (-10 dB) happens at the lowest frequencies between 50 and 100 Hz, and is more pronounced the stronger the fundamental frequency is. However, as the strongest energy at low frequencies are placed in the first formant (F1 is 250-300 Hz for an [i:] and 700-850 Hz for an [ɑ:], the discrepancy is less pronounced in real speech and singing, especially if the vowel used for the phonetograph recording is an [ɑ:].

The conclusion must be that it is acceptable to use the A weighting for recording of phonetograms, partly because the error made is not serious, and partly because of the great advantage that it is possible to make the recordings in noisy surroundings (up to 30-35 dB SPL (A) noise is acceptable).

Another main problem is the size of the 'open frequency range'. Some persons are not able to reproduce a given tone with the correct frequency. It is very difficult to make valid recordings with tone-deaf subjects. The investigator must have a really good tone perception (absolute pitch perception), if the phonetograms are made manually. In this case the best thing to do is to make an automatic registration of both frequency and level. When recording phonetograms of tone-deaf persones, the phonetograph must have a broad-band 'open frequency range' with at least half-octave filters. Such broad filters are also necessary when making voice range profiles of reading. For singers it is sufficient to set up frequency windows of a semitone each. We would recommend an 'open frequency range' which could be switched between a semitone, a tone, two tones, half octave, and no limitation.

Information on the accuracy of frequency measurements is seldom given by

the various manufacturers of phonetographs, but we would recommend ± 3% as a standard, because it corresponds to the distance between the semitone steps.

It is also an advantage to be able to adjust the recording time. The responses from some types of patients are often very short, and therefore, a short recording time is necessary. On the other hand, singers have often a pronounced frequency and/or intensity vibrato; in such cases it is an advantage that the phonetograph is able to integrate the measurements over a few seconds. Finally, if phonetograms are recorded when reading a text aloud (in order to find the gravity point and distribution of the voice area profile in reading), it is necessary to have a recording time of at least 1 minute. We would recommend a 'record duration' selectable from 0.5 second over 1, 2, 3, 5, 15, 30, to 60 seconds.

The registration of the sound pressure level is integrated and is, therefore, often called sound intensity even if the measurements are specified in the measured dB values. Most researchers and builders of phonetographs use a fast logarithmic RMS measurement, which we also would recommend. The frequency range of the electronics varies in different phonetographs. We would recommend a limited range of 40-15,000 Hz which is sufficient as no voices have much energy above 15,000 Hz. The digital solution is in most instruments set at 0.5 dB before calculation and 1.0 dB after calculation, which we also can recommend as a sufficiently exact measure for practical use.

Some phonetographs have facilities for curve editing. Simple calculations on the recorded phonetograms can also be made. We consider it important that all phonetographs are able to calculate the voice profile area. The unit used must be semitones times decibel. A print-out of a voice range profile must also contain information about the lowest and the highest tone, the pitch range, the weakest and the strongest tone, and the dynamic range. Some phonetographs are equipped with specialities such as simultaneous information in the voice range area of the singer's formant, of jitter, shimmer, or voice quality. It is convenient to have this information, but we do not consider these parameters important for the voice range profile. It is also useful to have an averaging program which can average a series of phonetograms. Fig. 4 shows such an average of 18 good female voices (all trained logopeds). Notice the dotted lines on both sides of the solid curve. These lines indicate the standard deviation on the average. The lowest and the highest frequencies are shown with standard deviations.

Finally, the poster will contain a table which compares the specifications for some of the most commonly used and commercially available phonetographs.

REFERENCES

[1] Bloothooft, G. (1981), Voice profiles, vowel spectra and vocal registers, CoMeT Conference, Amsterdam.

[2] Pabon, J. P. H. (1989), Automatic phonetogram recording supplemented with acoustical voice-quality parameters, *Proc. from Intern. Voice Symp.*, Edinburgh.

[3] Seidner, W. (1985), Objektive Qualitätsbeurteilung der Stimme mittels Dynamikmessungen, *Zeitschr. klin. Medizin 40.*

[4] Hacki, T. (1988), Die Beurteilung der quantitativen Sprechstimmleistung, *Folia Phoniatr. 40*, pp 190-196.

[5] Gramming, P. (1988), *The Phonetogram, an experimental & clinical study,* Malmö, Sweden.

[6] Pedersen, M. F. et al. (1986), Computerized phonetograms for clinical use, *Proc. XXII, IALP Congress,* Tokyo, Japan.

[7] Ohlsson, A.-C. et al. (1986), Phonetograms of normal and pathological voices, *Working Papers in Logopedics & Phoniatrics 3*, Lund, Sweden.

[8] Schutte, H. K. et al. (1983), Recommandation by UEP: Standardization of Voice Area Measurements, *Folia Phoniatrica, 33.*