



## EFFECT OF VOCAL LOUDNESS VARIATION ON THE VOICE SOURCE

J. Sundberg

**Abstract:** The physiological correlate of perceived vocal loudness is the overpressure of air under the glottis, or the subglottal pressure  $P_{sub}$ . Variation of vocal loudness, i.e., of  $P_{sub}$  has strong effects on the waveform of the transglottal airflow, also called the voice source. As the voice source is the primary sound itself, which after filtering by the vocal tract resonator is radiated through the lip opening, variation of  $P_{sub}$  has strong effects on the voice timbre. The waveform of the voice source, called the flow glottogram, can be obtained by inverse filtering, which implies that the vocal sound is filtered by the inverted frequency curve of the vocal tract. A flow glottogram is characterized by triangular air pulses, occurring when the vocal folds open the glottis and allows an airstream to pass. These air pulses are interleaved by episodes of zero airflow occurring when the vocal folds close the glottis, arresting the air stream. Important flow glottogram characteristics are (1) the relative duration of the closed phase, or  $Q_{closed}$ , (2) the peak amplitude of the flow pulse and (3) the maximum flow declination rate corresponding to the steepness of the trailing end of the flow pulse. These voice source characteristics show a reasonably simple relationship to acoustic properties of vocal sounds. The peak-to-peak amplitude of the flow pulse is strongly correlated with the amplitude of the lowest spectrum partial, the fundamental.

The maximum flow declination rate determines the sound level and  $Q_{closed}$  is strongly correlated with the dominance of the fundamental in the spectrum.

When  $P_{sub}$  is increased from very low to low,  $Q_{closed}$  increases markedly, but an increase from a high to a very high  $P_{sub}$  does not affect  $Q_{closed}$  appreciably. An increase of  $P_{sub}$  also leads to an increase of maximum flow declination rate and generally also of the peak-to-peak amplitude. Another consequence of an increased  $P_{sub}$  is that the higher partials in the spectrum gain more in sound level than the lower partials. Thus, a 10 dB increase of the overall sound level of a vowel is typically accompanied by a 15 dB increase of the partials near 300 Hz. This means that the slope of the spectrum, and hence also of a long-term-average spectrum varies with. All these effects of  $P_{sub}$  variation on the voice source imply that comparisons of acoustic spectrum characteristics of a voice, e.g., before and after treatment, must be made for the same degree of vocal loudness. If this condition is not met, the effect of the loudness difference between the recordings compared must be compensated for.