Technology in the Voice Studio

Brian Gill

Associate Professor of Music (Voice)
Jacobs Scholl of Music
Indiana University

Filipa M.B. Lã

Assistant Professor of Voice
Institute of Interdisciplinary Research, Center for Social Sciences, University of Coimbra

Johan Sundberg

Professor of Music Acoustics
KTH (retired)

Outline

Describing things that we can see is quite easy, while describing sound is an almost impossible task. Today we are able to visualize sound in real time. The technological development over the past 40 years has produced a variety of tools that can be useful in today's voice studios.

The purpose of this workshop is to demonstrate six such tools in segments, approximately 35 minutes each:

1. *Respitrack* shows your breathing habits

Inhalation is achieved by opening the airways and increasing the volume of the lungs, either by expanding the ribcage, by contracting the diaphragm, or a combination of these. Exhalation results from compressing the lung volume with open airways, either by compressing the ribcage, or by contracting the abdominal wall, or a combination of both.

These respiratory movements can be visualized in real-time by a *Respitrack* (Johan Stark, email johan.stark@ling.su.se), thus helping students to direct their attention and become aware of their breathing behavior. Elastic belts are placed around the ribcage and around the abdominal wall. The cross-sectional areas encircled by these belts are shown as two curves on the computer screen. A third curve shows the sum of these areas which can be adjusted to reflect how much air there is in the lungs.

The *Respitrack* curves show how inhalation and exhalation are performed. Great differences typically occur between people. For example, during inhalation, some expand the abdominal wall first and then the ribcage; others do the opposite; and many first expand and then contract the abdominal wall while the ribcage expands. A great variability is typically observed also during reading and singing, even among professionals.

NOTES

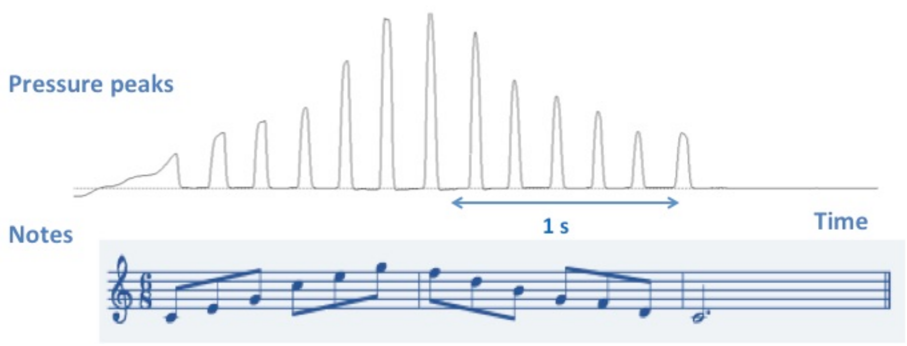
NOTES

2. Take a peek at your subglottal pressure peaks

Subglottal pressure, glottal adduction and length and tension of the vocal folds are the three main physiological control parameters for the voice source, the sound produced by the pulsating glottal airflow during phonation. Subglottal pressure is our tool for varying vocal loudness. However, in singing it also needs to be tailored to pitch: the higher the note, the higher pressure needed.

Subglottal pressure can be measured as the oral pressure during the occlusion for the consonant /p/, captured by a thin plastic tube that is held in the corner of the mouth. The PG-20E, manufactured and sold by the Glottal Enterprises in Syracuse NY, shows this pressure on a display, when the syllable /pa/ is repeated at a rate of about twice per second.

PG-20E also has an exit that can be connected to a display program, such that the subglottal pressure can be visualized in real-time. It is interesting to see how subglottal pressure is varied, e.g., while singing an ascending tonic triad followed by a descending dominant-seventh triad. Typically, a quite beautiful pattern is then produced on the computer screen, see below.



Ascending and descending nine-note *arpeggio*, showing the corresponding intraoral pressure peaks. Note that the highest note in the *arpeggio* does not have the highest-pressure value.

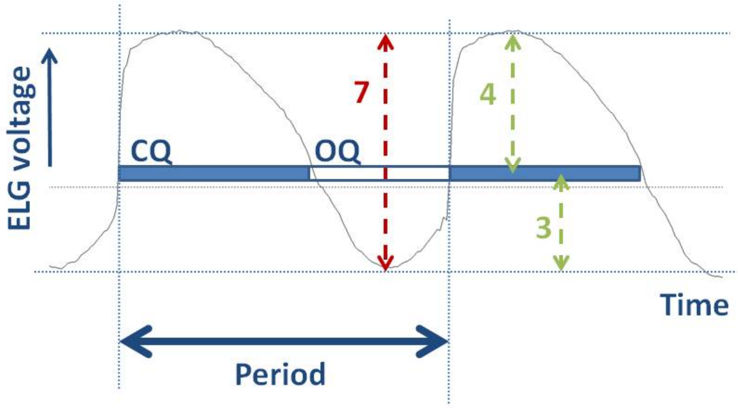
NOTES

NOTES

3. ELG - visualizing vocal fold contact

When the vocal folds contact during vibration, their ability to conduct an electrical voltage increases. This is the basic idea behind electroglottography/electrolaryngography.

Electrode plates are applied on the left and right sides of the thyroid cartilage and an electrical signal between them is displayed on the computer screen. This allows real-time inspection of vocal fold contact and an estimation of the closed phase in the vocal fold vibration cycle. This workshop demonstrates the Electrolaryngograph/ELG (Laryngograph Ltd, London). It displays a measure of the duration of vocal fold contact, $CQ_{Contact}$, expressed as a percentage of the period. This measurement is related, but not identical, to the closed quotient (CQ_{Closed}); the latter can be measured in flow glottograms. Factors affecting $CQ_{Contact}$ are pitch, register, phonation type and vocal loudness.



Speech Studio display of ELG signal (Lx), showing the main phases in one cycle: CQ = contact quotient – time in one vibratory cycle during which the vocal folds stay in contact; OQ = open quotient – time in one vibratory cycle during which the vocal folds stay open; Period = duration of one vibratory cycle.

NOTES

NOTES

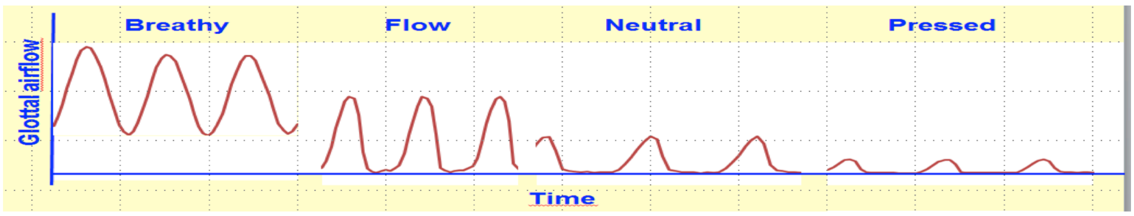
4. Looking at the voice source by inverse filtering

The voice source is the sound generated by the pulsating airflow through the glottis when the vocal folds vibrate. Vocal loudness, pitch and phonation type are all determined by the voice source, which will be displayed in real-time using an inverse filter from Glottal Enterprises, Syracuse, NY.

The underlying idea is to filter the sound captured at the lip opening by a frequency curve that is the upside-down version of frequency curve of the vocal tract. In this way the contributions of the vocal tract to the voice properties can be eliminated, allowing inspection of the voice source.

The voice source is displayed as a flow glottogram, showing glottal airflow versus time. It is characterized by triangular pulses separated by horizontal or nearly horizontal segments.

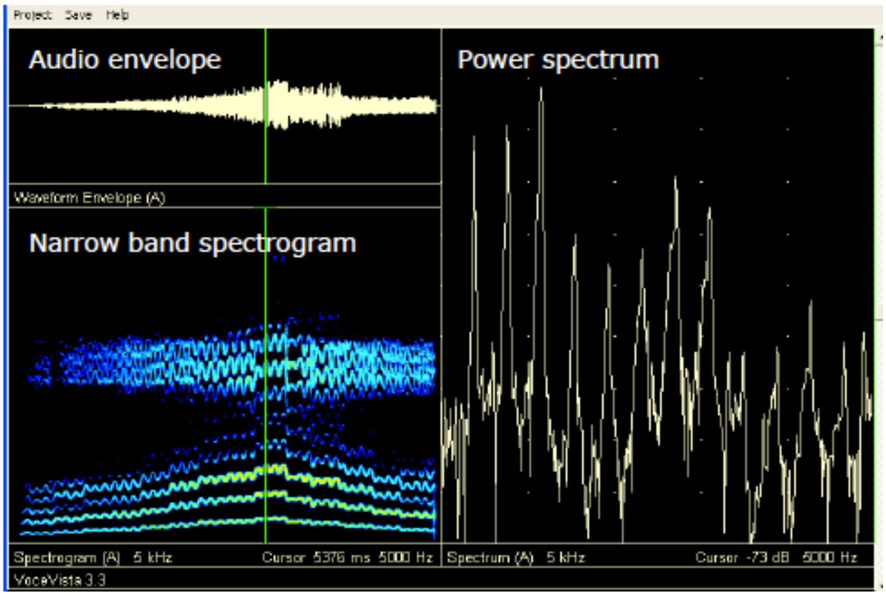
The flow glottogram reflects pitch, vocal loudness and phonation type. Pitch is seen as the separation of the air pulses in time. Vocal loudness corresponds to the steepness of the trailing end of the air pulses. Pressed, Neutral, Flow, Breathy phonation produce quite different flow glottogram characteristics, as shown below.



NOTES

NOTES

5. *VoceVista* - see the sound of your voice



Spectrogram, audio envelope and power spectrum displays available on *VoceVista*.

VoceVista and other spectrograph programs enable us to take an objective look at the acoustic output of a voice. It provides us with a visual of how our formants (resonances of the vocal tract) are interacting with the harmonics arising from the vibrating vocal folds as they modulate the transglottal airflow.

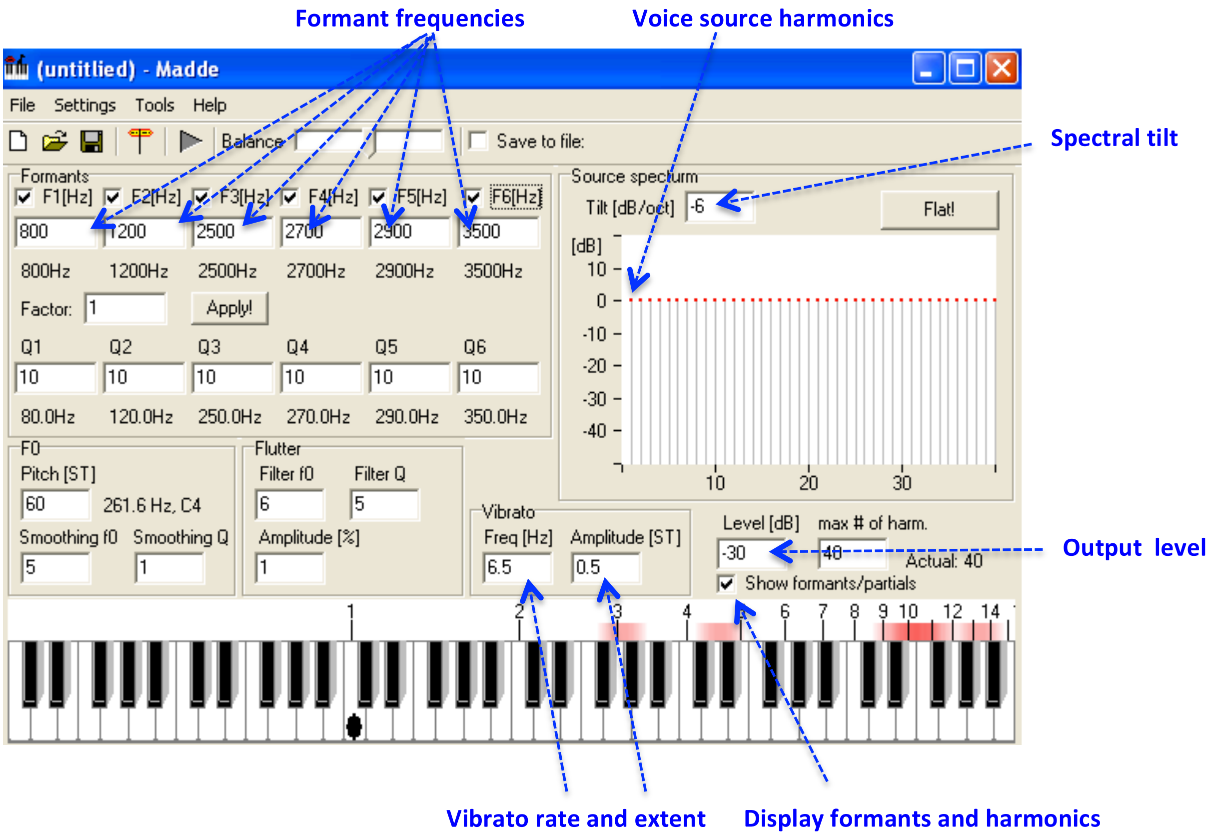
This part of the workshop will help participants develop a deeper understanding of voice acoustics. The details of the variables on each screen will be discussed as well as the basic theory of source-filter interaction. Topics covered: harmonics, partials and overtones; formants; how to measure vibrato rate and extent; and vocal tract tuning/detuning strategies in different styles of music.

NOTES

NOTES

6. How close can a formant synthesizer mimic a voice?

Madde is a voice synthesizer that allows you to test the timbral effects of various voice source and vocal tract properties. One can adjust formant frequencies (1), vibrato rate and extent, the spectral tilt and the amplitude of the individual harmonics of voice source. The frequencies of the formants and the harmonics can be displayed as bars and red markers just above the keyboard. In this workshop segment, we synthesize participants' voices, so as to discover to what extent personal voice quality can be exhaustively described in terms of formants and voice source.



Madde display

NOTES

NOTES