About sine wave, formants, vowels, consonants and overtone singing ...

G2Demo sings overtones
Thanks to sound researchers as Hermann von Helmholtz, Carl Stumpf, Wolfgang Köhler and Karl Erich Schumann. They unraveled the structure of (speech) sounds and thus discovered the essential identity of the vowels and consonants.

The fascination for the human voice is of all times. As early as 1791, Wolfgang von Kempelen writes about a *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*. This Hungarian built a mechanical model of the human speech organ that according to tradition yielded amazing results.

On this link, www.ling.su.se/staff/hartmut/kemplne.htm, you will find a detailed description with pictures of that speaking machine. Of course, sound registration did not exist at that time. But from the thirties of the last century it was possible. The following link, www.icsi.berkeley.edu/eecs225d/klatt.html from the University of Berkeley, summarizes the history of speech synthesis in 36 audio mp-3 examples.

Sine waves, the fundamental building blocks of sound
As early as 1811, Jean-Baptiste Joseph Fourier provided mathematical evidence that any periodic complex vibration can be dissected into a number of basic harmonic vibrations of distinct frequency, amplitude and phase. Fourier's theory came to live in sound experiments by Hermann von Helmholtz. With his Hemholtz synthesizer named after him, he built complex waveforms consisting of ten harmonic tuned partials.

The tuning forks synthesizer
At `http://physics.kenyon.edu/EarlyApparatus/Rudolf_Koenig_Apparatus/Fourier_Synthesis/Fourier_Synthesis.html`, you will find images of this instrument. Ten tuning forks tuned in the harmonic series were continuously kept in vibration by electricity. In a similar way as now the E-bow keeps guitar strings in continuous vibration.

These tuning forks only provide a very gentle sound. If you don’t hold them close by your ear or place them on the bottom of the skull, you will hardly hear anything. Each tuning fork was equipped with a resonator in the form of a closed cylinder with a small hole at the front. The resonance can be tuned through the contents of such a cylinder.

Now, each resonator was tuned in exactly the same way with the corresponding tuning fork. A movable disc was placed in front of each resonator hole, with which the resonator hole could be closed off less or more. So nothing happens at full closure. Only the tuning fork vibrates, but so soft that you do not suffer from it. When the hole is completely open, however, spontaneous resonance arises and the tuning fork vibration is considerably amplified and clearly audible.

Why the tuning fork as the basis for the sound generation you can wonder. The answer is very simple. If you continuously oscillate a tuning fork according to the E-bow principle, only the first vibration mode is activated. This results in a very nice pure sine wave.

Hermann von Helmholtz
In 1863 he published *Die Lehre von den Tonempfindungen als physiological Grundlage für die Theorie der Musik*. Tone perception as a physiological basis for the theory of music. With this work he established his name as the founder of modern sound research. He awakened the beautiful mathematical descriptions of Fourier into sounding results.

Helmholtz is also the first to give the concept of timbre a scientific foundation. He draws up two theories concerning timbre: The *Relativ* and the *Absolut* theory. The first theory explained the timbre of musical instruments, the second of the human voice, the vowels and consonants.

He discovered through analysis and synthesis experiments that the sound of vowels is determined by certain resonances at fixed frequencies, regardless of the pitch at which such a vowel is formed. Take for example the vowel ah. Pronounced by successively woman and man, is clearly recognizable as an ah vowel.

The reason for this is seen in the image `ah-formant.pdf` (after Werner Kaegi's book, Was ist elektronische Musik). Here you see the spectral analysis of the vowel ah. Here you can distinguish a clear resonance maximum at 1000. These energy maximmum that arise through resonance is called a formant. Such a frequency characteristic, or the spectrum envelope is one of the important identity characteristics of a sound.

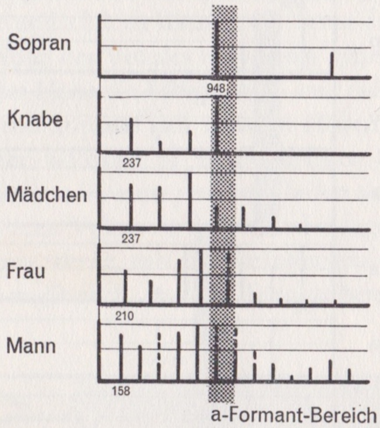

Abb. 3.6 Das harmonische Spektrum des Vokales a ist charakterisiert durch einen a-Formanten im absoluten Frequenzbereich zwischen etwa 900 und 1000 Hz. Auch einer Sopranstimme gelingt es also in jeder Lage sehr leicht, den Vokal a zu singen.

The formants of vowels and consonants
Depending on how accurate your analysis is, you can discover up to four formants in speech sounds. However, for the formation of the vowels and consonants such as those of importance for intelligibility, all of them can be formed with the two lowest formants.
Load the patches, `Vocals(m).pch2`, `Vocals(v).pch2` and `Vocals(k).pch2` and compare the synthesized vowel sounds for a male voice (m), female voice (v) and child voice (k). If you load patch `Consonanten.pch2` you can listen to simulations of the consonants: s, sch (ah)ch, (ee)ch and f.

All of these synthesis examples are based on the formant areas as defined by Dr. Werner Kaegi in *Was ist elektronische Musik*. In these example patches, a simple vending machine plays a tune over two octaves. This is generated by a LFOA followed by a KeyQuantize module and a Glide, which gives it a portamento effect.

Experiment with turning off the portamanto effect and turning off the pitch control on the Pitch input of the Excitation Oscillator. You will hear that as the signal becomes more static, the strength of formant perception decreases. With the Switch module you can also opt for a noise input instead of Excitation Osc. You then obtain the vocal sound colors in whisper version.

The actual formants are formed by two parallel FltNord filters in band pass mode. A Pulse module is connected behind the Excitation Osc. This ensures that the ExcOsc triggers the adjustable pulse duration with each cycle. This is a somewhat simplified representation of the vocal cord signal. In reality, these are a kind of asymmetrical rounded triangle pits with a fixed duration. Only the repetition periodicity is varied by the ExcOsc. Such a signal with a fixed pulse duration independent of the repetition periodicity ensures that a fixed frequency characteristic is produced in the signal. A spectral 'fingerprint', fixed spectrum envelope, as we already encountered as with the formants. Load FixSpectrumEnv.pch2 and tweak the pulse duration of the vocal cord signal and hear how the timbre changes through different pulse duration settings.

Speaking and singing

We are usually unaware of it, but voice and vocal sounds are always provided with a rich modulation in pitch. Would that be completely absent then we hear sounds like with an extremely tuned autotune plug in. Load patch `Singer-Expr.pch2` and listen to the eight different variations. Then load `Singer.pch2` and listen now again. You now hear twice the eight identical examples. Well, almost identical then. In the first example, all fine structure modulation has been omitted. What you hear is only a static signal, although it plays a melody and has a vocal formant structure, but it has nothing to do with singing.

Ernst Bonis

This article was originally published in Dutch as 'Over sinustonen, formanten, spraakvocalen en zang…' in Interface 111 Septmember 2007.

literature

H. Helmholtz
The Lehre von den Tonempfindungen as physiological Grundlage für die Theorie der Musik 1863

Hermann Helmholtz
On the Sensations of Tone (Dover Books on Music) Paperback – June 1, 1954

P. H. Mertens
The Schumannschen Klangfarbengesetze und ihre Bedeutung für die Übertragung von Sprache und Musik
1975 Frankfurt / M
ISBN 3 92011254 7

Werner Kaegi
Was ist electronic Musik
Orell Füssli Verlag
Zurich 1967

internet links

Excellent tutorials about the human voice
http://hyperphysics.phy-astr.gsu.edu/hbase/music/vowel.html

Von Kempelen's speaking machine (1791)
www.ling.su.se/staff/hartmut/kemplne.htm

http://www2.ling.su.se/staff/hartmut/farkas.htm

History of speech synthesis in MP3 audio samples.
www.icsi.berkeley.edu/eecs225d/klatt.html

About formants
http://en.wikipedia.org/wiki/Formant

Hermann von Helmholtz
http://en.wikipedia.org/wiki/Hermann_von_Helmholtz