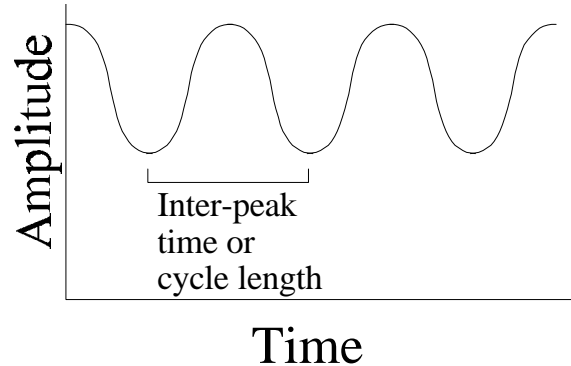


## The real *basics* of sound analysis

What are sounds? A sound is a pressure wave propagated through air. Like any other wave, it can be represented as a waveform. The time between peaks of amplitude is called the *cycle length*, which is in units of time per cycle. The inverse of cycle length, in units of cycles per time, is called the *frequency* of the sound. Most real sounds are complex mixtures of waves of many frequencies and amplitudes.



What is a sonogram? A sonogram is a method for viewing sounds, which plots the frequency of a sound against time (this is NOT the same as the graph above. Why not?). The amplitude of the wave at a given frequency is shown crudely by the intensity of the mark (in black and white) or the shading (in color). With some practice, it is possible to roughly guess what a sonogram will 'sound' like (or conversely, what a sound will 'look' like).

How are sonograms made? These days sonograms are made by sampling the analog (continuously-varying) signal at a large number of predetermined frequencies at small time intervals during a recording. The number of frequency samples at a given time point is called the *sample rate*. These constitute the basic information about the sound. Then, for a predetermined number of frequencies (set by the user), the amplitude of the signal for that frequency is measured and displayed on the sonogram. This process is repeated for each time interval. Because measuring the amplitude of a signal always requires measuring two points per cycle (the peak and the trough values), the highest frequency for which the amplitude can be measured is always one half of the sample rate. This maximum frequency is called the *Nyquist frequency*. Using more frequencies at a given time point takes longer to process, but yields better frequency resolution (more information on the amplitude of sounds at each frequency). Using shorter time intervals also takes longer to process, but gives better temporal resolution (more information on the changes in amplitude and frequency over time).

Some technical terminology: 1) signal-to-noise ratio: the amplitude (in decibels) of what you want to listen to or visualize on a sonogram relative to the amplitude of noise you would rather not listen to or see on a sonogram; 2) fundamental frequency: the frequency (or average frequency) of the lowest sounds made during a given vocalization; 3) harmonics: 'copies' of sounds at the fundamental frequency, which are at exact doublings of the fundamental frequency (for physical reasons -- go look up your old physics textbook); 4) formants: the pattern of amplitudes of the harmonics of complex speech or vocalization sounds. This pattern is determined by the resonance ('echoes') produced by the vocal tract (or instrument) is permeates virtually any sound produced by that instrument. Thus it can serve as a 'signature' of the instrument or speaker. Recognition of the identity of sound producers depends largely on the formants.