

Pressure is the essence of sound. Pressure variations are called **compressions** (increases in pressure) and **rarefactions** (decreases in pressure). When the air pressure inside our ear changes, a thin membrane called the **ear drum** responds by vibrating in and out. These vibrations are transmitted to a chain of small bones called the **hammer, anvil, and stirrup**. The vibrations from the ear drum travel through the bones, causing the stirrup to bang against the **cochlea** (or **inner ear**), which is filled with fluid and tiny hairs. As the cochlea is jostled by the stirrup, the cochlear fluid sloshes around, causing the hairs to sway back and forth. These hairs are connected to the **auditory nerve**, which transforms the vibrations of the hairs into electrical impulses, which are then passed along to the brain for processing.

Any pattern of compression and rarefaction is called a **compression wave**. In particular, a compression wave that passes through an acoustic medium like air or water is called a **sound wave**. If the pattern of a wave repeats, it is called a **periodic wave**. Otherwise, it is called an **aperiodic wave**. Aperiodic waves come in two types: **white noise**, which is a continuous aperiodic wave that has a 'hissing' sound, and a **transient**, which is an instantaneous aperiodic wave that has a 'pop' or 'click' sound. Sound waves found in speech can be either periodic or aperiodic, though more frequently, they are some combination of the two: technically aperiodic with some nearly periodic properties.

The amount of time it takes for a periodic wave to repeat itself is called its **period**, usually symbolized by T . The number of waves that pass by a given point in a specific amount of time is called the **frequency** of the wave, usually symbolized by f . Frequency is usually measured in **Hertz** (Hz), which is equal to 1 wave per second. The frequency of a sound wave corresponds to its pitch. Humans can hear in the range of about 20–20,000 Hz. The period and frequency of a wave are related by the formula $f = 1/T$.

The physical space separating repetitions of a periodic wave is called the **wavelength** of the wave, symbolized by λ . Though frequency is usually the more important property of sound waves that we care about for phonetics, wavelength can play an important role. The **speed** of a wave, symbolized by s , is the rate at which a wave travels. Speed, frequency, and wavelength are related by the formula $s = f\lambda$. The speed of a wave depends on the nature of the medium it is traveling through: density, temperature, elasticity, etc. For warm, moist air, as is typically found in the human mouth, the speed of sound is about 35,000 cm/sec.

The **amplitude** of a wave, symbolized by A is the magnitude of variation in the wave's pattern. For sound waves, the amplitude corresponds to loudness or volume. Since waves are patterns of variation by definition, the amplitude of a wave changes over time. The formula for a **simple wave** can be written as $A(t) = A_{\max} \sin(2\pi \cdot f \cdot t)$, where $A(t)$ is the amplitude at time t , A_{\max} is the maximum amplitude of the wave, and f is the frequency.

Waves may not be simple however, and most speech sounds are actually **complex waves**, combinations of multiple simple waves. These component waves each have their own frequencies and amplitudes, and they are added together to yield the complex wave. The general formula for a complex wave is $A(t) = A_{w_1} \sin(2\pi \cdot f_{w_1} t) + \dots + A_{w_n} \sin(2\pi \cdot f_{w_n} t)$, where $A(t)$ is the amplitude of the complex wave at time t , A_{w_1} and f_{w_1} are the maximum amplitude and the frequency of the first simple component wave, up through A_{w_n} and f_{w_n} , which are the maximum amplitude and the frequency of the n th simple component wave.

In addition to their component frequencies, complex waves also have a **fundamental frequency**, usually symbolized as f_0 . (Note: for simple waves, we often say that their frequency f is their fundamental frequency f_0 ; i.e., $f_0 = f$.) The fundamental frequency of a complex wave is calculated by finding the **greatest common factor** of the set of all of the component frequencies. That is, $f_0 = \text{gcf}(f_{w_1}, \dots, f_{w_n})$. The gcf is the largest number that divides all of the components frequencies evenly. Use traditional methods from middle school mathematics to find the gcf of frequencies that

are integers. If the frequencies are not integers, multiply them all by the same integer N to turn them into integers, then find the gcf of the new numbers, and finally, divide that result by N to find the actual gcf. For example, if the component frequencies are 44.4 Hz and 55.5 Hz, multiply both frequencies by 10 to get 444 and 555. The gcf of 444 and 555 is 111. Divide this by 10 to under the previous multiplication, and the result is 11.1 Hz as the fundamental frequency.

All sound waves naturally produce extra sound waves at higher frequencies called **harmonics**. The first harmonic is the sound wave with a frequency equal to the fundamental frequency of the original sound wave. In the case of a simple wave, the simple wave is its own first harmonic. The n th harmonic has a frequency equal to n times the fundamental frequency. The amplitude of the harmonics falls off rather quickly. The maximum amplitude of the first harmonic is equal to the maximum amplitude of the original sound wave, while the maximum amplitude of the n th harmonic is equal to $1/n^2$ times the maximum amplitude of the original sound wave. Putting this all together, the formula for the n th harmonic is $A_n = \frac{1}{n^2} A_0 \sin(2\pi \cdot n f_0 t)$.

Sound waves can be generated by any vibrating object. The natural frequencies at which an object vibrates are called its **resonant frequencies**, which depend on the shape of the object and the material it's made of. For a body of air in a **tube closed at both ends** with a length L , the first resonant frequency is $s/2L$, and in general, the n th resonant frequency is $ns/2L$. For a body of air in a **tube closed at one end and open at the other** with a length L , the first resonant frequency is $s/4L$, and in general, the n th resonant frequency is $(2n - 1)s/4L$.

Objects can act as **filters** for sound waves by enhancing waves with particular frequencies and dampening others. Generally, objects acting as filters enhance frequencies near their own resonant frequencies and dampen frequencies farther away from their resonant frequencies. The frequencies that are enhanced the most by a filter are called its **center frequencies**. Filters usually have a **bandwidth**, which is the size of the range of frequencies around the center frequency that are also significantly enhanced. For example, if a filter has a center frequency of 500 Hz and a bandwidth of 100 Hz, then frequencies between 450 Hz and 550 Hz will be enhanced, while all others will be dampened.

The idea of filtering plays an important role in acoustically analyzing speech, since the vocal tract acts as a filter. By approximating the various cavities in the vocal tract as tubes, we can give a rough but accurate estimate for the acoustic properties of different speech sounds based on how they are articulated.