The Third Lecture:

Simple and Complex Sounds

The movement of a tuning fork is very simple that is back and forth, so that the wave of the sound it produces is very simple. The sound which is made in this way is a pure tone of a single frequency (like 440 Hz).But the vibrations of the vocal tract and the sounds it creates are more complex. Speech sounds are made according to the source-filter theory of speech production. When air passes out of the windpipe(trachea) and over the vocal folds, the folds start to vibrate. They flap open and closed at a frequency between around 100 times per second for a large adult male and 300times per second for a child. But on the top of this basic flapping movement, there are many various sub ripples in the moving vocal folds.They do not flap up and down firmly, but wave in complex patterns. Each of these little waves contribute its own pattern to the sound making "overtones," or harmonics, the main pith of the speaker's voice as well. The basic rates of vibration, the fundamental frequency, identify the pitch, but the overtones create the various qualities of different sounds. Whereas the vocal folds vibrate, the air inside the vocal tract begins to vibrate in the same complicated way, then the air in the vocal tract filters the harmonic structure of the sound made at the vocal folds. Certain waves(certain harmonics)are amplified that depend on the shape of air column inside the vocal tract, and other waves are diluted. Different air particles tend to vibrate at different frequencies. Harmonics that are "in tune" with the characteristic frequencies of a particular vocal tract shape will be amplified, while those which are not in tune will be reduced. The filter is controlled by moving the tongue and lips to various positions by the speaker. He amplifies some harmonics and withholding others. The most strongly amplified frequencies are called formants.In the acoustic analysis of speech sounds it has been found that in addition to the glottal pitch of voiced sounds (and it has been calculated that about eighty per cent of the speech sounds of connected utterance in English are voiced), perceptible phonetic differences are the result of sound waves generated at different areas or 'bands' of frequency. In vowels and certain continuant sounds the features that distinguish one from another are caused by a combination of a small number of separate bands of frequencies(formants) of which two are the most important, ranging from around 200

cycles per second to around 750 cycles per second and from around 700 cycles per second to around 3,000 cycles per second respectively.The sound waves move through the air at the rate of around 340 meters per second, until they clash with a membrane tuned to receive them, such as the eardrum. They move down the air in the ear canal until they reach the eardrum that begins to vibrate. The patterns of vibration are transferred from the eardrum through the bones of the middle ear to the inner ear. This additional stage of transfer, through the middle ear, helps to amplify very soft sounds, and tone down very loud sounds. The process of hearing occurs actually in the inner ear (cochlea). Because the cochlear membrane varies in shape along its length, it thick at one end and thin at the other, different places along the membrane respond to different frequencies of vibration. The thick end vibrates in tune to low-pitched sounds, the thin end in tune to high to mid-range sounds. In response to a given pattern of vibration, sending signals to the brain about the frequencies present in the received wave. The brain reunifies the frequency information it recognizes into the sounds of language.

Several theories of hearing have been submitted since the subject was first systematically investigated in the mid-19th century. *Resonance* or *place theory* is the classical theory of pitch perception, deriving from the work of German scientist. Individual fibers in the cochlea were thought to reecho to a particular frequency and if the frequency changes, the place of vibration along the basilar membrane will change too. The basilar membrane vibrates along most of its length. *Temporal* or *frequency theory* is proposed by William Rutherford in 1886.The frequency of a wave was thought to be transferred by the number of pulses per second in a nerve fibre. The cochlea is working as a type of telephone transmitter, directly passing on frequency information to the auditory nerve. After the discovery of no nerve fibre can release at more than 1,000Hz.That most releases happen at much lower rates. As human beings can respond to speech frequencies ranging up to 20,000 Hz, a purely temporal theory is insufficient.

Volley theory is suggested by E.G Wever in 1949, represents a compromise solution between place and temporal theories. It suggests that below 5,000 Hz temporal patterning is important, with pitch perception being dependent on the synchronized action of several nerve

fibers, release in volleys. Above 5,000 Hz, place analysis is well maintained.

Voiced and Voiceless Sounds

There is an articulatory process called voicing in which the vocal flaps are assigned in vibration by leaving column of air. The lungs push the air and the airflow faces a controlled resistance at the larynx. The resistance can be controlled by different positions and tensions in the vocal cords. The free flow of air is allowed to and from the lungs during the process of quiet breathing when the cords are relaxed and spread apart, but during swallowing the cords are drawn tightly together to prevent foreign material from entering the lungs. The most important feature of the vocal cords for speech is that they can be made to vibrate if the airflow between them is rapid enough and if they have the appropriate tension and closeness to each other. The vocal cords are brought close tightly and the airstream vibrates them in rapid succession during voicing .Rapid closing and opening of these vocal cords several times a second is made. Consequently, the sounds that can be produced without the vibration of the vocal cords are called unvoiced or voiceless sounds, while those produced with the vibration of the vocal cords are called voiced sounds. In English /g, b, d, dʒ, v, z, ʒ, ð, m, n, ŋ, l,j w, r/ consonants and all the vowels are voiced sounds. The voiceless are/k, p, t, tʃ, f, s, ʃ, θ, h/. Frequency of the vocal cords vibration is associated with pitch level, low and high tones and voice amplitude. The usual frequency of which in human speech is from 80 to 400 cycles per second. The vocal cords of adult males are larger in size than females and children; therefore, their frequency of vibration is relatively lower than the frequency of vibration in females and children. So the pitch of adult males' voices is lower than that of them. Voicing has a critical function in speech production. It is a basic factor in the fundamental classification of speech sounds into two functional categories, the voiced and voiceless sounds. The dichotomy signifies grouping of sounds according to the degree of muscular tension. *"Fortis*" refers to voiceless sounds that are pronounced with greater muscular energy and breath ,whilst in *"lenis*" the muscular energy and breath are markedly reduced and mostly voiced sounds are *lenis*.