"Voice and Articulation".

In Section 1 of this course you will cover these topics:

- An Introduction
- Mechanisms For Speech

Topic : An Introduction

Topic Objective:

At the end of this topic student would be able to:

- Understand and give an Overview of Voice and Speech
- Discuss the Types of communication
- Understand the Influences of Voice and Speech
- Discuss the Voice Modulation in Spoken Language
- Discuss on how to Improve the Voice

Definition/Overview:

The topic discusses that Communication is commonly defined as "the imparting or interchange of thoughts, opinions, or information by speech, writing, or signs...", , 1: an act or instance of transmitting and 3 a: "a process by which information is exchanged between individuals through a common system of symbols, signs, or behavior ... also: exchange of information". Communication can be perceived as a two-way process in which there is an exchange and progression of thoughts, feelings or ideas towards a mutually accepted[clarification needed] goal or direction. Communication as an academic discipline relates to all the ways we communicate, so it embraces a large body of study and knowledge.

Key Points:

1. Overview

Communication is a process whereby information is encoded and imparted by a sender to a receiver via a channel/medium. The receiver then decodes the message and gives the sender a feedback. Communication requires that all parties have an area of communicative

commonality. There are auditory means, such as speaking, singing and sometimes tone of voice, and nonverbal, physical means, such as body language, sign language, paralanguage, touch, eye contact, by using writing.

Communication is thus a process by which we assign and convey meaning in an attempt to create shared understanding. This process requires a vast repertoire of skills in intrapersonal and interpersonal processing, listening, observing, speaking, questioning, analyzing, and evaluating. If you use these processes it is developmental and transfers to all areas of life: home, school, community, work, and beyond. It is through communication that collaboration and cooperation occur.

Communication is the articulation of sending a message through different media, whether it is verbal or nonverbal, so long as a being transmits a thought provoking idea, gesture, action, etc. Communication is a learned skill. Most babies are born with the physical ability to make sounds, but must learn to speak and communicate effectively. Speaking, listening, and our ability to understand verbal and nonverbal meanings are skills we develop in various ways. We learn basic communication skills by observing other people and modeling our behaviors based on what we see. We also are taught some communication skills directly through education, and by practicing those skills and having them evaluated.

There are also several common barriers to successful communication, two of which are message overload (when a person receives too many messages at the same time), and message complexity

2. Types of communication

There are three major parts in human face to face communication which are body language, voice tonality, and words. According to the research:

- 55% of impact is determined by body language--postures, gestures, and eye contact,
- 38% by the tone of voice, and
- 7% by the content or the words used in the communication process.

Although the exact percentage of influence may differ from variables such as the listener and the speaker, communication as a whole strives for the same goal and thus, in some

cases, can be universal. System of signals, such as voice sounds, intonations or pitch, gestures or written symbols which communicate thoughts or feelings. If a language is about communicating with signals, voice, sounds, gestures, or written symbols, can animal communications be considered as a language? Animals do not have a written form of a language, but use a language to communicate with each another. In that sense, an animal communication can be considered as a separate language.

Human spoken and written languages can be described as a system of symbols (sometimes known as lexemes) and the grammars (rules) by which the symbols are manipulated. The word "language" is also used to refer to common properties of languages. Language learning is normal in human childhood. Most human languages use patterns of sound or gesture for symbols which enable communication with others around them. There are thousands of human languages, and these seem to share certain properties, even though many shared properties have exceptions.

There is no defined line between a language and a dialect, but the linguist Max Weinreich is credited as saying that "a language is a dialect with an army and a navy". Constructed languages such as Esperanto, programming languages, and various mathematical formalisms are not necessarily restricted to the properties shared by human languages.

1.1. Dialogue or verbal communication

A dialogue is a reciprocal conversation between two or more entities. The etymological origins of the word concepts like flowing-through meaning) do not necessarily convey the way in which people have come to use the word, with some confusion between the prefix διά-(di-,through) and the prefix δι- (di-,two) leading to the assumption that a dialogue is necessarily between only two parties.

1.2. Nonverbal communication

Nonverbal communication is the process of communicating through sending and receiving wordless messages. Such messages can be communicated through gesture, body language or posture; facial expression and eye contact, object communication such as clothing, hairstyles or even architecture, or symbols and infographics, as well as through an aggregate of the above, such as behavioral communication. Speech may

also contain nonverbal elements known as paralanguage, including voice quality, emotion and speaking style, as well as prosodic features such as rhythm, intonation and stress. Likewise, written texts have nonverbal elements such as handwriting style, spatial arrangement of words, or the use of emoticons.A portmanteau of the English words emotion (or emote) and icon, an emoticon is a symbol or combination of symbols used to convey emotional content in written or message form.

Other communication channels such as telegraphy fit into this category, whereby signals travel from person to person by an alternative means. These signals can in themselves be representative of words, objects or merely be state projections. Trials have shown that humans can communicate directly in this way without body language, voice tonality or words.

1.3. Visual communication

Visual communication as the name suggests is communication through visual aid. It is the conveyance of ideas and information in forms that can be read or looked upon. Primarily associated with two dimensional images, it includes: signs, typography, drawing, graphic design, illustration, colour and electronic resources. It solely relies on vision. It is form of communication with visual effect. It explores the idea that a visual message with text has a greater power to inform, educate or persuade a person. It is communication by presenting information through Visual form. The evaluation of a good visual design is based on measuring comprehension by the audience, not on aesthetic or artistic preference. There are no universally agreed-upon principles of beauty and ugliness. There exists a variety of ways to present information visually, like gestures, body languages, video and TV. Here, focus is on the presentation of text, pictures, diagrams, photos, et cetera, integrated on a computer display. The term visual presentation is used to refer to the actual presentation of information. Recent research in the field has focused on web design and graphically oriented usability. Graphic designers use methods of visual communication in their professional practice.

3. Influences on Your Voice and Speech

The human voice consists of sound made by a human being using the vocal folds for talking, singing, laughing, crying, screaming, etc. Human voice is specifically that part of

human sound production in which the vocal folds (vocal cords) are the primary sound source. Generally speaking, the mechanism for generating the human voice can be subdivided into three parts; the lungs, the vocal folds within the larynx, and the articulators. The lung (the pump) must produce adequate airflow and air pressure to vibrate vocal folds (this air pressure is the fuel of the voice). The vocal folds (vocal cords) are a vibrating valve that chops up the airflow from the lungs into audible pulses that form the laryngeal sound source. The muscles of the larynx adjust the length and tension of the vocal folds to fine tune pitch and tone. The articulators (the parts of the vocal tract above the larynx consisting of tongue, palate, cheek, lips, etc.) articulate and filter the sound emanating from the larynx and to some degree can interact with the laryngeal airflow to strengthen it or weaken it as a sound source.

The vocal folds, in combination with the articulators, are capable of producing highly intricate arrays of sound. The tone of voice may be modulated to suggest emotions such as anger, surprise, or happiness. Singers use the human voice as an instrument for creating music.

4. Voice types and the folds (cords) themselves

Adult men and women have different vocal folds sizes; reflecting the male-female differences in larynx size. Adult male voices are usually lower-pitched and have larger folds. The male vocal folds (which would be measured vertically in the opposite diagram), are between 17 mm and 25 mm in length. The female vocal folds are between 12.5 mm and 17.5 mm in length.

As seen in the illustration, the folds are located just above the trachea (the windpipe which travels from the lungs). Food and drink do not pass through the cords but instead pass through the esophagus, an unlinked tube. Both tubes are separated by the epiglottis, a "flap" that covers the opening of the trachea while swallowing.

The folds in both sexes are within the larynx. They are attached at the back (side nearest the spinal cord) to the arytenoid cartilages, and at the front (side under the chin) to the thyroid cartilage. They have no outer edge as they blend into the side of the breathing tube (the illustration is out of date and does not show this well) while their inner edges or "margins" are free to vibrate (the hole). They have a three layer construction of an

epithelium, vocal ligament, then muscle (vocalis muscle), which can shorten and bulge the folds. They are flat triangular bands and are pearly white in color. Above both sides of the vocal cord is the vestibular fold or false vocal cord, which has a small sac between its two folds (not illustrated).

The difference in vocal folds size between men and women means that they have differently pitched voices. Additionally, genetics also causes variances amongst the same sex, with men and women's singing voices being categorized into types. For example, among men, there are basses, baritones and tenors, and among women, contraltos, mezzo-sopranos and sopranos. There are additional categories for operatic voices, see voice type. This is not the only source of difference between male and female voice. Men, generally speaking, have a larger vocal tract, which essentially gives the resultant voice a lower tonal quality. This is mostly independent of the vocal folds themselves.

5. Voice Modulation in Spoken Language

Human spoken language makes use of the ability of almost all persons in a given society to dynamically modulate certain parameters of the laryngeal voice source in a consistent manner. The most important communicative, or phonetic, parameters are the voice pitch (determined by the vibratory frequency of the vocal folds) and the degree of separation of the vocal folds, referred to as vocal fold adduction (coming together) or adduction (separating).

The ability to vary the ab/adduction of the vocal folds quickly has a strong genetic component, since vocal fold adduction has a life-preserving function in keeping food from passing into the lungs, in addition to the covering action of the epiglottis. Consequently, the muscles that control this action are among the fastest in the body. Children can learn to use this action consistently during speech at an early age, as they learn to speak the difference between utterances such as "apa" (having an abductory-adductory gesture for the p) as "aba" (having no abductory-adductory gesture). Surprisingly enough, they can learn to do this well before the age of two by listening only to the voices of adults around them who have voices much different than their own, and even though the laryngeal movements causing these phonetic differentiations are deep in the throat and not visible to them.

If an abductory movement or adductory movement is strong enough, the vibrations of the vocal folds will stop (or not start). If the gesture is abductory and is part of a speech sound, the sound will be called [Voiceless]. However, voiceless speech sounds are sometimes better identified as containing an abductory gesture, even if the gesture was not strong enough to stop the vocal folds from vibrating. This anomalous feature of voiceless speech sounds is better understood if it is realized that it is the change in the spectral qualities of the voice as abduction proceeds that is the primary acoustic attribute that the listener attends to when identifying a voiceless speech sound, and not simply the presence or absence of voice (periodic energy).

An adductory gesture is also identified by the change in voice spectral energy it produces. Thus, a speech sound having an adductory gesture may be referred to as a "glottal stop" even if the vocal fold vibraions do not entirely stop. for an example illustrating this, obtained by using the inverse filtering of oral airflow.)

Other aspects of the voice, such as variations in the regularity of vibration, are also used for communication, and are important for the trained voice user to master, but are more rarely used in the formal phonetic code of a spoken language.

6. Improving Your Voice

6.1 Physiology and vocal timbre

The sound of each individual's voice is entirely unique not only because of the actual shape and size of an individual's vocal cords but also due to the size and shape of the rest of that person's body, especially the vocal tract, and the manner in which the speech sounds are habitually formed and articulated. (It is this latter aspect of the sound of the voice that can be mimiced by skilled performers.) Humans have vocal folds which can loosen, tighten, or change their thickness, and over which breath can be transferred at varying pressures. The shape of chest and neck, the position of the tongue, and the tightness of otherwise unrelated muscles can be altered. Any one of these actions results in a change in pitch, volume, timbre, or tone of the sound produced. Sound also resonates within different parts of the body, and an individual's size and bone structure can affect somewhat the sound produced by an individual.

Singers can also learn to project sound in certain ways so that it resonates better within their vocal tract. This is known as vocal resonation. Another major influence on vocal sound and production is the function of the larynx which people can manipulate in different ways to produce different sounds. These different kinds of laryngeal function are described as different kinds of vocal registers. The primary method for singers to accomplish this is through the use of the Singer's Formant, which has been shown to be a resonance added to the normal resonances of the vocal tract above the frequency range of most instruments and so enables the singer's voice to carry better over musical accompaniment

6.2 Vocal registration

Vocal registration refers to the system of vocal registers within the human voice. A register in the human voice is a particular series of tones, produced in the same vibratory pattern of the vocal folds, and possessing the same quality. Registers originate in laryngeal function. They occur because the vocal folds are capable of producing several different vibratory patterns. Each of these vibratory patterns appears within a particular range of pitches and produces certain characteristic sounds. The term register can be somewhat confusing as it encompasses several aspects of the human voice. The term register can be used to refer to any of the following:

A particular part of the vocal range such as the upper, middle, or lower registers.

A resonance area such as chest voice or head voice.

A phonatory process

A certain vocal timbre

A region of the voice which is defined or delimited by vocal breaks.

A subset of a language used for a particular purpose or in a particular social setting.

In linguistics, a register language is a language which combines tone and vowel phonation into a single phonological system.

Within speech pathology the term vocal register has three constituent elements: a certain vibratory pattern of the vocal folds, a certain series of pitches, and a certain type of sound. Speech pathologists identify four vocal registers based on the physiology of laryngeal function: the vocal fry register, the modal register, the falsetto register, and the whistle register. This view is also adopted by many vocal pedagogists

6.3 Vocal resonation

Vocal resonation is the process by which the basic product of phonation is enhanced in timbre and/or intensity by the air-filled cavities through which it passes on its way to the outside air. Various terms related to the resonation process include amplification, enrichment, enlargement, improvement, intensification, and prolongation, although in strictly scientific usage acoustic authorities would question most of them. The main point to be drawn from these terms by a singer or speaker is that the end result of resonation is, or should be, to make a better sound. There are seven areas that may be listed as possible vocal resonators. In sequence from the lowest within the body to the highest, these areas are the chest, the tracheal tree, the larynx itself, the pharynx, the oral cavity, the nasal cavity, and the sinuses

7. Influences of the human voice

The twelve-tone musical scale, upon which the majority of the world's music is based, may have its roots in the sound of the human voice during the course of evolution, according to a study published by the New Scientist. Analysis of recorded speech samples found peaks in acoustic energy that mirrored the distances between notes in the twelve-tone scale.

8. Voice disorders

There are many disorders which affect the human voice; these include speech impediments, and growths and lesions on the vocal folds. Talking for improperly long periods of time causes vocal loading, which is stress inflicted on the speech organs. When

vocal injury is done, often an ENT specialist may be able to help, but the best treatment is the prevention of injuries through good vocal production. Voice therapy is generally delivered by a Speech-language pathologist. Hoarseness or breathiness that lasts for more than two weeks is a common symptom of an underlying voice disorder and should be investigated medically.

9. Speech production

Speech refers to the processes associated with the production and perception of sounds used in spoken language. A number of academic disciplines study speech and speech sounds, including acoustics, psychology, speech pathology, linguistics, cognitive science, communication studies, otolaryngology and computer science. In linguistics (articulatory phonetics), manner of articulation describes how the tongue, lips, and other speech organs are involved in making a sound make contact. Often the concept is only used for the production of consonants. For any place of articulation, there may be several manners, and therefore several homorganic consonants.

10. Speech perception

Speech perception refers to the processes by which humans are able to interpret and understand the sounds used in language. The study of speech perception is closely linked to the fields of phonetics and phonology in linguistics and cognitive psychology and perception in psychology. Research in speech perception seeks to understand how human listeners recognize speech sounds and use this information to understand spoken language. Speech research has applications in building computer systems that can recognize speech, as well as improving speech recognition for hearing- and language-impaired listeners.

Topic : Mechanisms For Speech

Topic Objective:

At the end of this topic student would be able to:

- Discuss the Mechanisms for Speech.
- Understand the concept of Breathing
- Discuss the Phonation techniques

- Understand the concept of Resonation
- Understand the concept of Articulation

Definition/Overview:

The topic discusses that breathing takes oxygen in and carbon dioxide out of the body. Aerobic organisms require oxygen to create energy via respiration, in the form of the metabolism of energy-rich molecules such as glucose. The medical term for normal relaxed breathing is eupnea. Breathing is only part of the processes of delivering oxygen to where it is needed in the body and removing carbon dioxide waste. The process of gas exchange occurs in the alveoli by passive diffusion of gases between the alveolar gas and the blood passing by in the lung capillaries. Once in the blood the heart powers the flow of dissolved gases around the body in the circulation. As well as carbon dioxide, breathing also results in loss of water from the body. Exhaled air has a relative humidity of 100% because of water diffusing across the moist surface of breathing passages and alveoli.

Key Points:

1. Mechanics

In mammals, breathing in, or inhaling, is usually an active movement, with the contraction of the diaphragm muscle. This is known as negative pressure breathing. Normally, the diaphragm's relaxed position recoiled (decreasing the thoracic volume) whereas in the contracted position it is pulled downwards (increasing the thoracic volume). This process works in conjunction with the intercostal muscles connected to the rib cage. Contraction of these muscles lifts the rib cage, thus aiding in increasing the thoracic volume. Relaxation of the diaphragm compresses the lungs, effectively decreasing their volume while increasing the pressure inside them. The intercostal muscles simultaneously relax, further decreasing the volume of the lungs. With a pathway to the mouth or nose clear, this increased pressure forces air out of the lungs. Conversely, contraction of the diagraphm increases the volume of the (partially empty) lungs, decreasing the pressure inside, which creates a partial vacuum. Environmental air then follows its pressure gradient down to fill the lungs.

In amphibians, the process used is positive pressure breathing. Muscles lower the floor of the oral cavity, enlarging it and drawing in air through the nostrils (which uses the same mechanics - pressure, volume, and diffusion - as a mammalian lung). With the nostrils and mouth closed, the floor of the oral cavity is forced up, which forces air down the trachea into the lungs.

At rest, breathing out, or exhaling, is a combination of passive and active processes powered by the elastic recoil of the alveoli, similar to a deflating balloon, and the contraction of the muscular body wall. The following organs are used in respiration: the mouth; the nose and nostrils; the pharynx; the larynx; the trachea; the bronchi and bronchioles; the lungs; the diaphragm; and the terminal branches of the respiratory tree, such as the alveoli.

2. Control of breathing

Breathing is one of the few bodily functions which, within limits, can be controlled both consciously and unconsciously. Conscious control of breathing is common in many forms of meditation, specifically forms of yoga for example pranayama unlike anapana which is only awareness of breath. In swimming, cardio fitness, speech or vocal training, one learns to discipline one's breathing, initially consciously but later sub-consciously, for purposes other than life support. Human speech is also dependent on conscious breath control.

Unconsciously, breathing is controlled by specialized centers in the brainstem, which automatically regulate the rate and depth of breathing depending on the bodys needs at any time. When carbon dioxide levels increase in the blood, it reacts with the water in blood, producing carbonic acid. Lactic acid produced by anaerobic exercise also lowers pH. The drop in the blood's pH stimulates chemoreceptors in the carotid and aortic bodies in the blood system to send nerve impulses to the respiration centre in the medulla oblongata and pons in the brain. These, in turn send nerve impulses through the phrenic and thoracic nerves to the diaphragm and the intercostal muscles, increasing the rate of breathing. Even a slight difference in the bloods normal pH, 7.4, could cause death, so this is an important process. This automatic control of respiration can be impaired in premature babies, or by drugs or disease.

2.1 Examples

For instance, while exercising, the level of carbon dioxide in the blood increases due to increased cellular respiration by the muscles, which activates carotid and aortic bodies and the respiration center, which ultimately cause a higher rate of respiration. During rest, the level of carbon dioxide is lower, so breathing rate is lower. This ensures an appropriate amount of oxygen is delivered to the muscles and other organs. It is important to reiterate that it is the buildup of carbon dioxide making the blood acidic that elicits the desperation for a breath much more than lack of oxygen.

3. Interaction

It is not possible for a healthy person to voluntarily stop breathing indefinitely. If we do not inhale, the level of carbon dioxide builds up in our blood, and we experience overwhelming air hunger. This irrepressible reflex is not surprising given that without breathing, the body's internal oxygen levels drop dangerously low within minutes, leading to permanent brain damage followed eventually by death. However, there have been instances where people have survived for as long as two hours without air; this is only possible when submerged in cold water, as this triggers the mammalian diving reflex.

If a healthy person were to voluntarily stop breathing (i.e. hold his or her breath) for a long enough amount of time, he or she would lose consciousness, and the body would resume breathing on its own. Because of this one cannot suffocate oneself with this method, unless one's breathing was also restricted by something else (e.g. water, see drowning)

Hyperventilating causes a drop in CO_2 below normal levels, lowering blood acidity to trick the brain into thinking it has more oxygen than is actually present. Hyperventilating can cause blood oxygen levels to go to dangerous levels.

4. Relationship to death

Most organisms breathe to avoid death by asphyxiation. Breath is sometimes used as a metaphor for life itself, and often "last breath" is the most obvious sign that death has occurred. The association between the end of life and breathing is not absolute, as scientists have discovered the brain/mind can continue to function for many minutes

without the continuation of oxygen and/or its properties. Though thought to "restart" breathing, cardiopulmonary resuscitation (CPR) only circulates blood through the body.

5. Phonation

Phonation has slightly different meanings depending on the subfield of phonetics. Among some phoneticians, phonation is the process by which the vocal folds produce certain sounds through quasi-periodic vibration. This is the definition used among those who study laryngeal anatomy and physiology and speech production in general. Other phoneticians, though, call this process quasi-periodic vibration voicing, and they use the term phonation to refer to any oscillatory state of any part of the larynx that modifies the airstream, of which voicing is just one example. As such, voiceless and supra-glottal phonations are included under this definition, which is common in the field of linguistic phonetics.

5.1 Voicing

The phonatory process, or voicing, occurs when air is expelled from the lungs through the glottis, creating a pressure drop across the larynx. When this drop becomes sufficiently large, the vocal folds start to oscillate. The minimum pressure drop required to achieve phonation is called the phonation threshold pressure, and for humans with normal vocal folds, it is approximately 23 cm H2O. The motion of the vocal folds during oscillation is mostly in the lateral direction, though there is also some superior component as well. However, there is almost no motion along the length of the vocal folds. The oscillation of the vocal folds serves to modulate the pressure and flow of the air through the larynx, and this modulated airflow is the main component of the sound of most voiced phones.

The sound that the larynx produces is a harmonic series. In other words, it consists of a fundamental tone (called the fundamental frequency, the main acoustic cue for the percept pitch) accompanied by harmonic overtones which are multiples of the fundamental frequency . According to the Source-Filter Theory, the resulting sound excites the resonance chamber that is the vocal tract to produce the individual speech sounds.

The vocal folds will not oscillate if they are not sufficiently close to one another, are not under sufficient tension or under too much tension, or if the pressure drop across the larynx is not sufficiently large. In linguistics, a phone is called voiceless if there is no phonation during its occurrence. In speech, voiceless phones are associated with vocal folds that are elongated, highly tensed, and placed laterally (abducted) when compared to vocal folds during phonation .

Fundamental frequency, the main acoustic cue for the percept pitch, can be varied through a variety of means. Large scale changes are accomplished by increasing the tension in the vocal folds through contraction of the cricothyroid muscle. Smaller changes in tension can be effected by contraction of the thyroarytenoid muscle or changes in the relative position of the thyroid and cricoid cartilages, as may occur when the larynx is lowered or raised, either volitionally or through movement of the tongue to which the larynx is attached via the hyoid bone . In addition to tension changes, fundamental frequency is also affected by the pressure drop across the larynx, which is mostly affected by the pressure in the lungs, and will also vary with the distance between the vocal folds. Variation in fundamental frequency is used linguistically to produce intonation and tone.

There are currently two main theories as to how vibration of the vocal folds is initiated: the myoelastic theory and the aerodynamic theory . These two theories are not in contention with one another and it is quite possible that both theories are true and operating simultaneously to initiate and maintain vibration. A third theory, the neurochronaxic theory, was in considerable vogue in the 1950s, but has since been largely discredited

5.2 Myoelastic and aerodynamic theory

The myoelastic theory states that when the vocal cords are closed and breath pressure is applied to them, the cords remain closed until the pressure beneath themthe subglottic pressureis sufficient to push them apart, allowing air to escape and reducing the pressure enough for the muscle tension to pull the folds back together again. Pressure builds up once again until the cords are pushed apart, and the whole cycle keeps repeating itself. The rate at which the cords open and closethe number of cycles per seconddetermines the pitch of the phonation.

The Aerodynamic theory is based on the Bernoulli effect. The theory states that breath is flowing through the glottis while the arytenoid cartilages are being pulled together by the action of the interarytenoid muscles. Due to the Bernoulli Effect, the breath flowing past the vocal folds causes them to be sucked into vibration before the arytenoids are fully together. When the arytenoids have been pulled together, this same air flow sucks the glottis closed, thus cutting off the air flow until breath pressure pushes the folds apart and the flow starts up again, causing the cycles to repeat. The Myoelastic-Aerodynamic Theory of Phonation by Janwillem van den Berg is a combination of the two above theories

5.3 Neurochronaxic theory

This theory states that the frequency of the vocal fold vibration is determined by the chronaxy of the recurrent nerve, and not by breath pressure or muscular tension. Advocates of this theory thought that every single vibration of the vocal folds was due to an impulse from the recurrent laryngeal nerves and that the acoustic center in the brain regulated the speed of vocal fold vibration. Speech and voice scientists have long since left this theory as the muscles have been shown to not be able to contract fast enough to accomplish the vibration. In addition, persons with paralyzed vocal folds can produce phonation, which would not be possible according to this theory. Phonation occurring in excised larynges would also not be possible according to this theory.

5.4 As the state of the glottis

In linguistic phonetic treatments of phonation, such as those of Peter Ladefoged, phonation was considered to be a matter of points on a continuum of tension and closure of the vocal cords. More intricate mechanisms were occasionally described, but they were difficult to investigate, and until recently the state of the glottis and phonation were considered to be nearly synonymous.

If the vocal cords are completely relaxed, with the arytenoid cartilages apart for maximum airflow, the cords do not vibrate. This is voiceless phonation, and is extremely common with obstruents. If the arytenoids are pressed together for glottal closure, the vocal cords block the airstream, producing stop sounds such as the glottal

stop. In between there is a sweet spot of maximum vibration. This is modal voice, and is the normal state for vowels and sonorants in all the world's languages. However, the aperture of the arytenoid cartilages, and therefore the tension in the vocal cords, is one of degree between the end points of open and closed, and there are several intermediate situations utilized by various languages to make contrasting sounds.

For example, Gujarati has vowels with a partially lax phonation called breathy voice or murmured, while Burmese has vowels with a partially tense phonation called creaky voice or laryngealized. Both of these phonations have dedicated IPA diacritics, an under-umlaut and under-tilde. The Jalapa dialect of Mazatec is unusual in contrasting both with modal voice in a three-way distinction. (Note that Mazatec is a tonal language, so the glottis is making several tonal distinctions simultaneously with the phonation distinctions.)

6. Resonation.

Vocal resonation is the process by which the basic product of phonation is enhanced in timbre and/or intensity by the air-filled cavities through which it passes on its way to the outside air. Various terms related to the resonation process include amplification, enrichment, enlargement, improvement, intensification, and prolongation, although in strictly scientific usage acoustic authorities would question most of them. The main point to be drawn from these terms by a singer or speaker is that the end result of resonation is, or should be, to make a better sound.

6.1 A physiological understanding of resonation

In a technical sense resonance is a relationship that exists between two bodies vibrating at the same frequency or a multiple thereof. In other words, the vibrations emanating from one body cause the other body to start vibrating in tune with it. A resonator may be defined as a secondary vibrator which is set into motion by the main vibrator and which adds its own characteristics to the generated sound waves.

There are two basic kinds of resonance: sympathetic resonance (or free resonance) and conductive resonance (or forced resonance). The essential difference between them lies in what causes the resonator to start vibrating. In sympathetic resonance

there is no physical contact between the two bodies. The resonator starts functioning because it receives vibrations through the air and responds to them sympathetically. In conductive resonance the resonator starts vibrating because it is in physical contact with a vibrating body.

Both types of resonance are at work in the human voice during speaking and singing. Much of the vibration felt by singers while singing is a result of conductive resonance. The vibrations created by the vocal folds travel along the bones, cartilages, and muscles of the neck, head, and upper chest, causing them to vibrate. There is little evidence that these vibratory sensations make any significant contribution to the external sound.

These same conductive vibrations, however, are good sensation guides for the singer, regardless of their effect on the external sound. These sensations provide evidence to the singer that his vocal folds are forming strong primary vibrations which are being carried from them to the head and chest. Thus these vibratory sensations can supply feedback about the efficiency of the phonatory process to the singer.

In contrast, the sound a person hears listening to a singer is a product of sympathetic resonance. Vibrations created by the vocal cords travel through air from the larynx into the cavities of the throat and head, setting them into vibration. This is sympathetic resonance, for there is no physical contact between these cavities and the vocal cords. The vocal resonator is not a sounding board of some sort, as comparisons with stringed instruments would make it, but a column of air, whose shape is not only complex, but highly variable. Vennard says the following:

"Thus it may vibrate as a whole or in any of its parts. It should not be too hard to think of it as vibrating several ways at once. Indeed most vibrators do this, otherwise we would not have timbre, which consists of several frequencies of different intensities sounding together. Air is fully as capable of this as any other medium; indeed, the sounds of many diverse instruments are carried to the ear by the same air, are funnelled into the same tiny channel, and can still be heard as one sound or as sounds from the individual sources, depending upon the manner in which we give attention."

6.2 Factors affecting resonators

There are a number of factors which determine the resonance characteristics of a resonator. Included among them are the following: size, shape, type of opening, composition and thickness of the walls, surface, and combined resonators. The quality of a sound can be appreciably changed by rather small variations in these conditioning factors. In general, the larger a resonator is, the lower the frequency it will respond to; the greater the volume of air, the lower its pitch. But the pitch also will be affected by the shape of resonator and by the size of opening and amount of lip or neck the resonator has. A conical shaped resonator, such as a megaphone, tends to amplify all pitches indiscriminately. A cylindrical shaped resonator is affected primarily by the length of the tube through which the sound wave travels. A spherical resonator will be affected by the amount of opening it has and by whether or not that opening has a lip.

Three factors relating to the walls of a resonator will affect how it functions: the material it is made of, the thickness of its walls, and the type of surface it has, The resonance characteristics of a musical instrument obviously will vary with different materials and the amount of material used will have some effect.

Of special importance to singing is the relationship of the surface of a resonator to its tonal characteristics. Resonators can be highly selective-meaning that they will respond to only one frequency (or multiples of it)-or they can be universal-meaning that they can respond to a broad range of frequencies. In general, the harder the surface of the resonator, the more selective it will be, and the softer the surface, the more universal it will become. A hard resonator will respond only when the vibrator contains an overtone that is exactly in tune with the resonator, while a soft resonator permits a wide range of fundamentals to pass through un-dampened but adds its own frequency as on overtone, harmonic or inharmonic as the case may be.

Hardness carried to the extreme will result in a penetrating tone with a few very strong high partials. Softness carried to the extreme will result in a mushy, non-directional tone of little character. Between these two extremes lies a whole gamut of tonal possibilities.

The final factor to be mentioned is the effect of joining two or more resonators together. In general the effect of joining two or more resonators is that resonant frequency or each is lowered in different proportions according to their capacities, their orifices, and so forth. The rules governing combined resonators apply to the human voice, for the throat and mouth and sometimes the nose function in this manner.

6.3 The vocal resonators

There are seven areas that may be listed as possible vocal resonators. In sequence from the lowest within the body to the highest, these areas are the chest, the tracheal tree, the larynx itself, the pharynx, the oral cavity, the nasal cavity, and the sinuses

6.4 The chest

The chest is not an effective resonator. Although strong vibratory sensations may be experienced in the upper chest, and although numerous voice books refer to chest resonance, the chest, by virtue of its design and location, can make no significant contribution to the resonance system of the voice. The chest is on the wrong side of the vocal folds and there is nothing in the design of the lungs that could serve to reflect sound waves back toward the larynx

6.5 The tracheal tree

The tracheal tree makes no significant contribution to the resonance system except for a negative effect around its resonant frequency. The trachea and the bronchial tubes combine to form an inverted Y-shaped structure known as the tracheal tree. It lies just below the larynx, and, unlike the interior of the lungs, has a definite tubular shape and comparatively hard surfaces. The response of the tracheal tree is the same for all pitches except for its own resonant frequency. When this resonant frequency is reached, the response of the subglottic tube is to act as acoustical impedance or interference which tends to upset the phonatory function of the larynx. Research has placed the resonant frequency of the subglottal system or tracheal tree around the E-flat above "middle C" for both men and women, varying somewhat with the size of the individual.

6.6 The larynx

Due to its small size, the larynx acts as a resonator only for high frequencies. Research indicates that one of the desirable attributes of good vocal tone is a prominent overtone lying between 2800 and 3200 hertz, with male voices nearer the lower limit and female voices nearer the upper. This attribute is identified as brilliance, or more frequently as ring or the singer's formant. There are several areas in or adjacent to the larynx which might resonate such a high pitch. Among them are the collar of the larynx, the ventricles of Morgani, the vallecula, and the pyriform sinuses. The larynx is not under conscious control, but whatever produces "ring" can be encouraged indirectly by awareness on the part of the student and the teacher of the sounds which contain it.

6.7 The pharynx

The pharynx is the most important resonator by virtue of its position, size, and degree of adjustability. It is the first cavity of any size through which the product of the laryngeal vibrator passes, the other supraglottal cavities have to accept whatever the pharynx passes on to them. Greene states: "The supraglottic resonators being in the main muscular and moveable structures must be voluntarily controlled to produce conditions of optimal resonance either by varying degrees of tension in their walls, or by alterations in the size of their orifices and cavities during the articulatory movements."

7. Articulation

In music, articulation refers to the direction or performance technique which affects the transition or continuity on single note or between multiple notes or sounds.

There are many different kinds of articulation, each having a different effect on how the note is played. Some articulation marks include the slur, phrase mark, staccato, staccatissimo, accent, sforzando, rinforzando, and legato. Each articulation is represented by a different symbol placed above or below the note (depending on its position on the staff). Woodwind and brass instruments generally articulate by tonguing, the use of the tongue to break the airflow into the instrument. Stringed instruments use different bowing techniques to achieve different articulations. When staccato marks are combined with a slur, the result is portato, also known as articulated legato. Tenuto markings under a slur

are called (for bowed strings) hook bows. This name is also less commonly applied to staccato or martellato (martel) markings.

In Section 2 of this course you will cover these topics:

- Language
- Pitch

Topic : Language

Topic Objective:

At the end of this topic student would be able to:

- Define the Language
- Discuss the History of English.
- Understand the Language Development in the Individual.

Definition/Overview:

The topic discusses that Language is a form of symbolic communication in which elements are combined to represent something other than themselves. The term "language" can also refer to particular systems of communal communication ("languages").

Strictly speaking, language is considered to be an exclusively human mode of communication. Although other animals make use of quite sophisticated communicative systems, sometimes casually referred to as animal language, none of these are known to make use of all of the properties that linguists use to define language.

When discussed more technically as a general phenomenon, "language" always implies a very particular way of human thinking which which can be present even when communication is not the result, and this way of thinking is also sometimes treated as indistinguishable from language itself.

In Western Philosophy for example, language has long been closely associated with reason, which is also a uniquely human way of using symbols. In Ancient Greek philosophical terminology, the same word, logos, was used as a term for both language or speech and reason, and the philosopher Thomas Hobbes used the English word "speech" so that it

similarly could refer to reason, as will be discussed below. More commonly though, the English word "language", derived ultimately from lingua, Latin for tongue, typically refers only to expressions of reason which can be understood by other people, most obviously by speaking.

Key Points:

1. Properties of language

A set of commonly accepted signs (indices, icons or symbols) is only one feature of language; all languages must define (i) the structural relationships between these signs in a system of grammar, (ii) the context wherein the signs are used (pragmatics) and (iii) dependent on their context the content specifity, i.e. its meaning (semantics). Rules of grammar are one of the characteristics sometimes said to distinguish language from other forms of communication. They allow a finite set of signs to be manipulated to create a potentially infinite number of grammatical utterances.

Another property of language is that its symbols are arbitrary. Any concept or grammatical rule can be mapped onto a symbol. In other words, most languages make use of sound, but the combinations of sounds used do not have any necessary and inherent meaning they are merely an agreed-upon convention to represent a certain thing by users of that language. For instance, there is nothing about the Spanish word nada itself that forces Spanish speakers to convey the idea of "nothing". Another set of sounds (for example, the English word nothing) could equally be used to represent the same concept, but all Spanish speakers have acquired or learned to correlate this meaning for this particular sound pattern. For Slovenian, Croatian, Serbian or Bosnian speakers on the other hand, nada means something else; it means "hope".

This arbitrariness even applies to words with an onomatopoetic dimension (i.e. words that to some extent simulate the sound of the token referred to). For example, several animal names (e.g. cuckoo, whip-poor-will, katydid) are derived from sounds the respective animal makes, but these forms did not have to be chosen for these meanings. Non-onomatopoetic words can stand just as easily for the same meaning. For instance, the katydid is called a "bush cricket" in British English, a term that bears no relation to the sound the animal makes. In time, onomatopoetic words can also change in form, losing

their mimetic status. Onomatopoetic words may have an inherent relation to their referent, but this meaning is not inherent, thus they do not violate arbitrariness.

2. Origin of language

Even before the theory of evolution made discussion of more animal-like human ancestors common place, philosophical and scientific speculation casting doubt on the use of early language has been frequent throughout history. In modern Western Philosophy, speculation by authors such as Thomas Hobbes, and later Jean-Jacques Rousseau led to the Acadmie franaise declaring the subject off-limits.

The origin of language is of great interest to philosophers because language is such an essential characteristic of human life. In classical Greek philosophy such inquiry was approached by considering the nature of things, in this case human nature. Aristotle, for example, treated humans as creatures with reason and language by their intrinsic nature, related to their natural propensities to be "political", and dwell in city-state communities (Greek: poleis).

Hobbes followed by John Locke and others claimed that language is an extension of the "speech" which humans have within themselves, which in a sense takes the classical view that reason is one of the most primary characteristics of human nature. Others have argued the opposite - that reason developed out of the need for more complex communication. Rousseau, despite writing before the publication of Darwin's theory of evolution, claimed that there had once been humans who had no language or reason and who developed language first, rather than reason - the development of which things he explicitly described as a mixed blessing, with many negative characteristics.

Since the arrival of Darwin the subject has been approached more often by scientists than philosophers. For example, neurologist Terrence Deacon in his Symbolic Species has argued that reason and language "coevolved". Merlin Donald sees language as a later development building upon what he refers to as mimetic culture, emphasizing that this coevolution depended upon the interactions of many individuals. He writes that: A shared communicative culture, with sharing of mental representations to some degree, must have come first, before language, creating a social environment in which language would have been useful and adaptive. The specific causes of the natural selection that led to language

are however still the subject of much speculation, but a common theme which goes right back to Aristotle is that many theories propose that the gains to be had from language and/or reason were probably mainly in the area of increasingly sophisticated social structures.

In more recent times a theory of mirror neurons has emerged in relation to language, Ramachandran has gone so far as to claim that "mirror neurons will do for psychology what DNA did for biology: they will provide a unifying framework and help explain a host of mental abilities that have hitherto remained mysterious and inaccessible to experiments". Mirror neurons are located in the human inferior frontal cortex and superior parietal lobe, and are unique in that they fire when completing an action and also when witnessing an actor performing the same action. Various studies have proposed a theory of mirror neurons related to language development

3. Human languages

Human languages are usually referred to as natural languages, and the science of studying them falls under the purview of linguistics. A common progression for natural languages is that they are considered to be first spoken, then written, and then an understanding and explanation of their grammar is attempted.

Languages live, die, move from place to place, and change with time. Any language that ceases to change or develop is categorized as a dead language. Conversely, any language that is in a continuous state of change is known as a living language or modern language.

Making a principled distinction between one language and another is usually impossible. For instance, there are a few dialects of German similar to some dialects of Dutch. The transition between languages within the same language family is sometimes gradual.

Some like to make parallels with biology, where it is not possible to make a well-defined distinction between one species and the next. In either case, the ultimate difficulty may stem from the interactions between languages and populations. The concepts of Ausbausprache, Abstandsprache and Dachsprache are used to make finer distinctions about the degrees of difference between languages or dialects.

4. History of English.

English is a West Germanic language that originated in Anglo-Saxon England and has lingua franca status in many parts of the world as a result of the military, economic, scientific, political and cultural influence of the British Empire in the 18th, 19th and early 20th centuries and that of the United States from the mid 20th century onwards. It is used extensively as a second language and as an official language in Commonwealth countries and is the preferred language of many international organizations.

Historically, English originated from several dialects (now called Old English) which were brought to Britain by Anglo-Saxon settlers beginning in the 5th century. The language was heavily influenced by the Old Norse language of Viking invaders. After the Norman conquest, Old English developed into Middle English, borrowing heavily from the Norman (Anglo-French) vocabulary and spelling conventions. Modern English developed from there and continues to adopt foreign words, especially from Latin and Greek.

Modern English, sometimes described as the first global lingua franca, is the dominant international language in communications, science, business, aviation, entertainment, radio and diplomacy. Its spread beyond the British Isles began with the growth of the British Empire, and by the late nineteenth century its reach was truly global. It is the dominant language in the United States, whose growing economic and cultural influence and status as a global superpower since World War II have significantly accelerated the language's adoption across the planet.

A working knowledge of English has become a requirement in a number of fields, occupations and professions such as medicine and computing; as a consequence over a billion people speak English to at least a basic level (see English language learning and teaching). It is also one of six official languages of the United Nations.

Linguists such as David Crystal recognize that one impact of this massive growth of English, in common with other global languages, has been to reduce native linguistic diversity in many parts of the world, most particularly in Australasia and North America, and its huge influence continues to play an important role in language attrition. Similarly, historical linguists, aware of the complex and fluid dynamics of language change, are

always aware of the potential English contains through the vast size and spread of the communities that use it and its natural internal variety, such as in its creoles and pidgins, to produce a new family of distinct languages over time.

English is a West Germanic language that originated from the Anglo-Frisian and Lower Saxon dialects brought to Britain by Germanic settlers and Roman auxiliary troops from various parts of what is now northwest Germany and the Northern Netherlands in the 5th century. One of these German tribes were the Angles, who may have come from Angeln, and Bede wrote that their whole nation came to Britain, leaving their former land empty. The names 'England' (or 'Aenglaland') and English are derived from the name of this tribe.

The Anglo Saxons began invading around 449 AD from the regions of Denmark and Jutland, before the Anglo-Saxons arrived in England the native population spoke Brythonic, a Celtic language. Although the most significant changes in dialect occurred after the Norman invasion of 1066, the language retained its name and the pre-Norman invasion dialect is now known as Old English.

Initially, Old English was a diverse group of dialects, reflecting the varied origins of the Anglo-Saxon Kingdoms of Great Britain. One of these dialects, Late West Saxon, eventually came to dominate. The original Old English language was then influenced by two waves of invasion. The first was by language speakers of the Scandinavian branch of the Germanic family; they conquered and colonized parts of the British Isles in the 8th and 9th centuries. The second was the Normans in the 11th century, who spoke Old Norman and developed an English variety of this called Anglo-Norman. (Over the centuries, this lost the specifically Norman element under the influence of Parisian French and, later, of English, eventually turning into a distinctive dialect of Anglo-French.) These two invasions caused English to become "mixed" to some degree (though it was never a truly mixed language in the strict linguistic sense of the word; mixed languages arise from the cohabitation of speakers of different languages, who develop a hybrid tongue for basic communication).

Cohabitation with the Scandinavians resulted in a significant grammatical simplification and lexical supplementation of the Anglo-Frisian core of English; the later Norman occupation led to the grafting onto that Germanic core of a more elaborate layer of words from the Italic branch of the European languages. This Norman influence entered English

largely through the courts and government. Thus, English developed into a "borrowing" language of great flexibility and with a huge vocabulary. The emergence and spread of the British Empire as well as the emergence of the United States as a superpower helped to spread the English language around the world.

5. Classification and related languages

The English language belongs to the western sub-branch of the Germanic branch of the Indo-European family of languages. The closest living relative of English is either Scots, spoken primarily in Scotland and parts of Northern Ireland, or Frisian. As Scots is viewed by linguists as either a separate language or else as a group of dialects of English, Frisian rather than Scots is often said to be the next closest. After those are other Germanic languages, namely the West Germanic languages (Dutch, Afrikaans, Low German, High German), and the North Germanic languages Swedish, Danish, Norwegian, Icelandic, and Faroese. With the exception of Scots, none of these languages is mutually intelligible with English, because of divergences in lexis, syntax, semantics, and phonology.

Lexical differences with the other Germanic languages arise predominantly because of the heavy usage in English of words taken from Latin (for example, "exit", vs. Dutch uitgang) (literally "out-gang" with "gang" as in "gangway") and French ("change" vs. German nderung, "movement" vs. German Bewegung) (literally "othering" and "be-way-ing" ("proceeding along the way")). The syntax of German and Dutch is also significantly different from English, with different rules for setting up sentences (for example, German Ich habe noch nie etwas auf dem Platz gesehen, vs. English "I have still never seen anything in the square"). Semantics causes a number of false friends between English and its relatives. Phonology differences obscure words which actually are genetically related ("enough" vs. German genug), and sometimes both semantics and phonology are different (German Zeit, "time", is related to English "tide", but the English word has come to mean gravitational effects on the ocean by the moon).

Finally, English has been forming compound words and affixing existing words separately from the other Germanic languages for over 1500 years and has different habits in that regard. For instance, abstract nouns in English may be formed from native words by the suffixes -hood, -ship, -dom and -ness. All of these have cognate suffixes in most or all other Germanic languages, but their usage patterns have diverged, as German "Freiheit"

vs. English "freedom" (the suffix -heit being cognate of English -hood, while English -dom is cognate with German -tum).

Many written French words are also intelligible to an English speaker (though pronunciations are often quite different) because English absorbed a large vocabulary from Norman and French, via Anglo-Norman after the Norman Conquest and directly from French in subsequent centuries. As a result, a large portion of English vocabulary is derived from French, with some minor spelling differences (word endings, use of old French spellings, etc.), as well as occasional divergences in meaning of so-called false friends. The pronunciation of most French loanwords in English (with exceptions such as mirage or phrases like coup dtat) has become completely anglicized and follows a typically English pattern of stress. Some North Germanic words also entered English due to the Danish invasion shortly before then (see Danelaw); these include words such as "sky", "window", "egg", and even "they" (and its forms) and "are" (the present plural form of "to be"

6. Language Development in the Individual

Language development is a process starting early in human life, when a person begins to acquire language by learning it as it is spoken and by mimicry. Children's language development moves from simple to complex. Infants start without language. Yet by four months of age, babies can read lips and discriminate speech sounds. The language that infants speak is called 'gibberish'.

Usually, language starts off as recall of simple words without associated meaning, but as children grow, words acquire meaning, with connections between words formed. In time, sentences start as words are joined together to create logical meaning. As a person gets older, new meanings and new associations are created and vocabulary increases as more words are learned.

Infants use their bodies, vocal cries and other preverbal vocalizations to communicate their wants, needs and dispositions. Even though most children begin to vocalize and eventually verbalize at various ages and at different rates, they learn their first language without conscious instruction from parents or caretakers. It is a seemingly effortless task

that grows increasingly difficult with age. Of course, before any learning can begin, the child must be biologically and socially mature enough.

Linguists do not all agree on the biological factors contributing to language development, however most do agree that our ability to acquire such a complicated system is specific to the human species. Furthermore, our ability to learn language may have been developed through the evolutionary process and that the foundation for language may be passed down genetically.

From birth to one year, comprehension (the language we understand) develops before production (the language we use). There is about a 5 month lag in between the two. Babies have an innate preference to listen to their mother's voice. Babies can recognize familiar words and use preverbal gestures. From 1-2 years, vocabulary grows to several hundred words. There is a vocabulary spurt between 18-24 months, which includes fast mapping. Fast mapping is the babies' ability to learn a lot of new things quickly. The majority of the babies' new vocabulary consists of object words (nouns) and action words (verbs). By 3-5 years, children usually have difficulty using words correctly. Children experience many problems such as underextensions, taking a general word and applying it specifically (for example, 'blankie')and overextensions, taking a specific word and applying it too generally (example, 'car' for 'van'). However, children coin words to fill in for words not yet learned (for example, someone is a cooker rather than a chef because a child will not know what a chef is). Children can also understand metaphors. From 6-10 years, children can understand meanings of words based on their definitions. They also are able to appreciate the multiple meanings of words and use words precisely through metaphors and puns. Fast mapping continues. There are three major theories of language development.

The behaviorist theory, proposed by B. F. Skinner (father of behaviorism) says that language is learned through operant conditioning (reinforcement and imitation). This perspective sides with the nurture side of the nature-nurture debate. This perspective is not widely accepted today because there are many criticisms. These criticisms include that the perspective is too specific, encourages incorrect phrases and is not entirely possible. In order for this to be possible, parents would have to engage in intensive tutoring in order for language to be taught properly.

The nativist theory, proposed by Noam Chomsky, says that language is a unique human accomplishment. Chomsky says that all children have whats called an LAD, an innate language acquisition device that allows children to produce consistent sentences once vocabulary is learned. He also says that grammar is universal. This theory, while there is much evidence supporting it (language areas in the brain, sensitive period for language development, children's ability to invent new language systems) is not believed by all researchers.

The last theory, the interactionist perspective, consists of two components. This perspective is a combination of both the nativist and behaviorist theories. The first part, the information-processing theories, tests through the connectionist model, using statistics. From these theories, we see that the brain is excellent at detecting patterns.

The second part of the interactionist perspective, is the social-interactionist theories. These theories suggest that there is a native desire to understand others as well as being understood by others.

Topic : Pitch

Topic Objective:

At the end of this topic student would be able to:

- Understand the Perception of pitch
- Discuss the Inflection: The Paralinguistics of Pitch.
- Understand the concept of Intonation.
- Discuss the Extralinguistic Features.
- Discuss the Pitch Variety.

Definition/Overview:

The topic discusses that Pitch represents the perceived fundamental frequency of a sound. It is one of the three major auditory attributes of sounds along with loudness and timbre. When the actual fundamental frequency can be precisely determined through physical measurement, it may differ from the perceived pitch because of overtones, also known as partials, harmonic or otherwise, in the sound. The human auditory perception system may also have trouble

distinguishing frequency differences between notes under certain circumstances. According to ANSI acoustical terminology, it is the auditory attribute of sound according to which sounds can be ordered on a scale from low to high.

Key Points:

1. Perception of pitch

Pitch is a subjective sensation in which a listener assigns perceived tones to relative positions on a musical scale based primarily on the frequency of vibration. The just noticeable difference (jnd, the threshold at which a change is perceived) depends on the tone's frequency and is about 4.3 cents (hundredths of a semitone) or about 0.36 Hz in frequency within the octave of 1,0002,000 Hz but within the octave 62125 Hz the jnd is much coarser with some 40 cents or about 2 Hz between perceived pitch changes. The jnd is typically tested by playing two tones in quick succession with the listener asked if there was a difference in their pitches. The jnd becomes smaller if the two pitches are played simultaneously as the listener is then able to discern beat frequencies. The total number of perceptible pitch steps in the range of human hearing is about 1,400; the total number of notes in the equal-tempered scale is 120.

Pitch depends to lesser degree on the sound pressure level (loudness, volume) of the tone, especially at frequencies below 1,000 Hz and above 2,000 Hz. The pitch of lower tones gets lower as sound pressure increases. For instance, a tone of 200 Hz that is very loud will seem to be one semitone lower in pitch than if it is just barely audible. The pitch of higher tones gets higher as the sound gets louder. In this way, pitch perception is like other human senses which respond to the intensity of the stimulus as stated in the Weber-Fechner law.

The relative perception of pitch can be fooled, resulting in "aural illusions". There are several of these, such as the tritone paradox, but most notably the Shepard scale, where a continuous or discrete sequence of specially formed tones can be made to sound as if the sequence continues ascending or descending forever

2. Repetition pitch

A special type of pitch often occurs in free nature when the sound of a sound source reaches the ear of an observer directly and also after being reflected against a sound-

reflecting surface. This phenomenon is called repetition pitch, because the addition of a true repetition of the original sound to itself is the basic prerequisite. The technique is most associated with the twelve-tone technique, created by its "total thematicism" where a tone-row (melody) generates all (harmonic) material. It was also used by Alexander Scriabin, though from a diametrically opposed direction, created by his use of extremely slow harmonic rhythm which eventually led to his use of unordered pitch-class sets, usually hexachords (of six pitches) as harmony from which melody may also be created.

It may also be observed in Igor Stravinsky's Russian period, such as in Les Noces, derived from his use of folk melodies as generating material and influenced by shorter pieces by Claude Debussy, such as Voiles, and Modest Mussorgsky. Bla Bartk's Bagatelles, and several of Alfredo Casella's Nine Piano Pieces such as No. 4 "In Modo Burlesco" the close intervallic relationship between motive and chord creates or justifies the great harmonic dissonance.

3. Definite and indefinite pitch

Not all musical instruments make notes with a clear pitch; percussion instruments are often distinguished by whether they do and or do not. A sound or note of definite pitch is one of which it is possible or relatively easy to discern the pitch. Sounds with definite pitch have harmonic frequency spectra or close to harmonic spectra. A sound or note of indefinite pitch is one of which it is impossible or relatively difficult to discern a pitch. Sounds with indefinite pitch do not have harmonic spectra or have altered harmonic spectra.

It is still possible for two sounds of indefinite pitch to clearly be higher or lower than one another, for instance, a snare drum invariably sounds higher in pitch than a bass drum, though both have indefinite pitch, because its sound contains higher frequencies. In other words, it is possible and often easy to roughly discern the relative pitches of two sounds of indefinite pitch, but any given sound of indefinite pitch does not neatly correspond to a given definite pitch.

4. Labeling pitches

Pitches are often labeled using scientific pitch notation or some combination of a letter and a number representing a fundamental frequency. For example, one might refer to the A above middle C as "A4" or "A440." However, there are two problems with this practice. First, in standard Western equal temperament, the notion of pitch is insensitive to spelling: the description "G4 double sharp" refers to the same pitch as "A4." Second, human pitch perception is logarithmic with respect to fundamental frequency: the perceived distance between the pitches "A220" and "A440" is the same as the perceived distance between the pitches "A440" and "A880."

To avoid these problems, music theorists sometimes represent pitches using a numerical scale based on the logarithm of fundamental frequency. For example, one can adopt the widely used MIDI standard to map fundamental frequency f to a real number p as follows

This creates a linear pitch space in which octaves have size 12, semitones (the distance between adjacent keys on the piano keyboard) have size 1, and A440 is assigned the number 69. Distance in this space corresponds to musical distance as measured in psychological experiments and understood by musicians. (An equal-tempered semitone is subdivided into 100 cents.) The system is flexible enough to include "microtones" not found on standard piano keyboards. For example, the pitch halfway between C (60) and C♯(61) can be labeled 60.5.

5. Scales

The relative pitches of individual notes in a scale may be determined by one of a number of tuning systems. In the west, the twelve-note chromatic scale is the most common method of organization, with equal temperament now the most widely used method of tuning that scale. In it, the pitch ratio between any two successive notes of the scale is exactly the twelfth root of two (or about 1.05946). In well-tempered systems (as used in the time of Johann Sebastian Bach, for example), different methods of musical tuning were used. Almost all of these systems have one interval in common, the octave, where

the pitch of one note is double the frequency of another. For example, if the A above middle C is 440 Hz, the A an octave above that will be 880 Hz (info).

6. Other musical meanings of pitch

In atonal, twelve tones, or musical set theory a "pitch" is a specific frequency while a pitch class is all the octaves of a frequency. Pitches are named with integers because of octave and enharmonic equivalency (for example, C♯and D ♭are the same pitch, while C4 and C5 are functionally the same, one octave apart).

Discrete pitches, rather than continuously variable pitches, are virtually universal, with exceptions including "tumbling strains" and "indeterminate-pitch chants Gliding pitches are used in most cultures, but are related to the discrete pitches they reference or embellish.

7. Pre-19th Century

Until the 19th century there was no concerted effort to standardize musical pitch, and the levels across Europe varied widely. Pitches did not just vary from place to place, or over timepitch levels could vary even within the same city. The pitch used for an English cathedral organ in the 17th century, for example, could be as much as five semitones lower than that used for a domestic keyboard instrument in the same city.

Even within one church, the pitch used could vary over time because of the way organs were tuned. Generally, the end of an organ pipe would be hammered inwards to a cone, or flared outwards, to raise or lower the pitch. When the pipe ends became frayed by this constant process they were all trimmed down, thus raising the overall pitch of the organ.

Some idea of the variance in pitches can be gained by examining old pitchpipes, organ pipes and other sources. For example, an English pitchpipe from 1720 plays the A above middle C at 380 Hz, (info) while the organs played by Johann Sebastian Bach in Hamburg, Leipzig and Weimar were pitched at A = 480 Hz, (info) a difference of around four semitones. In other words, the A produced by the 1720 pitchpipe would have been at the same frequency as the F on one of Bach's organs.

From the early 18th century, pitch could be also controlled with the use of tuning forks (invented in 1711), although again there was variation. For example, a tuning fork associated with Handel, dating from 1740, is pitched at A = 422.5 Hz, (info) while a later one from 1780 is pitched at A = 409 Hz, (info) almost a semitone lower. Nonetheless, there was a tendency towards the end of the 18th century for the frequency of the A above middle C to be in the range of 400 (info) to 450 Hz. (info)

The frequencies quoted here are based on modern measurements and would not have been precisely known to musicians of the day. Although Mersenne had made a rough determination of sound frequencies as early as the 1600s, such measurements did not become scientifically accurate until the 19th century, beginning with the work of German physicist Johann Scheibler in the 1830s. The unit hertz (Hz), replacing cycles per second (cps), was not introduced until the twentieth century.

8. Pitch inflation

During historical periods when instrumental music rose in prominence (relative to the voice), there was a continuous tendency for pitch levels to rise. This "pitch inflation" seemed largely a product of instrumentalists' competing with each other, each attempting to produce a brighter, more "brilliant", sound than that of one's rivals. (In string instruments, this is not all acoustic illusion: when tuned up, they actually sound objectively brighter because the higher string tension results in larger amplitudes for the harmonics.) This tendency was also prevalent with wind instrument manufacturers, who crafted their instruments to play generally at a higher pitch than those made by the same craftsmen years earlier.

It should be noted too that pitch inflation is a problem only where musical compositions are fixed by notation. The combination of numerous wind instruments and notated music has therefore restricted pitch inflation almost entirely to the Western tradition.

On at least two occasions, pitch inflation has become so severe that reform became needed. At the beginning of the 17th century, Michael Praetorius reported in his encyclopedic Syntagma musicum that pitch levels had become so high that singers were experiencing severe throat strain and lutenists and viol players were complaining of snapped strings. The standard voice ranges he cites show that the pitch level of his time,

at least in the part of Germany where he lived, was at least a minor third higher than today's. Solutions to this problem were sporadic and local, but generally involved the establishment of separate standards for voice and organ ("Chorton") and for chamber ensembles ("Kammerton"). Where the two were combined, as for example in a cantata, the singers and instrumentalists might perform from music written in different keys. This system kept pitch inflation at bay for some two centuries.

The advent of the orchestra as an independent (as opposed to accompanying) ensemble brought pitch inflation to the fore again. The rise in pitch at this time can be seen reflected in tuning forks. An 1815 tuning fork from the Dresden opera house gives A = 423.2 Hz (info), while one of eleven years later from the same opera house gives A = 435 Hz (info). At La Scala in Milan, the A above middle C rose as high as 451 Hz (info).

9. 19th and 20th century standards

The most vocal opponents of the upward tendency in pitch were singers, who complained that it was putting a strain on their voices. Largely due to their protests, the French government passed a law on February 16, 1859 which set the A above middle C at 435 Hz. This was the first attempt to standardize pitch on such a scale, and was known as the diapason normal. It became quite a popular pitch standard outside of France as well, and has also been known at various times as French pitch, continental pitch or international pitch (the last of these not to be confused with the 1939 "international standard pitch" described below).

The diapason normal resulted in middle C being tuned at approximately 258.65 Hz (info). An alternative pitch standard known as philosophical or scientific pitch, which fixed middle C at exactly 256 Hz (info) (that is, 28 Hz), and resulted in the A above it being tuned to approximately 430.54 Hz (info), gained some popularity due to its mathematical convenience (the frequencies of all the Cs being a power of two) . This never received the same official recognition as A = 435 Hz, however, and was not as widely used.

British attempts at standardisation in the 19th century gave rise to the so-called old philharmonic pitch standard of about A = 452 Hz (different sources quote slightly different values), replaced in 1896 by the considerably "deflated" new philharmonic pitch at A = 439 Hz. The high pitch was maintained by Sir Michael Costa for the Crystal Palace

Handel Festivals, causing the withdrawal of the principal tenor Sims Reeves in 1877, though at singers' insistence the Birmingham Festival pitch was lowered (and the organ retuned) at that time. At the Queen's Hall in London, the establishment of the diapason normal for the Promenade Concerts in 1895 (and retuning of the organ to A = 439 at 15 C (59 F), to be in tune with A = 435.5 in a heated hall) caused the Royal Philharmonic Society and others (including the Bach Choir, and the Felix Mottl and Artur Nikisch concerts) to adopt the continental pitch thereafter.

In 1939, an international conference recommended that the A above middle C be tuned to 440 Hz, now known as concert pitch. This standard was taken up by the International Organization for Standardization in 1955 (and was reaffirmed by them in 1975) as ISO 16. The difference between this and the diapason normal is due to confusion over which temperature the French standard should be measured at. The initial standard was A = 439 Hz (info), but this was superseded by A = 440 Hz after complaints that 439 Hz was difficult to reproduce in a laboratory owing to 439 being a prime number.

Despite such confusion, A = 440 Hz is arguably the most common tuning used around the world. Many, though certainly not all, prominent orchestras in the United States and United Kingdom adhere to this standard as concert pitch. In other countries, however, higher pitches have become the norm: A = 442 Hz is common in certain continental European and American orchestras (the Boston symphony being the best-known example), while A = 445 Hz is heard in Germany, Austria, and China.

In practice, as orchestras still tune to a note given out by the oboe, rather than to an electronic tuning device (which would be more reliable), and as the oboist may not have used such a device to tune in the first place, there is still some variance in the exact pitch used. Solo instruments such as the piano (which an orchestra may tune to if they are playing together) are also not universally tuned to A = 440 Hz. Overall, it is thought that the general trend since the middle of the 20th century has been for standard pitch to rise, though it has been rising far more slowly than it has in the past.

Many modern ensembles which specialize in the performance of Baroque music have agreed on a standard of A = 415 Hz, an even-tempered semitone lower (rounded to the nearest integer Hz) than A = 440 Hz. (An exact even-tempered semitone lower than A=440 would be 440/21/12=415.3047 Hz.) At least in principle, this allows for playing

along with modern fixed-pitch instruments if their parts are transposed down a semitone. It is, however, common performance practice, especially in the German Baroque idiom, to tune certain works to Chorton, approximately a semitone higher than A-440 (460-470 Hz) (e.g., Pre-Leipzig period cantatas of Bach).

10. Paralinguistics of Pitch

'Paralanguage refers to the non-verbal elements of communication used to modify meaning and convey emotion. Paralanguage may be expressed consciously or unconsciously, and it includes the pitch, volume, and, in some cases, intonation of speech. Sometimes the definition is restricted to vocally-produced sounds. The study of paralanguage is known as paralinguistics. The term paralanguage is sometimes used as a cover term for body language, which is not necessarily tied to speech, and paralinguistic phenomena in speech. The latter are phenomena that can be observed in speech (Saussure's parole) but that do not belong to the arbitrary conventional code of language (Saussure's langue). The paralinguistic properties of speech play an important role in human speech communication. There are no utterances or speech signals that lack paralinguistic properties, since speech requires the presence of a voice that can be modulated. This voice must have some properties, and all the properties of a voice as such are paralinguistic. However, the distinction linguistic vs. paralinguistic applies not only to speech but to writing and sign language as well, and it is not bound to any sensory modality. Even vocal language has some paralinguistic as well as linguistic properties that can be seen (lip reading, McGurk effect), and even felt, e.g. by the Tadoma method.

One can distinguish the following aspects of speech signals and perceived utterances: Speech signals that arrive at a listeners ears have acoustic properties that may allow listeners to localize the speaker (distance, direction). Sound localization functions in a similar way also for non-speech sounds. The perspectival aspects of lip reading are more obvious and have more drastic effects when head turning is involved.

The speech organs of different speakers differ in size. As children grow up, their organs of speech become larger and there are differences between male and female adults. The differences concern not only size, but also proportions. They affect the pitch of the voice and to a substantial extent also the formant frequencies, which characterize the different speech sounds. The organic quality of speech has a communicative function in a restricted

sense, since it is merely informative about the speaker. It will be expressed independently of the speakers intention.

The properties of the voice and the way of speaking are affected by emotions and attitudes. Typically, attitudes are expressed intentionally and emotions without intention, but attempts to fake or to hide emotions are not unusual. Expressive variation is central to paralanguage. It affects loudness, speaking rate, pitch, pitch range and, to some extent, also the formant frequencies.

These aspects are the main concern of linguists. Ordinary phonetic transcriptions of utterances reflect only the linguistically informative quality. The problem of how listeners factor out the linguistically informative quality from speech signals is a topic of current research.

Some of the linguistic features of speech, in particular of its prosody, are paralinguistic or pre-linguistic in origin. A most fundamental and widespread phenomenon of this kind is known as the "frequency code". This code works even in communication across species. It has its origin in the fact that the acoustic frequencies in the voice of small vocalizers are high while they are low in the voice of large vocalizers. This gives rise to secondary meanings such as 'harmless', 'submissive', 'unassertive', which are naturally associated with smallness, while meanings such as 'dangerous', 'dominant', and 'assertive' are associated with largeness. In most languages, the frequency code also serves the purpose of distinguishing questions from statements. It is universally reflected in expressive variation, and it is reasonable to assume that it has phylogenetically given rise to the sexual dimorphism that lies behind the large difference in pitch between average female and male adults. In text-only communication such as email, chatrooms and instant messaging, paralinguistic elements can be displayed by emoticons, font and color choices, capitalization and the use of non-alphabetic or abstract characters. Nonetheless, paralanguage in written communication is limited in comparison with face-to-face conversation, sometimes leading to misunderstandings.

11. Intonation

Intonation is variation of pitch while speaking which is not used to distinguish words. (Compare tone.) Intonation and stress are two main elements of linguistic prosody.

All languages use pitch semantically, that is, as intonation, for instance for emphasis, to convey surprise or irony, or to pose a question. Tonal languages such as Chinese and Hausa use pitch to distinguish words in addition to intonation. Rising intonation means the pitch of the voice increases over time; falling intonation means that the pitch decreases with time. A dipping intonation falls and then rises, whereas a peaking intonation rises and then falls. The classic example of intonation is the question-statement distinction. For example, northeastern American English, like very many languages, has a rising intonation for echo or declarative questions (He found it on the street?), and a falling intonation for wh- questions (Where did he find it?) and statements (He found it on the street.). Yes or no questions (Did he find it on the street?) often have a rising end, but not always. The Chickasaw language has the opposite pattern, rising for statements and falling with questions. Dialects of British and Irish English vary substantially, with rises on many statements in urban Belfast, and falls on most questions in urban Leeds

12. Extralinguistic Features.

Spontaneous speech consists of both lexical and non-lexical (extra-linguistic) elements. Sundaram and Naryanan classify characteristics of spontaneous speech into 3 extra-linguistic groups: paralinguistic cues (falsetto, whisper, creak, laughter giggle, cry/sob etc), disuency patterns (words such as okay, oh, so and well; repetitions pauses such as uh and um) and re- exes (throat clearing, sni_/gulp, toungue clicking, lip smacking and breathing). These extra-linguistic features are considered to be a major factor in discriminating spontaneous speech from written text. Historically, there have been three main ways of viewing extra-linguistic sounds. The oldest view is the one promoted by Chomsky , who claims extra-linguistic sounds to be errors that lie outside language proper. According to this view, extra-linguistic sounds should be excluded from linguistic theory. The second view, states that although these sounds are errors, they are worthy of study for what they reveal about performance. The third view is one that takes some of these extra-linguistic sounds to be genuine parts of a language

Filled pauses are used to provide time for lexical choice-making, planning of the upcoming discourse segment, or to transmit some implicit message, such as a request for attention. The more options there are, the more likely it is that the speaker uses filled pauses. These pauses can thus be indicators of the strength of association between sequential linguistic units, but can also, alternatively, be interpreted in more cognitive

terms as time for choosing among word or phrase options, or for making decisions about the next thought. It is empirically supported that extra-linguistic sounds precede unpredictable lexical items rather than predictable ones. The presence of pauses may thus be an indication of word-searching problems, leading to conclusions such as words following a hesitation has a low transition probability and thus a high information value. This in turn helps listeners detect upcoming important linguistic materials . Clark and Fox show that the filled pauses uh and um are conventional English words, and speakers tend to plan for, formulate, and produce them as they do for any other word. They also report results from experiments on the London-Lund corpus of 170,000 words from 50 face to face conversations and show that there is a difference in their usage, um being followed by much longer delays and pauses than uh, and claim that um is preferably used in connection with larger decision making processes, whereas uh is used when lexical choices are needed. This info

13. Pitch Variety

Pitch is a subjective sensation in which a listener assigns perceived tones to relative positions on a musical scale based primarily on the frequency of vibration. The just noticeable difference (jnd, the threshold at which a change is perceived) depends on the tone's frequency and is about 4.3 cents (hundredths of a semitone) or about 0.36 Hz in frequency within the octave of 1,0002,000 Hz but within the octave 62125 Hz the jnd is much coarser with some 40 cents or about 2 Hz between perceived pitch changes. The jnd is typically tested by playing two tones in quick succession with the listener asked if there was a difference in their pitches. The jnd becomes smaller if the two pitches are played simultaneously as the listener is then able to discern beat frequencies. The total number of perceptible pitch steps in the range of human hearing is about 1,400; the total number of notes in the equal-tempered scale is 120.

Pitch depends to lesser degree on the sound pressure level (loudness, volume) of the tone, especially at frequencies below 1,000 Hz and above 2,000 Hz. The pitch of lower tones gets lower as sound pressure increases. For instance, a tone of 200 Hz that is very loud will seem to be one semitone lower in pitch than if it is just barely audible. The pitch of higher tones gets higher as the sound gets louder. In this way, pitch perception is like other human senses which respond to the intensity of the stimulus as stated in the Weber-Fechner law.

The relative perception of pitch can be fooled, resulting in "aural illusions". There are several of these, such as the tritone paradox, but most notably the Shepard scale, where a continuous or discrete sequence of specially formed tones can be made to sound as if the sequence continues ascending or descending forever.

In Section 3 of this course you will cover these topics:

- Volume
- Rate/Duration

Topic : Volume

Topic Objective:

At the end of this topic student would be able to:

- Discuss the Paralinguistic Factors of Loudness.
- Understand the Syllabic Stress.
- Understand the Projection.
- Discuss the Variation Techniques

Definition/Overview:

The topic discusses that the most immersive interactive environments are those which engage the user in a diverse sensory experience. Sound is now a major medium in most interactive tools and it has a perceptible effect in augmenting a wide range of visual applications. Recently, the importance of sound in creating immersive experiences has emerged as being of comparable importance to visual media. However, most programming tools and multimedia applications are still set up so that the visual dominates and supersedes the acoustic.

Conversely, this paper explores the use of voice-input to control interactive media. Unlike voice-based interfaces which employ voice-recognition or speech-recognition to control an interface, this essay explores the use of paralinguistic features of voice to control interactive applications. The paralinguistic features of voice consist mainly of vocal qualities and emotive vocalisations. Vocal qualities refer to the components that make up the characteristics of voice such as pitch, loudness, duration, and timbre. On the other hand, emotive vocalisations refer to the emotional sounds created by letting air pass through the respiratory system in order to generate non-verbal emotional signals such as sighs, coughs,

laughs, breadth noises, and burps. Some vocalisations may also be produced by obstructing the airflow from the lungs through the mouth and the nose across different areas along the vocal tract.

Vocalisations also include sneezes, tongue clicks, and cracks. The main objective is to allow users to drive all or some aspects in interactive media applications through the real-time use of these paralinguistic vocal signals.

Key Points:

1. The Paralinguistic Factors of Loudness

The linguistic layer carries the semantic information encoded in the language (its grammar and phonological units) and its phonetic representation. It covers what we intend to say, i.e. the "text" of an utterance. The paralinguistic layer of communication is non-linguistic and non-verbal but tells the interlocutor about the speaker's current affective, attitudinal or emotional state. It also includes the speaker's regional dialect and characteristic sociolect. Control of the active participation of time sharing in the conversation is also conducted by this layer (Laver, 1994:21). The paralinguistic markers, however, are in part culturally determined and their respective interpretation must be learned. For example, the use of falsetto voice in English to mimic a male conversation participant's utterance counts as an accusation of effeminacy, whereas in Tzeltal (Mexico) the use of falsetto in greeting someone is a marker of deference (Laver, 1994:22). Paralinguistic communication is not so well structured as its linguistic counterpart; it is not sequential (it is obvious in which sequence the markers are produced) and relative - perception is based more on the fluctuation of a feature rather than on its actual value. The third layer, the extralinguistic behavior, in speech identifies the speaker, his/her sex, age, voice, the range of pitch and loudness as well as his/her habitual factors. In other words, it encompasses all physical and physiological (including all organic features involved in speech production) characteristics of a certain speaker.

2. Syllabic Stress

In linguistics, stress is the relative emphasis that may be given to certain syllables in a word. The term is also used for similar patterns of phonetic prominence inside syllables. The word accent is sometimes also used with this sense.

The ways stress manifests itself in the speech stream are highly language dependent. In some languages, stressed syllables have a higher or lower pitch than non-stressed syllables so-called pitch accent (or musical accent). In other languages, they may bear either higher or lower pitch than surrounding syllables (a pitch excursion), depending on the sentence type. There are also dynamic accent (loudness), qualitative accent (full vowels), and quantitative accent (length, known in music theory as agogic accent). Stress may be characterized by more than one of these characteristics. Further, stress may be realized to varying degrees on different words in a sentence; sometimes the difference between the acoustic signals of stressed and unstressed syllables may be minimal. In English, stress is most dramatically realized on focussed or accented words. In it, the stress-related acoustic differences between the syllables of "tomorrow" would be small compared to the differences between the syllables of "dinner", the emphasized word. In these emphasized words, stressed syllables such as "din" in "dinner" are louder and longer. They may also have a different fundamental frequency, or other properties. Unstressed syllables typically have a vowel, which is closer to a neutral position (the schwa), while stressed vowels are more fully realized. In contrast, stressed and unstressed vowels in Spanish share the same qualityunlike English, the language has no reduced vowels.

(Much literature emphasizes the importance of pitch changes and pitch motions on stressed syllables, but experimental support for this idea is weak. Nevertheless, most experiments do not directly address the pitch of speech, which is a subjective perceived quantity. Experiments typically measure the speech fundamental frequency, which is objectively measurable, and strongly correlated with pitch, but not quite the same thing.)

The possibilities for stress in tone languages is an area of ongoing research, but stress-like patterns have been observed in Mandarin Chinese. They are realized as alternations between syllables where the tones are carefully realized with a relatively large swing in fundamental frequency, and syllables where they are realized "sloppily" with typically a small swing.

Stressed syllables are often perceived as being more forceful than non-stressed syllables. Research has shown, however, that although dynamic stress is accompanied by greater respiratory force, it does not mean a more forceful articulation in the vocal tract.

Some languages have fixed stress. That is, stress is placed always on a given syllable, as in Finnish and Hungarian (stress always on the first syllable) or Quechua and Polish (stress always on the penult: one syllable before the last) or on third syllable counting backwards (the antepenult), as in Macedonian (see: Stress in Macedonian language). Other languages have stress placed on different syllables but in a predictable way, as in Classical Arabic and Latin (where stress is conditioned by the structure of the penultimate syllable). They are said to have a regular stress rule.

French words are sometimes said to be stressed on the final syllable, but actually French has no word stress at all. Rather, it has a prosody whereby the final or next-to-final syllable of a string of words is stressed. This string may be equivalent to a clause or a phrase. However, when a word is said alone, it receives the full prosody and therefore the stress as well.

There are also languages like English, Italian,Russian and Spanish, where stress is (at least partly) unpredictable. Rather, it is lexical: it comes as part of the word and must be memorized, although orthography can make stress unambiguous for a reader, as is the case in Spanish and Portuguese. In such languages, otherwise homophonous words may differ only by the position of the stress (e.g. incite and insight in English), and therefore it is possible to use stress as a grammatical device.

English does this to some extent with noun-verb pairs such as a rcord vs. to recrd, where the verb is stressed on the last syllable and the related noun is stressed on the first; record also hyphenates differently: a rc-ord vs. to re-crd. The German language does this with certain prefixes for example m-schrei-ben (to rewrite) vs. um-schri-ben (to paraphrase, outline) and in Russian this phenomenon often occurs with different cases of certain nouns (земли́/zemli (genitive case of the Earth, land or soil) and зе́мли (soils or lands plural form)).

It is common for dialects to differ in their stress placement for some words. For example, in British English, the word "laboratory" is pronounced with primary stress on the second syllable, while American English stresses the first.

3. Projection

Voice projection is the strength of speaking or singing whereby the voice is used loudly and clearly. It is a technique which can be employed to demand respect and attention, such as when a teacher is talking to the class, or simply to be heard clearly, as an actor in a theatre.

Breath technique is essential for proper voice projection. Whereas in normal talking one may use air from the top of the lungs, a properly projected voice uses air properly flowing from the expansion of the diaphragm. In good vocal technique, well-balanced respiration is especially important to maintaining vocal projection. The goal is to isolate and relax the muscles controlling the vocal folds, so that they are unimpaired by tension. The external intercostal muscles are used only to enlarge the chest cavity, whilst the counterplay between the diaphragm and abdominal muscles is trained to control airflow. Stance is also important, and it is recommended to stand up straight with your feet shoulder width apart and your upstage foot (right foot if right-handed etc) slightly forward. This improves your balance and your breathing. In singing voice projection is often equated with resonance, the concentrated pressure through which one produces a focused sound. True resonance will produce the greatest amount of projection available to a voice by utilizing all the key resonators found in the vocal cavity. As the sound being produced and these resonators find the same overtones, the sound will begin to spin as it reaches the ideal singer's formant at about 2800 Hz. The size, shape, and hardness of the resonators all factor into the production of these overtones and ultimately determine the projective capacities of the voice.

4. Variation

In music, variation is a formal technique where material is altered during repetition: reiteration with changes. The changes may involve harmony, melody, counterpoint, rhythm, timbre or orchestration. Variation forms include ground bass, passacaglia, chaconne, and theme and variations. Theme and variations is a musical form in which the

fundamental musical idea, or theme, is repeated in altered form or accompanied in a different manner. It can be used as a solo piece or as movement of a larger piece. Passacaglias and chaconnes are forms in which a repeating bass line or ostinatotypically shorter than a full-scale variation themeor constantly recurring harmonic progression is heard through the entire piece. Fantasia variation is a form which relies on variation but which repeats and incorporates material freely. Skilled musicians who know a theme well can often improvise variations on it. This was commonplace in the Baroque era, when the da capo aria, particularly when in slow tempo, required the performer to be able to improvise a variation during the return of the main material.

Musicians of the Classical era also could improvise variations; both Mozart and Beethoven made powerful impressions on their audiences when they improvised. Modern listeners can get a sense of what these improvised variations sounded like by listening to published works that evidently are written transcriptions of improvised performances

Topic : Rate/Duration

Topic Objective:

At the end of this topic student would be able to:

- Understand the Speaking/Reading Rate.
- Discuss the Pause effect
- Discuss the Speech Phrasing
- Understand the Variety within speech
- Discuss the Rate Assessments.

Definition/Overview:

The topic discusses that generally a faster speaking speed signals urgency, excitement, passion or raw emotion. It can lead the audience to expect something thrilling is going to occur. They hold their breaths and go for the ride with you. In contrast a slower speaking rate signals importance, seriousness, or significant ideas. It says: 'listen up! You need to know this.' A new concept or new and perhaps, complex sequential information may need to be

delivered slowly so the audience has time to grasp all of the ideas and their consequences before moving on. 'Slow' is also useful for summarizing material.

Key Points:

1. Read or recites part of a text

Read or recites part of a text you know and love quickly. If you can record yourself, do so. If not, listen and note the effect it has on you. If you've recorded yourself, play it back. Ask yourself where was the speed effective? Where was it detrimental? Mark those places on your script. (Use a highlighter: red for fast, blue for go slower) Read again incorporating your changes. Stuck for ideas on what to choose to work with? Try passages from the Bible or the text from a famous speech you know well. If you don't have copies, you can find them easily through a quick search on the net.

2. Pick an information loaded report from a Newspaper or Magazine

Go through it to familiarise yourself with the flow of material and then read aloud. Make a note of which passages need careful or slow reading and which can be taken at a faster rate. Re-read aloud until you feel you have the mix of speeds right. As an extension exercise read the report as if you were reading for an audience who knew nothing about the subject. Note what changes you made and why.

3. Time yourself reading or saying your speech at your normal speaking rate

Note the time down. Now go through again having marked passages for slower or faster treatment. Note the new time.

4. Practice with a partner.

Go through any of the exercises above. Explain what you doing and ask them to listen for effectiveness. Get them to note examples where you did well and where you needed to alter your rate and why.

5. Listen to speakers you admire.

They could be radio presenters, preachers...anybody accustomed to speaking in public. Note the different rates of speech they use over the course of their presentation and the effectiveness of them. (Try to listen to a variety so you have a broad range to draw inspiration from.) Take elements of their rate changes and experiment with them for yourself.

6. Speaking Fast

People talk so fast because others around them do this, because they think erroneously that others will not take the time to listen to them, and because they do not realize the listeners are struggling. In some cultures, speaking quickly is a sign of professional competence. The average speech rate in the Mid-Atlantic States is 120 - 140 words per minute. It is faster in some places such as New York City, and slower in other locales. What matters, is not how many words a speaker can get out, but how many (well-chosen) words are understood by the listener.

Speech rate becomes a problem in any location when the listener does not understand. The speaker either may have to repeat himself, or some information gets ignored. In a healthcare setting, this can be a real danger as instructions may get confused and patient compliance may slip. Young children and senior citizens may also process information more slowly.

If you are listening to a fast speaker, and do not understand the information, ask for the confusing parts to be repeated if possible. Tell the speaker exactly what was unclear, such as "the part after "Ben?s department" ", or "the name of the muscle". Make sure the speaker knows you want to understand. If you are the fast speaker, and you know this is a problem, start by listing the reasons you want to slow down. Maybe you want a promotion, but need clearer speech. Possibly your colleagues are getting frustrated with your speech. Perhaps you are tired of repeating yourself. Possibly phone calls are not returned because others cannot understand your telephone number or name. Remember the picture of your mouth as the driver? When your tongue is high revving it's accelerating away. When you're stuck in first gear, it's crawling forward one little word at

a time. Now you've got control over rate, it's time to turn your attention to the brakes. Skilled use of the brakes is the key to effective silence

7. Reading Rate

Reading is a complex cognitive process of decoding symbols for the purpose of deriving meaning (reading comprehension) and/or constructing meaning. Written information is received by the retina, processed by the primary visual cortex, and interpreted in Wernicke's area. Reading is a means of language acquisition, of communication, and of sharing information and ideas.

Readers use a variety of reading strategies to assist with decoding (to translate symbols into sounds or visual representations of language), and comprehension. Readers may use morpheme, semantics, and syntax and context clues to identify the meaning of unknown words. Readers integrate the words they have read into their existing framework of knowledge or schema (schemata theory). Other types of reading may not be text-based, such as music notation or pictograms. Reading text is now an important way for the general population in many societies to access information and make meaning

Although reading print text is now an important way for the general population to access information, this has not always been the case. With some exceptions, only a small percentage of the population in many countries was considered literate before the Industrial Revolution. With the use of computers where people read news and texts online, a new word has been coined -- "screening" -- to mean "reading" on a computer screen as opposed to reading on a paper surface.

Rates of reading include reading for memorization (fewer than 100 words per minute [wpm]); reading for learning (100200 wpm); reading for comprehension (200400 wpm); and skimming (400700 wpm). Reading for comprehension is the essence of the daily reading of most people. Skimming is for superficially processing large quantities of text at a low level of comprehension (below 50%).

Advice for choosing the appropriate reading-rate includes reading flexibly, slowing when concepts are closely presented, and when the material is new, and increasing when the material is familiar and of thin concept. Speed reading courses and books often encourage

the reader to continually accelerate; comprehension tests lead the reader to believe his or her comprehension is continually improving; yet, competence-in-reading requires knowing that skimming is dangerous, as a default habit.

The table to the right shows reading-rate varies with age , regardless of the period (1965 to 2005) and the language (English, French, German). The Taylor values probably are higher, for disregarding students who failed the comprehension test. The reading test by the french psychologist Pierre Lefavrais tested reading aloud, with a penalty for errors, and could, therefore, not be a rate greater than 150 wpm.

Studies have shown that American children who learn to read by the third grade are less likely to end up in prison, drop out of school, or take drugs. Seventy percent of prison inmates score in the bottom quarter on reading tests. Adults who read literature on a regular basis are nearly three times as likely to attend a performing arts event, almost four times as likely to visit an art museum, more than two-and-a-half times as likely to do volunteer or charity work, and over one-and-a-half times as likely to participate in sporting activities. Literacy rates in the United States are also more highly correlated to weekly earnings than IQ.

Reading requires more lighting than many other activities. Therefore, the possibility of comfortable reading in cafs, restaurants, buses, at bus stops or in parks greatly varies depending on available lighting and time of day. Starting in the 1950s, many offices and classrooms were over-illuminated. Since about 1990, there has been a movement to create reading environments with appropriate lighting levels (approximately 600 to 800 lux).

8. Pause

Verbal pauses are when you say um, ah, uh, you know, etc. While your brain is searching for the next words to say, your mouth keeps on going and blurts out meaningless extra syllables. Verbal pauses also include bridge words like and, but, and so. If you say one of these words and hang on it before you actually know what youre going to say next, its a bridge word. Another form of verbal pause is the repeated word. You keep repeating your last word until you figure out what to say next, like and and and. The person who fills this role changes every week, and his/her job is to count the verbal pauses of everyone who speaks during the meeting and then to tell each of us how we did at the end of the meeting.

The simplest replacement for a verbal pause is a silent pause. When your brain stops feeding intelligible words to your mouth, stop talking. Dont say um, ah, yknow, sooooo, etc. If youre addicted to verbal pauses, this may feel uncomfortable at first, but you eventually get used to it. Remember that you dont have to fill every minute of airtime with noise. Verbal pauses are distracting in communication. They can make you sound less intelligent and clear. They muddle your message. Verbal pauses are simply noise, not communication. You dont need them, and your communication will be more effective once you eliminate them. The way to eliminate verbal pauses in your communication is two-fold: awareness and practice. First, start becoming aware of verbal pauses by listening for them in others spoken communication. If you watch the news or any non-scripted talk show, listen for verbal pauses in the speakers. Its amazing how some people will have very few, and others will have many. The next time you speak, even if its simply in a conversation with coworkers and friends, ask someone to listen for your verbal pauses and to count them. Then at the end of your communication, ask them how you did. You may not even be aware of how this bad habit affects your communication. Once you gain an awareness of where you stand, practice to eliminate verbal pauses. Listen to yourself speak and notice when you blurt out that um, ah, or double-and. Have someone else observe you periodically (even if just to watch you speak in a conversation or on a phone call) to see if youre improving. With awareness and practice you can eliminate this bad habit and improve the clarity of your, uh, verbal communication.

9. Speech Phrasing

Speech Phrasing is a figure of speech is a use of a word that diverges from its normal meaning, or a phrase with a specialized meaning not based on the literal meaning of the words in it. Figures of speech often provide emphasis, freshness of expression, or clarity. However, clarity may also suffer from their use, as any figure of speech introduces an ambiguity between literal and figurative interpretation. A figure of speech is sometimes called rhetoric or a locution.

Scholars of classical Western rhetoric have divided figures of speech into two main categories: schemes and tropes. Schemes (from the Greek schēma, form or shape) are figures of speech that change the ordinary or expected pattern of words. For example, the phrase, "John, my best friend" uses the scheme known as apposition. Tropes (from the Greek tropein, to turn) change the general meaning of words. An example of a trope is

irony, which is the use of words to convey the opposite of their usual meaning ("For Brutus is an honorable man; / So are they all, all honorable men"). During the Renaissance, scholars meticulously enumerated and classified figures of speech. Henry Peacham, for example, in his The Garden of Eloquence (1577), enumerated 184 different figures of speech.

For simplicity, this article divides the figures between schemes and tropes, but does not further sub-classify them (e.g., "Figures of Disorder"). Within each category, words are listed alphabetically. Most entries link to a page that provides greater detail and relevant examples, but a short definition is placed here for convenience. Some of those listed may be considered rhetorical devices, which are similar in many ways.

10. Variety

Your greatest tool as a speaker is your voice. Every time you address an audience your mind, your body and your voice act as partners in the task of getting your message across to your listeners. When you speak, your voice is the primary link between you and your listeners - it's the medium of your message. The importance of having an effective voice isn't restricted to public speaking. A good, controlled voice is an asset in every contact with others. Your voice mirrors your personality with a language all its own - a language that people recognise and respond to immediately. A natural voice that projects an image of cordiality, cultivation and authority is a significant tool for personal success. It can help in gaining promotions, making sales, winning the respect of others and improving your social opportunities, as well as in speaking effectively to audiences. When you speak, your voice reflects your psychological and emotional state of mind. You can't hope to persuade or influence others - or even get them to listen in a positive way - if your tones are harsh, scolding and unfriendly. Such a voice can repel even when the speaker wishes to attract. The quality of friendliness is a prime requisite for a good speaking voice.

Think for a moment about musical woodwind and brass instruments. Their sound comes from the musician's breath and lip vibrations or the vibration of a reed in the mouthpiece. Because the chambers of these instruments differ in size and shape, their tone qualities are distinctive. Different parts of the original tone are increased, or resonated, and other parts are reduced. Human resonation is the increasing or modifying of sounds by the throat, nose and mouth. The sound waves created by the vibration of the vocal cords travel into

the upper part of the throat, then to the mouth and, at times, into the nose. As these waves bounce around within these structures, they are reinforced and amplified. The differences in people's voices arise from the size of the vocal cords and the effects that the resonators (throat, mouth, nasal passages) have on the vocal tone. To a certain extent a speaker can change the size, shape and surface tensions of the pharynx and the oral cavity; he or she may also use, partly use or close off the nasal cavities.

11. Rate Assessments

Vocal habits or tics represent a dysfluency in vocal behavior; the goal of treatment is to decrease the dysfluency and ultimately increase fluency. Habit reversal is an empirically supported procedure to reduce vocal habits and improve vocal fluency. Components of habit reversal may include awareness training, modeling competing responses, practicing competing responses, and reinforcing competing responses. Vocal dysfluency may occur while reading. Brief experimental analyses are recommended for identifying effective reading interventions for children who have difficulties with reading fluency. Common reading interventions include repeated readings, passage previewing, word error correction, reinforced rapid readings, and paired reading. Improved reading fluency is frequently the goal of reading intervention; brief assessments of oral reading fluency are reliable and valid indicators of reading skill and are sensitive to short-term instructional growth. Common reading interventions to improve reading fluency have components similar to habit reversal treatments. Parallels include additional practice with materials (repeated readings), modeling of appropriate responses (passage previewing and paired reading), and practice of correct and competing responses (word error correction).

In Section 4 of this course you will cover these topics:

- Quality
- Articulation

Topic : Quality

Topic Objective:

At the end of this topic student would be able to:

- Discuss the Vocal Quality aspect
- Understand the concept of Paralinguistics

- Understand the concept of Extralinguistics

Definition/Overview:

The topic discusses that the paralinguistic properties of speech play an important role in human speech communication. There are no utterances or speech signals that lack paralinguistic properties, since speech requires the presence of a voice that can be modulated. This voice must have some properties, and all the properties of a voice as such are paralinguistic. However, the distinction linguistic vs. paralinguistic applies not only to speech but to writing and sign language as well, and it is not bound to any sensory modality. Even vocal language has some paralinguistic as well as linguistic properties that can be seen (lip reading, McGurk effect), and even felt, e.g. by the Tadoma method.

Key Points:

1. Perspectival aspects

Speech signals that arrive at a listeners ears have acoustic properties that may allow listeners to localize the speaker (distance, direction). Sound localization functions in a similar way also for non-speech sounds. The perspectival aspects of lip reading are more obvious and have more drastic effects when head turning is involved.

2. Organic aspects

The speech organs of different speakers differ in size. As children grow up, their organs of speech become larger and there are differences between male and female adults. The differences concern not only size, but also proportions. They affect the pitch of the voice and to a substantial extent also the formant frequencies, which characterize the different speech sounds. The organic quality of speech has a communicative function in a restricted sense, since it is merely informative about the speaker. It will be expressed independently of the speakers intention.

3. Expressive aspects

The properties of the voice and the way of speaking are affected by emotions and attitudes. Typically, attitudes are expressed intentionally and emotions without intention,

but attempts to fake or to hide emotions are not unusual. Expressive variation is central to paralanguage. It affects loudness, speaking rate, pitch, pitch range and, to some extent, also the formant frequencies.

4. Linguistic aspects

These aspects are the main concern of linguists. Ordinary phonetic transcriptions of utterances reflect only the linguistically informative quality. The problem of how listeners factor out the linguistically informative quality from speech signals is a topic of current research.

Some of the linguistic features of speech, in particular of its prosody, are paralinguistic or pre-linguistic in origin. A most fundamental and widespread phenomenon of this kind is known as the "frequency code". This code works even in communication across species. It has its origin in the fact that the acoustic frequencies in the voice of small vocalizers are high while they are low in the voice of large vocalizers. This gives rise to secondary meanings such as 'harmless', 'submissive', 'unassertive', which are naturally associated with smallness, while meanings such as 'dangerous', 'dominant', and 'assertive' are associated with largeness. In most languages, the frequency code also serves the purpose of distinguishing questions from statements. It is universally reflected in expressive variation, and it is reasonable to assume that it has phylogenetically given rise to the sexual dimorphism that lies behind the large difference in pitch between average female and male adults.

In text-only communication such as email, chatrooms and instant messaging, paralinguistic elements can be displayed by emoticons, font and color choices, capitalization and the use of non-alphabetic or abstract characters. Nonetheless, paralanguage in written communication is limited in comparison with face-to-face conversation, sometimes leading to misunderstandings.

5. Extra linguistic reality

Extra linguistic reality is reflected by language and influences it. Concepts can change their meaning depending on changes in society, historic epoch, etc. They can even acquire opposite connotations.

Topic : Articulation

Topic Objective:

At the end of this topic student would be able to:

- Understand the Coarticulation.
- Discuss the Individual Differences/Dialects.
- Understand the Program for Improvement.

Definition/Overview:

The topic discusses that in music, articulation refers to the direction or performance technique which affects the transition or continuity on single note or between multiple notes or sounds.

There are many different kinds of articulation, each having a different effect on how the note is played. Some articulation marks include the slur, phrase mark, staccato, staccatissimo, accent, sforzando, rinforzando, and legato. Each articulation is represented by a different symbol placed above or below the note (depending on its position on the staff).

Woodwind and brass instruments generally articulate by tonguing, the use of the tongue to break the airflow into the instrument. Stringed instruments use different bowing techniques to achieve different articulations. When staccato marks are combined with a slur, the result is portato, also known as articulated legato. Tenuto markings under a slur are called (for bowed strings) hook bows. This name is also less commonly applied to staccato or martellato (martel) markings.

Key Points:

1. Coarticulation

Coarticulation in phonetics refers to two different phenomena: The assimilation of the place of articulation of one speech sound to that of an adjacent speech sound. For example, while the sound /n/ of English normally has an alveolar place of articulation, in the word tenth it is pronounced with a dental place of articulation because the following sound, /θ/, is dental. The production of a co-articulated consonant, that is, a consonant with two simultaneous places of articulation. An example of such a sound is the voiceless

labial-velar plosive /k͡p/ found in many West African languages. Co-articulation may also refer to the transition from one gesture to another.

Co-articulation tends to be stronger within syllables rather than across syllable boundaries. This greater co-articulation within syllables is evidence for the cognitive existence of the syllable as a fundamental unit of articulatory organisation. Greater degrees of co-articulation between the phonemes in a syllable increase the perceptual integration of syllables. That is, greater degrees of co-articulation increase the perception that the phonemes in a syllable are connected.

Vowels affect the articulation of adjacent consonants (and adjacent vowels). Consonants affect the articulation of adjacent vowels (and other adjacent consonants).

Some sounds are more resistant to co-articulation than other sounds. This may be due to differences in phoneme duration, differences in the inertia of contrastive articulators, differences in articulator movement distances or it may be due to the phoneme inventory effect (see below).

Co-articulation is greatest (for a given phoneme duration) when there is the greatest articulator movement (ie. greatest distance) between phonemes.

Many consonants have a high tongue position (eg. [k]) and so they are more likely to cause target undershoot in low vowels than in high vowels. This is because there is a longer distance to travel from a high consonant tongue position to a low vowel tongue position than there is to a high vowel tongue position.

Long vowels are more resistant to target undershoot than are short vowels as there is more time for the articulator to reach its target. Accented and stressed vowels are more resistant to undershoot than are unstressed vowels as they are even longer.

2. Individual Differences/Dialects

Not all talkers move the same way when they say the same thing. Despite this fact, popular production theories are often silent about such differences, and the differences themselves exert little influence on usual assumptions about the underlying serial structure of utterances, articulatory features of sounds, or suprasegmental manipulations

related to speaking rate, stress, clarity, syntax, and social context. Standard views about such topics are commonly expressed in speaker-general terms, based upon data from a few speakers. These standard views may reflect central tendencies for the movement tasks speakers must perform, but do little to define the nature and range of variations in speech kinematic behavior.

Individual differences in speech kinematics have been neglected partly because the dominant research problem of the past 30 years the description and explanation of systematic variation in articulatory and acoustic correlates of speech sounds can be addressed using single-speaker research designs. Broad goals associated with this so-called invariance problem include an understanding of control objectives for speech, their plausible internal expression, and the application of muscles, sensors, bones, connective tissue, and the central nervous system, to achieve those objectives rapidly and accurately in real time. Logically, these same goals might also be addressed from a cross-speaker perspective. This path has been chosen less often, probably for technical and practical reasons, rather than scientific ones. Only recently has it become possible to record physiological data of certain types from many speakers. Data storage, processing, and timely reduction have also been barriers.

Some or all of these facts may be related to an observation made by linguists for many years, to the effect that at least two distinct articulatory varieties of /small turned r/ co-exist among speakers. One type is the so-called retroflexed variety, during which the apex of the tongue is (presumably) curled somewhat upward and/or backward in the mouth. The other is the so-called bunched variety, where the apex is held low and drawn rearward while a relatively dorsal portion of the tongue is elevated toward the palate. In both varieties, speakers may achieve similar articulatory effects with the tongue, forming an oral constriction in the mid-palatal region, and exposing an ante/sub-lingual cavity. From a production point of view, /small turned r/ may therefore afford speakers an unusual degree of articulatory latitude. There are different places along the vocal tract where its constriction(s) can be formed, and for one of these places, significantly different ways of forming the constriction

3. Program for Improvement

Without appropriate coarticulation (e.g. in poor synthetic speech) the resulting speech sounds unnatural and is hard to understand. Two opposing forces are at work in speech production and perception. There is a tendency to simplify speech patterns to increase ease of articulation as simpler speech is easier to produce. This is opposed by the competing need ("constraint") to maintain phonological distinctiveness in speech perception. One of the reasons why we accent certain words (eg. words containing new information) is so that we can increase their duration, and therefore can avoid undershoot (ie. more time to reach targets).

Accented words are the words for which we most need to maintain maximum perceptual distinctiveness as we can't rely on context to perceive them. (Constraints maximised). We tend to not accent words representing given information. We don't require maximum perceptual distinctiveness as word identification is assisted by context. (Constraints reduced). Function words carry very little of the semantic load and so they are often reduced as there is no need to maintain their perceptual distinctiveness. (Constraints almost absent) These patterns of accenting certain words (new information) and reducing the other words (given information) in a sentence, provides a pattern of timing that results in the greatest time to reach articulatory targets in accented words and much greater chance of undershoot for other words (especially function words). If we always attempted to maximise phonological distinctiveness in speech perception (without the opposing tendency towards ease of production) then we could perhaps predict the degree to which targets are achieved or undershot from articulator inertia and the timing of each syllable in each accented or unaccented word. However, this isn't always possible.

We vary in the extent to which we maximise distinctiveness from one phonetic context to another. This variation may be language, dialect, local speech community and individual specific. That is, there are different constraints on the extent to which we can relax perceptual distinctiveness and increase ease of articulation. Assimilation is a language (or speech community) specific and also a phonetic context specific relaxation of these constraints. Assimilation constraints may allow some allophones in certain contexts but disallow other allophones in other contexts.

In some cases it may even be possible to increase ease of articulation to a point where a sound assimilates to such a great extent that it becomes more like another phoneme.

In Section 5 of this course you will cover these topics:

- Vowels
- Consonants

Topic : Vowels

Topic Objective:

At the end of this topic student would be able to:

- Discuss the Front Vowels.
- Understand the concept of Back Vowels.
- Understand the concept of Mid Vowels.
- Discuss the Diphthongs.
- Understand the Special Diphthong.

Definition/Overview:

In phonetics, a vowel is a sound in spoken language, such as English ah! [ɑː] or oh! [oʊ], pronounced with an open vocal tract so that there is no build-up of air pressure at any point above the glottis. This contrasts with consonants, where there is a constriction or closure at some point along the vocal tract. A vowel is also understood to be syllabic: an equivalent open but non-syllabic sound is called a semivowel. In all languages, vowels form the nucleus or peak of syllables, whereas consonants form the onset and (in languages which have them) coda.

Key Points:

1. Front Vowels

A front vowel is a type of vowel sound used in some spoken languages. The defining characteristic of a front vowel is that the tongue is positioned as far forward as possible in the mouth without creating a constriction that would be classified as a consonant. In the history of many Indo-European languages, front vowels altered preceding velar consonants, bringing them forward to a palatal, postalveolar, or alveolar place of

articulation. Similar changes, or sometimes ongoing allophonic variation, have occurred in many other languages, including Japanese. See palatalization. This historical palatalization is reflected in the orthographies of several European languages, including the "c" and "g" of Italian, Spanish, and French, the "k" in Norwegian and Swedish, and the "γ" in Greek. English follows the French pattern, but without as much regularity. However, for native or early borrowed words affected by palatalization, English has generally altered the spelling after the pronunciation (Examples include cheap, church, cheese, churn from *[k] yell, yarn, yearn, yeast.

2. Back vowel

A back vowel is a type of vowel sound used in some spoken languages. The defining characteristic of a back vowel is that the tongue is positioned as far back as possible in the mouth without creating a constriction that would be classified as a consonant

3. Mid vowel

A mid vowel is a vowel sound used in some spoken languages. The defining characteristic of a mid vowel is that the tongue is positioned mid-way between an open vowel and a close vowel. The only mid vowel with a dedicated symbol in the International Phonetic Alphabet is the mid central vowel [ə], a symbol which is also used for the vowel schwa. Few languages contrast all three heights of mid vowel, since it is rare for a language to distinguish more than four heights of true front or back vowels. One, the Amstetten dialect of Bavarian German, contrasts four heights of front unrounded, front rounded, and back vowels in addition to having an open central vowel.

4. Diphthongs

In phonetics, a diphthong, (also gliding vowel), is a contour vowelthat is, a unitary vowel that changes quality during its pronunciation, or "glides", with a smooth movement of the tongue from one articulation to another, as in the English words eye, boy, and cow. This contrasts with "pure" vowels, or monophthongs, where the tongue is held still, as in the English word papa.

Diphthongs often form when separate vowels are run together in rapid speech. However, there are also unitary diphthongs, as in the English examples above, which are heard by listeners as single vowel sounds (phonemes).

5. Special Diphthong

Falling (or descending) diphthongs start with a vowel quality of higher prominence (higher pitch or louder) and end in a semivowel with less prominence, like [aɪ̯] in "eye", while rising (or ascending) diphthongs begin with a less prominent semivowel and end with a more prominent full vowel, like [i̯a] in "yard". The less prominent component in the diphthong may also be transcribed as an approximant, thus [aj] in "eye" and [ja] in "yard". However, when the diphthong is analysed as a single phoneme, both elements are often transcribed with vowel letters (/aɪ̯/, /i̯a/). Note also that semivowels and approximants are not equivalent in all treatments, and in the English and Italian languages, among others, many phoneticians do not consider rising combinations to be diphthongs, but rather sequences of approximant and vowel. There are many languages (such as Romanian) that contrast one or more rising diphthongs with similar sequences of a glide and a vowel in their phonetic inventory.

In closing diphthongs, the second element is more close than the first (e.g. [ai]); in opening diphthongs, more open (e.g. [ia]). Closing diphthongs tend to be falling ([aɪ̯]), and opening diphthongs are generally rising ([i̯a]), because open vowels are more sonorous and therefore tend to be more prominent. However, exceptions to this rule are not rare in the world's languages. In Finnish, for instance, the opening diphthongs /ie̯/ and /uo̯/are true falling diphthongs, since they begin louder and with higher pitch and fall in prominence during the diphthong.

A centering diphthong is one that begins with a more peripheral vowel and ends with a more central one, such as [ɪə̯], [ɛə̯], and [ʊə̯] in Received Pronunciation or [iə̯] and [uə̯]in Irish. Many centering diphthongs are also opening diphthongs ([iə̯], [uə̯]).

Some languages contrast short and long diphthongs, the latter usually being described as having a long first element. Languages that contrast three quantities in diphthongs are extremely rare, but not unheard of: Northern Sami is known to contrast long, short and finally stressed diphthongs, the last of which are distinguished by a long second element.

Topic : Consonants

Topic Objective:

At the end of this topic student would be able to:

- Discuss the Bilabial Consonants.
- Discuss the Labio-Dental Consonants.
- Understand the Velar Consonants.
- Understand the Glottal Consonants.

Definition/Overview:

The topic discusses that in articulatory phonetics, a consonant is a speech sound that is articulated with complete or partial closure of the upper vocal tract, the upper vocal tract being defined as that part of the vocal tract that lies above the larynx. Consonants contrast with vowels. Since the number of consonants in the world's languages is much greater than the number of consonant letters in any one alphabet, linguists have devised systems such as the International Phonetic Alphabet (IPA) to assign a unique symbol to each attested consonant. In fact, the Latin alphabet, which is used to write English, has fewer consonant letters than English has consonant sounds, so digraphs like "ch", "sh", "th", and "zh" are used to extend the alphabet, and some letters and digraphs represent more than one consonant. For example, many speakers are not aware that the sound spelled "th" in "this" is a different consonant than the "th" sound in "thing". (In the IPA they are transcribed and θ, respectively.)

Key Points:

1. Bilabial Consonants

In phonetics, a bilabial consonant is a consonant articulated with both lips. The bilabial consonants identified by the International Phonetic Alphabet. The bilabial nasal is a type of consonantal sound used in almost all spoken languages. The symbol in the International Phonetic Alphabet that represents this sound is m, and the equivalent X-SAMPA symbol is m. The bilabial nasal occurs in English, and it is the sound represented by "m" in map and rum. It occurs nearly universally, and few languages (e.g., Mohawk) are known to lack this sound.

The voiceless bilabial plosive is a type of consonantal sound used in many spoken languages. The symbol in the International Phonetic Alphabet that represents this sound is p, and the equivalent X-SAMPA symbol is p. The voiceless bilabial plosive in English is spelled with 'p', as in pit or speed.

[p] is missing from about 10% of languages that have a [b]. (See voiced velar plosive for another such gap.) This is an areal feature of the "circum-Saharan zone" (Africa north of the equator, including the Arabian peninsula). It is not known how old this areal feature is, and whether it might be a recent phenomenon due to Arabic as a prestige language (Arabic lost its /p/ in prehistoric times), or whether Arabic was itself affected by a more ancient areal pattern. It is found in other areas as well; for example, in Europe, Proto-Celtic and Old Basque are both reconstructed as having [b] but no [p].

Nonetheless, the [p] sound is very common cross-linguistically. Most languages have at least a plain [p], and some distinguish more than one variety. Many Indo-Aryan languages, such as Hindi, have a two-way contrast between aspirated and plain [p].

2. Labio-Dental Consonants

In phonetics, labiodentals are consonants articulated with the lower lip and the upper teeth.

The voiceless labiodental plosive is a consonant sound produced like a [p], but with the lower lip contacting the upper teeth, as in [f]. This can be represented in the IPA as [p̪]. A separate symbol not recognized by the IPA that is often seen, especially in Bantu linguistics, is the qp monogram ɋ.

The voiceless labiodental plosive is not known to be phonemic in any language. However, it does occur allophonically. The XiNkuna dialect of Tsonga has affricates, [p̪͡f] and [b̪͡v] (that is, [ɋ͡f] and [ɗ͡v]), which unlike the bilabial-labiodental affricate [p͡f] of German are purely labiodental.

One reason that this sound may be so rare is that a person with uneven upper teeth, or gaps between the teeth, will not be able to completely block the flow of air out of the mouth, and therefore will tend to produce a fricative [f] rather than a plosive [p̪].

The voiced labiodental plosive is a consonant sound produced like a [b], but with the lower lip contacting the upper teeth, as in [v]. This can be represented in the IPA as [b̪]. A separate symbol not recognized by the IPA that is often seen, especially in Bantu linguistics, is the db monogram ɗ.

The voiced labiodental plosive is not known to be phonemic in any language. However, it does occur allophonically: In the Austronesian language Sika, this sound occurs as an allophone of the labiodental flap in careful pronunciation. The XiNkuna dialect of Tsonga has affricates, [p̪f̪] (voiceless labiodental affricate) and [b̪v̪] (voiced labiodental affricate) (that is, [ɓ͡f̪] and [ɗ͡v̪]), which unlike the bilabial-labiodental affricate [p͡f] of German are purely labiodental.

3. Velar consonant

Velars are consonants articulated with the back part of the tongue (the dorsum) against the soft palate, the back part of the roof of the mouth, known also as the velum).

Since the velar region of the roof of the mouth is relatively extensive and the movements of the dorsum are not very precise, velars easily undergo assimilation, shifting their articulation back or to the front depending on the quality of adjacent vowels. They often become automatically fronted, that is partly or completely palatal before a following front vowel, and retracted before back vowels. Palatalised velars (like English /k/ in keen or cube) are sometimes referred to as palatovelars. Many languages also have labialized velars, such as [kʷ], in which the articulation is accompanied by rounding of the lips. There are also labial-velar consonants, which are doubly articulated at the velum and at the lips, such as [k͡p]. This distinction disappears with the approximant [w], since labialization involves adding of a labial approximant articulation to a sound, and this ambiguous situation is often called labiovelar.

The velar consonant [k] is the most common consonant in human languages. The only language recorded to lack velarsindeed, to lack any dorsal consonant at allmay be Xavante. However, there are other languages which lack simple velars. An areal feature of the Pacific Northwest coast is that historical *k has become palatalized in many languages, in many languages becoming [kʲ], but in others, such as Saanich, Salish, and Chemakun becoming [tʃ]. (Likewise, historical *k has become [tʃ’] and historical *x has

become [ʃ]; there was no *g or *ŋ.)However, all three languages retain a labiovelar series [kʷ], [k'ʷ], [xʷ], [w], as well as a uvular series.

Apart from [g], none of the other velars are particularly common, not even [w] and [ŋ], which occur in English. [g] of course does not occur in languages like Mandarin Chinese which lack voiced stops, but it is sporadically missing elsewhere. About 10% of languages which otherwise have [p b t d k], such as Modern Standard Arabic, are missing [g].

The Pirah language has both a [k] and a [g] phonetically. However, the [k] does not behave as other consonants, and the argument has been made that it is phonemically /hi/, leaving Pirah with only [g]as an underlyingly velar consonant. Hawaiian does not distinguish [k] from [t]; the sound spelled k tends toward [k] at the beginnings of utterances, [t] before [i], and is variable elsewhere, especially in the dialect of Ni'ihau and Kaua'i. Since Hawaiian has no [ŋ], and w similarly varies between [w] and labial [v], it's not clear that it is meaningful to say that Hawaiian has velar consonants.

4. Glottal Consonants

Glottal consonants are consonants articulated with the glottis. Many phoneticians consider them, or at least the so-called fricatives, to be transitional states of the glottis without a point of articulation as other consonants have; in fact, some do not consider them to be consonants at all. However, the glottal stop at least behaves as a typical consonant in languages such as Tsou.

The "fricatives" are not true fricatives. This is a historical usage of the word. They instead represent transitional states of the glottis (phonation) without a specific place of articulation. [h] is a voiceless transition. [ɦ] is a breathy-voiced transition, and could be transcribed as [ḥ].

The glottal stop occurs in many languages. Often all vocalic onsets are preceded by a glottal stop, for example in German. The Hawaiian language writes the glottal stop as an opening single quote . Some alphabets use diacritics for the glottal stop, such as hamza <ʔ> in the Arabic alphabet; in many languages of Mesoamerica, the Latin letter <h> is used for glottal stop.