CHAPTER-6

VOICE QUALITY AND VOICE DYNAMICS

6.0 Previous chapters have dealt with the articulatory aspects of the aural medium which describe segments of syllables i.e. vowels and consonants. These segmental features, however do not by any means account for the whole of the medium. The existence of a number of non-segmental components or ingredients in the medium have not yet been properly studied. The present chapter is concerned with non-segmental components of the medium. These components can be divided into two groups: one of them consisting of the components which contribute to the general quality of the voice, and the other consisting of components which arise out of the way the voice is handled. Components of the first group concern voice quality and the second group concern voice dynamics. Thus there are three groups of components in the aural medium:

6.1 Segmental Features

6.2 Features of Voice Quality

6.3 Features of Voice Dynamics

These components are like three strands, all simultaneously and continuously present and together making up the totality of the medium. All three strands must be taken into account:
6.1 SEGMENTAL FEATURES

The strand consisting of the segmental features of an utterance is made up of complex auditory qualities which are in rapid fluctuation, reflecting the rapid succession of movements of the articulators.

6.2 FEATURES OF VOICE QUALITY

The term 'voice quality' refers to those characteristics which exist more or less all the time when a person talks; it is a quasi-permanent quality running throughout the sound produced. The components of voice quality are of two different kinds: a. those which are outside the speaker’s control and b. those which are within the speaker’s control. Some of the components which are outside the speaker’s control are innate and produced by physical characteristics. Among these anatomically derived components are the effects of such things as the bone-structure of the head and chest, the length of the vocal tract from larynx to lips, the size of the tongue, the shape and height of the palate, the size of the vocal cords, and so on. The strikingly obvious differences between the voice quality of a man, the voice quality of a woman and the voice quality of a child are largely the consequences of such physical characteristics; and as a child grows, the changes in its physique are accompanied by changes in voice quality.
There are other components of voice quality which are outside the speaker's control, some of which may be quite temporary. They may arise, for example, from such causes as adenoids, tonsillitis, laryngitis, pharyngitis or a common cold. These and other infections involve inflammation of the tissues of the vocal tract at various points, which usually result in modifying the quality of the sound which the vocal tract conveys.

The remaining components of voice quality are those which are within the speaker's voluntary control and do not derive from his/her physique. They originate in various muscular tensions maintained by a speaker during the time one talks and which keep certain of the organs of speech adjusted in a way which is not their relaxed position of rest. These adjustments give a kind of general 'set' or configuration of the vocal tract which affects the quality of sound.

Other continuing muscular tensions affect the adjustment. The mode of vibration of the vocal cords produces different types of phonation. For example: in 'breathy' phonation, they may be adjusted so that a lot of air escapes through them while they are vibrating and in 'tight' phonation, very little air escapes through them during their vibration. The process of phonation may be modified in other ways. Continuing muscular tensions can hold the larynx in a slightly raised or lowered position in
the throat. Others may constrict the pharynx, producing what is sometimes called a 'pulpit' voice, a kind of voice quality that is also associated with some tenors. Voice quality has a wide range of metaphorical adjectives drawn from various sensory fields. Thus a voice may crack or become dry, flat, hollow, husky, melodious, raucous, rough, tiny or thin. The relative importance of the learnt and the unlearnt in voice quality is difficult to assess. Voice quality is described as a quasi-permanent strand in the medium because it can in fact be altered. Probably most people are capable of making some changes in their voice quality. It is even possible to neutralize, by means of muscular adjustments, the components in voice quality which are anatomically derived, at least to some extent, and perhaps even given enough skill, entirely. There are many professional mimics on stage, radio and television who are able to give convincing imitations of their fellow actors and of public figures initiation in which the performer's even voice quality characteristics are effectively submerged. The ventriloquist, also, has to have command in several voice qualities. The extreme of virtuosity, probably is reached by a certain music hall performer, a larger middle aged man, who had learnt to produce, completely convincingly, the voice quality of a seven year-old girl showing that it is possible to compensate by muscular adjustments, for extreme anatomical differences.
Performances such as this show what the theoretical possibilities are in this direction. In practice, it is sometimes hard to say how much of voice quality is learnt. It is thus an institutionalized feature, common to a group of speakers. There is no doubt that a special voice quality is recognizable as characteristic of certain languages or dialects. In these cases, the learnt components predominate over the unlearnt ones.

6.3 FEATURES OF VOICE DYNAMICS

The third of the three strands is the constituents of the medium grouped under the heading ‘features of voice dynamics’. These features are under the speaker’s control and therefore can be acquired; consequently they tend to be copied from other people and so are capable of characterizing social groups as well as individuals. Some important features of voice dynamics can be considered under the following headings:

i. **Loudness**

ii. **Tempo**

iii. **Continuity**

iv. **Rhythm**

v. **Tessitura**

vi. **Register**

vii. **Pitch Fluctuation**

i. **Loudness:** Loudness, or the scale of the medium depends
on the degree of force with which air is expelled from the lungs by the pulmonic air-stream mechanism while the vocal cords are in vibration - the greater the force, the greater the loudness. The range of loudness of which the human voice is capable is very considerable. It is an easily controlled feature and speakers of all languages adjust automatically and immediately to suit the conditions under which they are talking.

ii. Tempo: Tempo meant is the speed of speaking, which is best measured by the rate of syllable succession. It is a feature, which like loudness, is varied from time to time by the individual speaker. Tempo can be used to express different attitudes. Fast tempo may be associated with anger, drawling or slower than normal tempo has been associated with relaxation. Some people employ more variation than others but everyone has a norm which is characteristic of his/her usual conversational style.

iii. Continuity: It is closely connected with tempo which refers to the incidence of pauses in the stream of speech, e.g., their location of pauses, their frequency and length. The incidence of pauses, whether they are hesitations or deliberate cessations of talking for the purpose of taking breath, seems to be a highly idiosyncratic matter and there is a lot of variation from speaker to speaker.
iv. Rhythm: Rhythm in speech as in other human activities arises out of the periodic recurrence of some sort of movement. The movements concerned in the rhythm of speech are those of the syllable-and stress producing processes which, together, make up the pulmonic air stream mechanism. Speech rhythm is essentially a muscular rhythm and the muscles concerned are the breathing muscles. It is the way in which the chest-pulses and the stress-pulses recur their mode of succession and co-ordination determining the rhythm of a language. Basically there are two different ways in which the chest-pulses and stress pulses can be combined and these give rise to two main kinds of speech rhythm, i.e. (i) syllable-timed rhythm and (ii) stress-timed rhythm. Every language in the world is spoken with either of these kinds of rhythm. In syllable-timed rhythm, the periodic recurrence of movement is supplied by the syllable producing process: the chest-pulses, and hence the syllables, recur at equal intervals of time—they are isochronous. French, Telugu and Yoruba are the examples of syllable-timed languages. In stress-timed rhythm, the periodic recurrence of movement is supplied by the stress-producing process: the stress pulses, and hence the stressed syllables are isochronous. English, Russian and Arabic are stress-timed languages.

When one of the two series of pulses is in isochronous succession, the other will not be. Thus in a syllable-timed
rhythm, the stress-pulses are unevenly spaced, and in a stress-timed rhythm the chest-pulses are unevenly spaced. Consider, however, an utterance in English, a language with a stress timed rhythm:

Which is the train for Delhi, please?

It contains four stressed syllables, which train, Delhi, please and their equal spacing in time can be made apparent by tapping with a pencil on a hard surface simultaneously with these four syllables as the sentence is spoken. The resulting taps will be clearly isochronous. But if one taps on every syllable (there are seven in all), the taps will be unequally spaced, some of them are coming more quickly than others. This is the fact that the number of unstressed syllables which separate the stresses from each other is constantly varying, as is made evident if the stresses are marked off by vertical lines, thus:

| Which is the | train for | Delhi | please? |

It can be seen that which is separated from train by two unstressed syllables, train from Delhi by one and Delhi from please by none: yet the interval of time separating them is the same in each case. The rate of syllable succession has thus to be continually adjusted in order to fit varying numbers of syllables into the same time interval. In other words, there is considerable variation in syllable-length in a language spoken
with a stress-timed rhythm, whereas in a language spoken with
a syllable-timed rhythm the syllables tend to be equally in length.

The rhythm of everyday speech is the foundation of verse in
most languages. Thus French verse is based on syllable-timed
rhythm and English verse on stress-timed rhythm.

v. Tessitura and Register- Tessitura is conveniently borrowed
from the terminology of musicians. ‘Voice’ in its technical
phonetic sense of sound resulting from phonation, is a musical
tone which has a fundamental frequency, and therefore a
recognizable pitch. The pitch of the voice is in continual
fluctuation during speech, but the fluctuations tend to take place
round a central point: if one disregards the occasional extremes,
a speaker has a characteristic range of notes or compass within
which the pitch fluctuation of his/her voice falls during normal
circumstances. The range or compass is the tessitura. It can
vary from person to person (it can be said that someone has a
‘low-pitched voice’ or a ‘high-pitched voiced’ and it is possible
that for everyone there is a tessitura which is best suited to the
strength, size and condition of the vocal cords.

Tessitura, in some communities, is an institutionalized
feature, copied from other people and therefore part of the
characteristic way of speaking a language or dialect; thus native
speakers of Tlingit use a markedly lower tessitura than speakers
of English. The term register, like tessitura, is borrowed from the terminology of musicians. One of its applications there is to a set of organ pipes having a certain tone-quality in common, but it is also used, in a somewhat analogous way, about voices in singing, and this is the source of our application of the term to the speaking voice. Registers of the singing voice are different qualities of sound arising from differences in the action of phonation.

Changes of register occur in many different circumstances. They may be used to express affective indices, signs of emotional states and attitudes of the speaker. Speakers of many languages will typically change register in order to express tenderness or irritability. The same registers carry the same affective indices in different cultures.

There are many languages in the world in which contrasts of register, in addition to providing affective indices are used to provide language-bearing patterns as well. For example, in Cambodian, every syllable is spoken with one of the two registers, which are mainly distinguished from each other by the position of the larynx in the throat. The same is true of Gujarati spoken in Surat, the difference here being between 'tight' and 'breathy' phonation. In various West African languages entire words are spoken with one of the two registers: Nzema, for example, in
Ghana and at least some of the varieties of Ijaw, in Nigeria. Register differences are found in some of the dialects of Scots Gaelic. In many of these languages, register differences are always accompanied by vowel-quality differences or pitch differences or both together.

vi. Pitch Fluctuation: Fluctuation in the pitch of the voice is probably the most important of the features of voice dynamics. It owes its importance partly to its outstanding role as a bearer of affective indices. These indices, together with affective indices conveyed by register differences make up what is meant by the 'tone of voice', and the flow of conversation much depends on them. Voice-pitch fluctuation, in this function, is very similar to gesture. In fact this function of pitch fluctuation might be called vocal gesture. But it owes its importance also to the fact that as well as being a vehicle for indexical signs. It constantly carries language bearing patterns, which operate simultaneously with, and interact with, the language bearing patterns which the segmental features and some of the other dynamic features, carry. Voice-pitch fluctuation thus has both an indexical and a linguistic function and the latter is basic in the sense that the indexical signs are superimposed on the language-bearing patterns. Pitch fluctuation in its linguistic function may conveniently be called speech melody. It is part of the spoken
form of a language. Speech melody is found in all languages, but there is the greatest diversity in the patterns which make it up and in the nature of the linguistic functions it performs. The linguistic functions of speech melody are very varied but of two fundamentally different kinds. In one case, the function of the speech melody patterns is to be part of the structure of sentences; in the other case, their function is to be part of the structure of words. In the former case, the patterns are called intonation and in the latter case they are called tone. In every language the function of speech melody is predominantly either of one kind or the other, so that the languages of the world can be divided into two classes, intonation languages and tone languages. It is only when analysis has established the linguistic function of speech melody in a language that it can be put in one category or the other. English is an example of an intonation language and Chinese is perhaps the most famous of all tone languages. To native English speakers, and other Europeans, tone languages are exotic and strange. Nevertheless a great number of tone languages exists and in fact speakers of tone languages may be in the majority in the population of the world. They are certainly in the majority in Asia and Africa and many tone languages are to be found in North and Central America.

Within the broad division into tone languages and intonation languages, there are further differences to be found in both the
function of the patterns of speech melody and their form. In some tone languages, for example, the patterns may serve to distinguish quite unrelated words which are otherwise alike, while in others the patterns may have a grammatical function, making the distinction between present and past tense in verbs. In a large number of tone languages, the two functions are combined. In intonation languages the speech melody pattern may determine sentence type, such as declarative or interrogative, command or request, dependent, or independent; or may indicate the principal point of information in the sentence. The lexical melody patterns in a tone language are not, of course, absolutely individual for each separate word. The contrasts of pitch on which they are based are limited, so that all the patterns can be analysed and described as arrangements of a small number of elements. Usually there is one tone element to each syllable. These elements may be different pitch levels, which contrast with each other as higher or lower as in many tone languages in West Africa or different pitch movements, which contrasts as rises, falls, or more complex movements such as rise-falls as in many tone languages in South-East Asia.

Both tone patterns and intonation patterns may be modified by the superimposition of vocal gesture. Tone patterns, are
usually not constant, but subject to variation according to the other words next to which they are placed. In addition to that, sometimes it can be found, in what is predominantly a tone language, a component of speech melody that has to be analysed as intonation, i.e. as part of the structure of sentences. This has been described in a number of languages and there are various ways in which the intonation may be related to the tone patterns. Similarly, it can be found in what is predominantly an intonation language a component of speech melody that has to be analysed as tone i.e. as part of the structure of words. This is a well-known feature of Swedish, Norwegian and Slovenian and Lewis dialect of Scots Gaelic. It is possible that languages may exist in which speech melody is made up of tone and intonation in equal proportions.

It can therefore be concluded that all three strands carry indexical signs of social affiliations in roughly equal degrees. Moreover, all three have aspects which are idiosyncratic and characterise the individual speaker recognition. From the linguistic point of view, segmental features are the most important strand in the medium. Hence phoneticians have given much more attention to these features than the features of voice quality and voice dynamics. This chapter will be devoted to giving at least some outlines of this neglected aspect of phonetics.