طرائق دراسة نوع الصوت

Approaches to Voice Quality

م.م عفراء عبد الحسين ناصر

قسم اللغة الانكليزية / كلية الآداب / جامعة المستنصرية

Instructor. Afraa AbdulHussein Nasser (PhD)

Department of English language and literature College of Arts

Mustansiriyah University

Affra_20063@uomustansiriyah.edu.iq

Abstract

Voice quality is a thriving topic. Recently, it has been extensively tackled due to its social, pathological, and forensic importance as well as due to the advances in technical tools that are utilized to investigate the human vocal apparatus. This is not to say that it is a new topic in the linguistic field. It has its roots that extend to the previous centuries. However, it is exposed explicitly as a theory by John Laver (1980). Laver proposed a systematic analysis of voice quality. He stated several features or setting that modulate the voice in a specific manner starting with features in the larynx passing through the pharynx and the tongue ending up to the lips, clarifying the influence of each organ on the modulation of voice. The present paper intends to reveal some of the most essential approaches of how to investigate Voice

Quality, specifically the social variations. There are three main approaches; Auditory or perceptual, articulatory and the acoustic approaches. The importance of this paper lies in the fact that it represents a point of departure for researchers and those who are interested in voice quality to reveal new secrets of the human voice seeking a systematic modulation of human voice as a whole.

Key Words : VQ , approach, articulatory , auditory , acoustic

المستخلص

إن نوع الصوت موضوع مزدهر. وقد تم تناوله مؤخرًا على نطاق واسع نظرًا لأهميته الاجتماعية والمرضية والجنائية وكذلك بسبب التقدم في الأدوات التقنية المستخدمة للبحث في الجهاز الصوتي البشري. وهذا لا يعني أنه موضوع جديد في المجال اللغوي. فله جذور تمتد إلى القرون السابقة. ومع ذلك، فقد تم عرضه كنظرية منظمه لجون لافر (١٩٨٠). اقترح لافر تحليلًا منهجيًا لنوع الصوت. وذكر العديد من السمات أو الإعدادات التي تحور الصوت بطريقة محددة بدءًا من السمات الموجودة في الحنجرة مروراً بالبلعوم واللسان وانتهاءً بالشفتين، موضحًا تأثير كل عضو على تحوير الصوت. تهدف هذه الورقة البحثيه إلى الكشف عن بعض أهم الأساليب للبحث في نوع الصوت، وتحديدًا الاختلافات الاجتماعية. هناك ثلاثة أساليب رئيسية؛ الأساليب السمعية أو الإدراكية، والأساليب النطقية والأساليب الفيزيائية .تكمن أهمية هذه الورقة البحثيه في أنها تمثل نقطة انطلاق للباحثين والمهتمين بنوع الصوت لكشف أسرار جديدة للصوت البشري سعياً إلى تغير منهجي للصوت البشري ككل.

1. Theoretical background of Voice Quality

Voice Quality (hence forth VQ) is not a new phonetic theory; as any other theories, sparks of its inception go back to the seventeenth century. Nonetheless, varied terms have been theorized. Wilkins (1668) is the first who characterize that a linguistic community has specific features of voice as a result of modulation and preference of a certain set of speech sounds. He claims that languages have different pronunciation depending on the nature of speech sounds of a language.

In the nineteenth-century, some phoneticians such as Sweet (1906, p.74) refers to the long-term patterns that characterize the accent as its organic basis or basis of articulation. He supposes that each language has "*general tendencies that govern its organic movement and positions which constitute the basis of articulation, and the acquisition of these organic movements assist greatly in the acquisition of that language*".

Following Sweet, (Abercrombie, 1967, p. 89) considers VQ features as part of speech components that are acquired over time and he assures that such features bears the speakers indexical social and idiosyncratic characteristics. He acknowledges the ignorance of such features at that time. Supporting Abercrombie, (Katamba, 1989, p. 247) highlights the indexical nature of VQ that reflects the social aspects of the speaker as sex, age, social group, regional origins or the social class. He argues that VQ features can hardly be changed by the speaker.

Though such claims have been difficult to prove previously, they provide an attractive theoretical point of departure. Chomsky and Halle (1968) investigate the preparatory state of articulation before speaking assuming that the physiological posture acquired over time is universal. Unlikely, results demonstrate that such physiological posture is language specific. It relies greatly on the language variety the speaker learns. Other researches have been conducted on the Russian language to identify the phonetic principles of 'articulatory base' (Eslin etal, 2019, p. 10) .

Since then, researches have been evolving to investigate the configurations that are specific to communities or language varieties as in the sociophonetic studies conducted by Knowles (1974), Trudgill (1974), and Esling (1978).

2. The Theory of Voice Quality

VQ is the core of speech production. It comprises the vibration of the vocal cords as the source of voice as well as all other modulations of voice that take place within the laryngeal area and all the modifications that ensue as the air passes through the oral vocal tract, and past the lips; i.e., it embraces the vibrations, strictures and resonating chambers that contribute to the speaker's VQ (Esling and Moisik, 2022, p. 237)

Abercrombie (1976) states that VQ are "*those characteristics which are present more or less all the time that a person is talking: it is the quasi-permanent quality running through all the sound that issues from his mouth*" (p.91). This definition conveniently sets the concept of VQ as a phonetic theory; i.e., it is the broader sense of VQ. He declares the notion of 'quasi-permanent' characteristics of speech and differentiates them from the fluctuating segmental and suprasegmental articulation (al, 2019, p. 1) . (Laver, 1980, p. 1) ensures the quasi-permanent quality of voice in his definition as well. He states that "*perceptually, it is a cumulative abstraction over a period of time of a speaker-characterizing quality*".

Esling (al., 2019, p. 1) , following Abercrombie's and Laver's definitions, state that VQ is "*the long-term quality of an individual's voice, the more or less permanent, habitually recurring, proportionately frequent characteristics of a person's speech pattern*". They all confirm that such characteristics are acquired as part of the individual's social and regional upbringing.

VQ is regarded as a property of accent (Esling et al., 2019). Since, accent generally reveals our social background, then it is a social marker (Trudgill, 1974, p. 2) . Individuals are able to change their VQ by altering the phonetic pronunciation of their vowels or consonants; however, their speech will have long –lasting and short-lasting components. The long lasting ones are those markers which are shared by some speakers that typify the linguistic, regional, or the social group of a person. It is identified as a specific marker of that group (Edmondson, 2010, p. 131) .

Hewlett and (Beck, 2006, p. 257), as well, contend that long term adjustments of the quality of phonation not only reflect the physical state of the speaker but also it identifies the norms of a speaker's sociolinguistic community i.e.; it has a powerful informative feature of speech which is utilized by the listener to make judgments about the speaker's social and geographical background, or his personality, overall size, physique, and mental state. Kreiman and Sidtis (2011) affirm that when saying "*it is not who we are that makes us sound as we do; it is how we sound that, over time, makes us who we are*" (Kreiman, p. 354) .

The study of VQ is characterized by specific features and settings. Their systematic classification is first proposed by Laver (1980). He introduces a new view of the vocal apparatus. Traditionally, the vocal apparatus is devoted to explain the phonetics and phonology of segments and suprasegmental features; however, the present theory reveals that the vocal apparatus are utilized in characterizing the identity of a person, a community or a social group through the voice. It presents new functionality and prominence to the larynx that traditionally was overlooked due to the search for phonological systematic features.

Esling and (Moisik, 2022, p. 239) contend that the broader sense of VQ indicates the holistic characteristics of the person's voice that are habitually acquired and learned throughout time. These features of voice start with the larynx to which they attribute a new insight. They consider the larynx as a complex articulator that gives the voice its identity unlike the traditional view of the larynx which considers it solely as a source of voice.

2.1. Features or Settings of voice Quality

There are several features of VQ. Laver (1980) refers to them as 'long time settings' in line with his definition of these settings. The number of these features or

settings varies greatly due to the phoneticians who modify them. The present study relies heavily on Laver's (1980) classification of these features. First; the laryngeal features which are those features that modulate the voice inside the larynx. Second; the Supralaryngeal features which are the features that modulate the voice above the larynx as the pharynx, the tongue, and the lips.

A. **Laryngeal Features (phonation types)**

Four main types of phonations are stated by Laver (1980) that relies basically on the vibration of the vocal folds as well as the constriction of the surrounding cartilages and membrane contained within the larynx.

First is the model or normal voice which is regarded the neutral type of phonation by many phoneticians as (Catford, 1977, p. 32) , (Laver, 1980 , pp. 95; 1991,200) , and Esling et al. (2019, p. 14) . There is quite consensus among phoneticians about the laryngeal characteristics of the model or normal voice for which full glottis is involved and both the ligamental and cartilaginous sections act as a single unit. The vocal folds have a regular periodic vibration with no audible friction that may result from any incomplete closure of the glottis.

Second is folseto. (Hollien, 1974, p. 320) suggests that falsetto occupies the highest fundamental frequency compared with model voice which is moderate. (Laver, 1991, p. 200) demonstrates the compromise in literature on perceptual and physiological description of falsetto. Auditorily, falsetto is characterized by the thin perceptual quality. This quality is derived from the high fundamental frequency along with the mode of vibration of the vocal folds. Physiologically, falsetto is produced when the vocal ligaments are made stiff and immobile. Accordingly, the vocal ligaments with the glottal edge of the vocal folds are all put under strong tension. This situation results in the vertical cross-section of the edges of the vocal fold becoming thin and

the glottis remains slightly apart. Furthermore, the sub-glottal air pressure is lower than the model voice (Laver, 1991, p. 200) .

Third is breathy or whispery voice. Breath is not a new term as it denotes the phonetic state of the sound /h/, aspiration of stop consonants as /p ʰ/, or voiceless fricative (al, 2019, pp. 41-42) . The larynx takes the triangular or (V) shape.

Fourth is Creak which is also known as vocal fry or glottal fry (Laver, 1991, p. 200) or as laryngealized voice (Gussenhoven and Jacobs, 2011, p.20). The essential auditory quality of this phonation type is the effect of separate taps in rapid sequence that has an acoustic effect of a train of discrete laryngeal pulses. Gussenhoven and (Jacobs, 2011, p. 20) demonstrate the auditory effect of creakiness metaphorically as the sound produced when running a fingernail across the teeth of a comb. Vocal fry or creakiness is noticed in British speakers at the end of their utterances. Esling and Edmondson (2010) document that it is a general dialect feature that is associated with RP.

Fifth is harshness that indicates rough auditory quality to a voice. “hoarse, rasping, raucous strident, and pressed voice” are also alternative terms to the rough quality of voice (laver, 1980, p. 128) (al, 2019, p. 77). It results in aperiodic irregularity in the acoustic waveform and noise in the spectrum; i.e., the wave form irregularity is associated with cycle to cycle variation in amplitude and fundamental frequency.

Finally, Ladefoged (1971) thoughts of phonation types as a continuum from the nil phonation and breathiness in which case the glottis is most open to the most closed condition identified as glottal closure as clarified below:

Most open  Most closed

Phonation type: voiceless Breathy Model Creaky Glottal closure

Figure 2.11. The continuum of phonation types.

B. **Supralaryngeal Features**

These include the long term shape of the vocal apparatus above the larynx, they are described as follows:

First; raising or lowering the larynx. The muscles that surround the larynx can be kept up or down throughout speech despite the momentary fluctuations that is resulted from the movement of the vocal cords in the production of speech. (Laver, 1991, p. 191) speculates that the interconnection of the surrounding laryngeal muscles affect the movement of the jaws, tongue, pharynx and the larynx. Sundberg and Nordstorm (1976) have demonstrated that raising the larynx shortens the longitudinal axis of the vocal tract which may results in the shortening of the pharynx which may affect the VQ and vice versa.

Second is the pharyngealized features. It can also be sub-divided according to which part of the pharynx is constricted, either the central part or the lower part of the pharynx. The constriction of the lower part of the pharynx is usually accompanied by the constriction of the upper part of the larynx with the center of the pharynx moving backwards and slightly downwards. Hence, the constriction of the pharynx also involves the shift of the larynx downwards, giving a lowered larynx component, and the velum is also pulled downwards giving extra nasalization. One of the effects of pharyngealization is the expansion of the front part of the vocal tract as the tongue retracts backwards and downwards. As such, velarization might give the same effect to a smaller extent (Laver, 1991, p. 196) .

Third is the velopharyngeal features. This setting is concerned with nasality as an auditory phenomenon. (Laver, 1991, p. 197) explains that it is a vital characteristics that results in audible resonance. Accordingly, the neutral setting involves the absence of nasality except for phonological purposes; i.e., only when nasal segments are articulated. The non- neutral settings involves the extremes of both sides. This involves the prevalent nasality within all segments, stops, fricatives, close or open vowels. (al, 2019, p. 18) claim that the degree of nasality is an acquired feature since it is under the speaker's control and it can be a strong indicator of regional or social background.

Fourth is the lingual features. There are five groups of lingual features. Three of them involve the tip and blade of the tongue, the body of the tongue, and the root of the tongue, whereas the fourth involves the range of movements of the tongue away from the neutral position. The fifth comprises a minor aspect which is the degree of lateral curvature of the tongue surface (Laver, 1991, p. 193). The fourth and fifth are related to overall muscular tension.

The non-neutral settings of the body, tip and blade of the tongue preserve these parts of the tongue displaced from its neutral position within the vocal tract, hence, either extending or constricting the oral cavity of the pharynx.

Fifth is the labial features are more salient visually in VQ, even more than other supralaryngeal settings as the lingual settings (al, 2019, p. 26). The constriction or expansion of the interlabial space in two dimensions of the coronal plane of the lips comprises these features (Laver, 1991, p. 192) classifies the front view of labial settings into nine conditions; neutral, horizontal expansion, vertical expansion, horizontal constriction, vertical constriction, horizontal and vertical expansion, horizontal and vertical constriction, horizontal expansion and vertical constriction, and horizontal constriction and vertical expansion. The combination of these features

results in more other features; however, (al, 2019, p. 26) state the maximally distinctive features describing them as close rounded voice, open rounded voice and spread lip voice.

Sixth is the mandibular settings interact with all the supralaryngeal settings. It is possible visually to distinguish the close jaws settings from open jaws or protruded jaws. Accordingly, voices are described as close jaw voice, open jaw voice and protruded jaw voice. (Eslin etal, 2019, p. 25) add that the extreme version of close jaw is clenched jaw. (Honikman, 1964, p. 75) studied jaw features of some languages stating that; "*Indian languages and Pakistan languages have their jaw held inert and open in continuous speech*".

3.Approaches to Voice Quality

Several approaches proposed for the investigation of VQ features either socially to clarify the social effect on the modulation of the individual's VQ, or pathologically to identify some sort of sickness in voice. However the main approaches that are characterized are the auditory or perceptual, articulatory and the acoustic approaches. Some studies focus on the broader sense of VQ utilizing the perceptual or auditory method; whereas other studies preferred to focus on narrower sense of VQ; i.e., investigating specific feature as in the phonation types or other specific feature as nasalization only.

3.1. Perceptual or Auditory Approach

VQ is essentially perceptual in nature (Kreiman, Gerratt, Kempster, Erman, &Berke, 1993). Therefore, perceptual approach is regarded the "golden standard" to which articulatory and acoustic approaches might be compared to (Segundo et al, 2019). Furthermore, Garellek (2019) assumes that "the study of voice should be driven by what humans can hear"(p.102). Additionally, perceptual approach is rapidly

performed, non- invasive and attainable because they do not demand extensive instrumental equipment as other articulatory and acoustic approaches. However, they are highly subjective. Therefore, they require extensive training in order for the results to be reliable. On the other hand, perceptual approach acquires great attention and it prevail for the assessment of VQ features (San Segundo et. al., 2019).

Perceptual analysis refers to the auditory perception of VQ features. Quality of voice has been identified as one of the basic domains that differentiate voices perceptually in addition to the other three domains as duration, pitch, and loudness (Laver, 1994, p. 27) . The perceptual assessments of VQ features are necessary in various areas of research as sociolinguistic studies, forensic studies and voice therapy. A number of protocols have been proposed to detect those features of social groups or those who have vocal pathological conditions. Such protocols typically demand trained listeners to rate VQ features by employing scalar degrees, interval scales, or visual analog scales. Forensic phoneticians, as well, make use of perceptual protocols to assess and analyze the voice recording of an offender to compare it with the voice sample of a suspect. Perceptual protocols are also prevalent in sociophonetic studies. As a result, a thorough description of several verities of English and other languages is ensued reflecting the association of VQ features with social variables as gender, age and social class (San Segundo et. al. 2019).

Perceptual protocols require the listener to document diverse aspects of speaker's VQ in a systematic way (Carding, P., Carlson, E., Epstein, R., Mathieson, L., and Shewell, C., 2001). The number of specific VQ features and the employed rating system varies greatly across protocols; this follows the research purpose or interest. However, most of them are originated in a clinical context focusing on the narrow sense of VQ targeting the non-neutral laryngeal features that characterize different forms of dysphonia and quantifying their severity. Kreiman et al. (1993) identify

around sixty varied perceptual protocols in the USA, the Buffalo III Voice Profile being the most widely used. On the other hand, most common perceptual protocols are: the Vocal Profile Analysis Scheme (VPA) , the Hammarberg scheme , the Buffalo Voice Profile Analysis(BVP), and the GRBAS scale. Combining two or three methods are also appropriate relying on the research objectives. Nonetheless, the GRBAS scale is recommended fundamentally for assessing voice clinically. It has been demonstrated that, on the basis of low intra-rater and inter-rater variances, the GRBAS scale parameters seem to be the most reliable and relevant perceptual VQ ratings. It assesses voices on scales for essentially the same qualities. The grades are equivalent to overall severity, roughness, breathiness, weakness, and strain. It uses four-point scales.

VPA, on the other hand, was first proposed by John Laver, Sheila Wirz, Janet Mackenzie Beck and Steve Hiller in the 1980s. This scheme is primarily based on the principles developed for the description of normal voices; however, its use extends to the description of more extreme voices encountered in speech pathology. Accordingly, it is regarded as a unified technique for the description of both normal and pathological aspects of VQ (Laver, 1991, p. 12). Laver (1980) stated that though these features have auditory foundations but they "*all have correlates on the three levels of analysis; articulatory, physiological and acoustic levels*" (p. 7).

Two further protocols have been designed more recently; the Stockholm Voice Evaluation Approach (SVEA) (Hammarbrtg 2000), and the Consensus Auditory Perceptual Evaluation (CAPE-V) (Kempster et. al. 2009). This protocol rates voices on visual analog scales for overall severity, roughness, breathiness, strain, pitch, loudness, and any additional scales the clinician may wish to add (Kreiman et. al. , 1993).

The major motives why these protocols are created are: (i) to obtain a comprehensive overview of the characteristics of a voice; (ii) to compare the information provided by the perceptual VQ evaluation with other voice assessment methods; (iii) as a basis for planning and monitoring therapy; and (iv) as a way of facilitating communication with other professionals. The choice of one scheme than the other is basically dependent on the research question (Beck 2005). Most of the earlier mentioned protocols are designed to detect vocal pathological conditions or forensic purposes. This includes the VPA protocol. However, the VPA is frequently used for sociophonetic studies as those of Esling's (1978) and Smith's (1993). It has appeared in a wider range of linguistic fields (San Segundo et. al. 2019).Furthermore, Beck (2018) stated that it is one of the most comprehensive schemes as it assesses the overall VQ features; supralaryngeal and laryngeal VQ features compared with the other perceptual protocols that seem to focus on the narrow sense of VQ. Some perceptual studies examine VQ and ethnicity (Trent, 1995, pp. 29–36). Other studies examined the relations of VQ and social factors as gender, age or social class have been implemented by Laver and Mackenzie Beck (2007), Stuart-Smith (1999), Mackenzie Beck and Schaeffler (2015).

3.2. Articulatory Approach

The articulatory approach to VQ is highly technical, i.e., it needs specific equipment as well as some experience with such equipment to be conducted in the appropriate manner. Nonetheless, the results would be highly valid and reliable.

Investigating the articulatory and physiological features of VQ requires specific techniques to image the larynx and examine the long term VQ. Earlier techniques as electroglottography (EGG) has been used to investigate the various phonatory patterns and it succeeded in distinguishing between phonatory changes and the vibration of various areas which are associated with the constriction in the lower vocal

tract. X-ray speech data was utilized and were popular from 1960s to the mid-1980s. It provided two dimensional projection of the vocal tract. Furthermore, X-ray images in Russell (1931) and Negus (1949) showed the entire vocal tract. Other techniques are accessible recently as in laryngoscopy, laryngeal ultrasound, cineradiography and Magnetic Resonance Imaging (MRI) which proved pivotal in illustrating the epilarynx vibratory behavior in speech as well as its association to the other areas of the vocal tract. However, such tools need experience and precautions when tackled. Furthermore, there are computer models of the larynx which stimulate the laryngeal constriction, as in a 3D Laryngeal constrictor model or a model of Vocal-Ventricular Fold Contact (al, 2019, pp. 83-87). Most articulatory studies can be conducted on small scale studies as in case studies on individuals. The reason is that not all people would confirm their consent on using such tools on them as they are afraid of getting infected of some disease as the Corona Virus or other types of disease. Moreover, these equipment are highly medical and they are mostly consumed in detecting medical conditions in the vocal apparatus and rarely used for vocal research specifically in Iraq.

One of the prominent articulatory models is that of Esling et. al.(2019): The Larngeal Articulator Model (LAM). In this model, the larynx is considered a complex articulator equivalent to the tongue. It is regarded as a constrictor mechanism that has great effect on voice production. Hence, the vibration of the vocal cords is not the only mechanism that is basic for the production of voice. The aryepiglottic trilling has an important role in producing phonemes as well as in modifying voice. They stated that unlike the idea that the tongue is the most complex articulator responsible for the production of sounds and the larynx is only the location of the vocal folds which are the only source of voice, " *all the pharyngeal and epiglottal sounds, pharyngealization effects, epiglottalization effects, laryngealization, glottalization and other lower vocal tract effects are all produced by the action of the aryepiglottic constrictor mechanism*

accompanied by tongue retraction and larynx raising" (p.5) As such, the larynx with muscles above the vocal folds is a complex articulator as well. Furthermore, it "*redefines the relationship between the vibratory structures in the laryngeal space and identifies a set of resonating cavities that shapes spectral energy in ways that are predictable based on the inherent folding properties of the articulator.* (p.5). The tongue is no more the only complex articulator as its retraction to the laryngeal articulation represent a specific relationship to the laryngeal articulator.

The LAM is fundamentally designed to interpret the VQ strand of accent. It assumes that there are other sets of folds that can produce a specific phonation types in addition to the vocal folds that produces voice. These are the ventricular folds and the aryepiglottic folds. These two other folds can enrich the aerodynamic / mechanical actions of the vocal fold vibration to produce , for instance, a harsh voice or a style of singing… etc. The aryepiglottic folds are the place of articulation of the pharyngeal sounds. Other sounds called epiglottals are pronounced in the same place of articulation because there is not any other place in the lower vocal tract that can form a stricture. Accordingly, all the three sets of folds can produce signals that are regarded as source of phonation (Eslin etal, 2019, p. 8) .

The LAM presents new complexities of voice that can add to the linguistic, paralinguistic and extralinguistic meaning. These complications were not interpreted previously. However, with LAM which shed lights on the complexities of articulation of the lower part of the vocal tract, researchers can describe the core of VQ even of infants as they are learning to speak because infants, as revealed recently, acquire the VQ features of their mother tongue before even pronouncing their first word (Eslin etal, 2019, p. 9) .

The articulatory approach presents new horizons for many researches on VQ features, specifically on Arabic language since if contains some laryngeal phonemes

which clarify clearly the role of the larynx as a complex articulator equivalent to the tongue and shows the vibratory patterns other than the vocal folds. Many case studies, social studies, forensic studies or pathological studies can be conducted utilizing the articulatory approach as the LAM as well as the techniques mentioned earlier to dig deeper and reveal the capabilities of the human apparatus to produce and modify voice.

3.3. The Acoustic Approach

Acoustic analysis of VQ is the most prevalent one as researchers and phoneticians are racing to detect the acoustic parameters that correlate with each feature of VQ whether supralaryngeal or laryngeal features. Laver (1980) impressionistically detects some of these parameters when he proposed his theory.

Thomas (2011) and other phoneticians alluded that the focus of sociophonetics is on the acoustic area of study due to the technological advancement of techniques as PRAAT (1992-2009) or PLOTNIK (2009). Most significantly, such techniques assist researchers to detect more reliably the acoustic parameters that correlate with supralaryngeal as the height of the larynx, as it affects the length of the pharynx, the height of the tongue, whether the soft palate is raised or lowered which denote the closing or opening the nasopharyngeal port and the configuration of the lips. Moreover, there are other acoustic parameters for the laryngeal features as breathiness, creakiness or harshness. Hence, acoustic measurement gains more valid and reliable results and confirm the results that are detected perceptually.

In measurement, it is crucial to decide the amount of speech to be measured in VQ as it varies in relation to specific feature to be measured. For instance, nasality and the fundamental frequency are pervasive features that extend on long amount of utterance, accordingly they demand longer sample of speech. Contrarily, Creakiness

is confined to the last syllable before a pause, i.e., the glottal pulses that indicate creak in the speaker's voice is only detected in that context. As such, the amount of the recorded speech should correlate with the specific research interest of VQ feature (Thomas, 2011, p. 225) .

There are many acoustic parameters of VQ features that can be measured. As for the laryngeal features, the fundamental frequency (F_0 hence forth) range can be measured since people differ in using the range of F_0, however, researchers should isolate the social variation from the idiosyncratic variation of that range. 'Key', which is the degree of F_0 variation in one's voice, can reveal social variation, for instance, studies reveal that white females in USA use a wider key than African American females or even males of both ethnic group. 'Register' which is concerned with the average F_0, can gauge dialectal difference. However, researchers have to distinguish the social variation of F_0 register from physiologically controlled influences (Thomas, 2011, p. 225) . Another acoustic technique to measure VQ features is the Long-Term Average Spectrum (LTAS) analysis which is an early quantitative acoustic approach to characterize long-term VQ. It captures both the supralaryngeal as well as the laryngeal effects; however, it is specifically useful as a way of interpreting oral resonance postures (Eslin etal, 2019, p. 83) .

Phonation types can be measured acoustically as well by comparing the pulses in the waveform and the spectrogram of all irregular phonations as breathiness, harshness and creakiness with those of the model voice or identifying the conditions or states and amplitude of pulses; for instance, breathiness can be detected in the waveform as the pulses are better defined than for model voicing and on the spectrogram, the glottal pulses have a fuzzy appearance as the vocal folds fail to close. Harshness is characterized acoustically by the irregularity in the amplitude and frequency of glottal pulses. Creakiness, on the other hand, is characterized by slowing of the glottal

pulses and the amplitude of pulses alternating between high and low. These measurements are all accompanied by measurement of $F_{\circ}$ for each phonation type. Actually, these phonations can be controlled by the speaker and they may convey social meanings. Various acoustic techniques can be utilized to measure phonation types. However, methods that measure breathiness may not be similar to those used for measuring harshness or creakiness though there might be some exceptions as the spectral tilt that can be used to measure various phonations (Thomas, 2011, pp. 227-228) .

Supralaryngeal features can be measured acoustically as such. Lindblom and Sundberg (1971) extract measurement of larynx height by comparing formant values of modal voice and lowered larynx voice of a male producing Swedish vowels. The results showed that F_1 falls consistently by about 5 percent when the larynx is lowered. Similar results are extracted from other measurement of F_1. Furthermore, changes in the distances between formants as F_1 and F_2 or between F_5 and F_4 can be used to identify larynx height setting.

Concerning tongue settings, (Laver, 1980, p. 48) stated that such setting is not necessarily be pathological; these settings might characterize a dialectal variation. In accordance with Laver, Stuart-Smith (1999) found rich correlations of tongue setting with gender, age, and social class. However, her study was basically auditory. Nonetheless, Nolan (1983) provides acoustic measurement of mean formant values detected for lingual settings. In short, fronting of the tongue results in a general raising of F_2 values and a fall in F_1 and retraction of the tongue reflects the opposite; i.e., a rise in F_1 and a fall in F_2. However, the exact technique of measuring these settings acoustically needs experience and a lot of computing to be implemented (Thomas, 2011, pp. 241-242) .

As for labial settings, protrusion and constriction of the lips may imply the same effect acoustically which is lowering of all formants, whereas lip spreading may raise formants slightly. This raising and falling all compared with neutral lip setting.

Nasality, as well, can be gauged acoustically by several methods. An older method is to subtract the amplitude of the first harmonic whose frequency is F_0 from the amplitude of F_1. This value would be higher for non-nasal vowels than for nasal vowels as nasalization increases the bandwidth of F_1 and consequently decreasing its amplitude. However, this method has been substituted by several updated methods.

Acoustic parameters of VQ settings are numerous and continuously updated as in investigation jiter, shimmer or capestral peak and many other acoustic parameters which identify VQ features or settings acoustically. Some meticulous voice features have been investigated acoustically by Harmegnies (1987) and Harmegnies et al (1989) on the distribution of the energy peak in the spectrum integrated over time, in which he relies on the phonetically balanced texts to compare long-term settings. Other acoustic studies have been conducted on bilingual speakers as Bruyninckx et al. (1994). There are researches on gender styles and speaker identity as Podesva, 2007 and Podesva and Callier 2015, and ultrasound studies as Schaeffler and Scobbie 2010 (Eslin etal, 2019, p. 84) .

Conclusions

In a nutshell, VQ is 'uncharted territory in sociolinguistics' as (Thomas, 2011, p. 224) puts it. It is still a mysterious domain and regarded a fresh area for researchers to investigate. It is a vast field of research because new concepts and theories in the production of voice are yet to be fully investigated. The phonetic approaches stated earlier in this paper represent a point of departure to dig deeper in order to depict the new capabilities of human vocal apparatus to produce and modify

voice. Sociolinguistic, forensic and pathological studies of VQ are still evolving specifically with the advancement of articulatory and acoustic equipment that will enhance and provide objective measurement of various VQ features.

References

(Abercrombie, D. (1967) Elements of general phonetics, Edinburgh: Edinburgh University Press.)

(Beck, J. M. (2018). Vocal profile analysis scheme: a user's manual. Edinburgh: Queen Margaret University.)

(Beck, J. M.(2005). Perceptual analysis of voice quality: The place of vocal profile analysis. In Hardcastle, W. J., and Beck, J. M. (Eds.), A figure of speech: A festschrift for John Laver (pp.285-322). New Jersey: Lawrence Erlbaun Associates.)

(Catford, J. C. (1977). Fundamental problems in phonetics. Edinburgh: Edinburgh University Press.)

(Esling, J. H. (1978). Voice quality in Edinburgh: A sociolinguistic and phonetic study.)Doctoral Dissertation(Edinburgh: University of Edinburgh.)

(Esling, J. H. (1978). The identification of features of voice quality in social groups. Journal of the International Phonetic Association. Cambridge: Cambridge University Press, 8(1-2), 18-23. Retrieved from http://journals.combridge.org/abstract_S002510)

(Esling, J.H., and Edmondson, J. A. (2010). Acoustical analysis of voice quality for sociophonetic purposes. In M. Di Paolo and M. Yaeger-Dror (Eds), Sociophonetics: A student's guide (pp.131-148). London: Routledge.)

(Esling, J.H., Moisik, S. R., Banner, A., Buchman, L. C. (2019). Voice quality: The laryngeal articulator model, Cambridge: Cambridge University Press.)

(Esling, J. H., and Moisik, S. R. (2022). Voice quality. In R. A. Knight and J. Setter (Eds). The Cambridge Handbook of phonetics (pp.237-309). Cambridge: Cambridge University Press.)

(Gussenhoven, C., and Jacobs, H. (2011). Understanding phonology. Great Britain: Hodder Arnold.)

(Hewlett, N. and Beck, J. (2006). An introduction to the science of phonetics. New Jersey: Lawrence Erlbaum Associates.)

(Hollien, H. (1974). On vocal registers: a proposal. Journal of Phonetics. 2(2). 125-143. doi: 10.1016/s0095-4470(19)31188-x.)

(Honikman, B. (1964). Articulatory settings. In D. Avercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott and J. L. M. Trim (Eds). In honour of Danial Jones (pp.73-84). London: Longman. Retrieved at:)

(Katamba, F. (1989). An Introduction to Phonology. UK: Longman Group Limited.)

(Kreiman J, Gerratt BR, Kempster GB, Erman A, and Berke G. S. (1993) Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. Journal of Speech Hearing Research, 36(1), 21-40. doi:org/10.1044/jshr.3601.21.)

(Kreiman, J. and Sidtis, D. (2011). Foundation of voice studies: An interdisciplinary approach to voice production and perception, UK: Wiley-Blackwell Publishing.)

(Ladefoged, P. (1971). Preliminaries to Linguistic Phonetics, Chicago: The University of Chicago Press.)

(Laver, J. (1980). The phonetic description of voice quality. Cambridge: Cambridge University Press.)

(Laver, J. (1991). The Gift of Speech, Edinburgh: Edinburgh University Press.)

(Laver, J. (1994). Principles of phonetics, Cambridge: Cambridge University Press.)

(San Segundo, E., Foulkes, P., French, P., Harrison, P., Hughes, V., & Kavanagh, C. (2019). The use of the Vocal profile analysis for speaker characterization: Methodological proposals. Journal of the International Phonetic Association, 49(3), 353-380.)

(Sweet, H. (1906). A Primer of Phonetics (3rd ed, revised), Oxford: Clarendon Press.)

(Thomas, E. R. (2011). Sociophonetics: An introduction, USA: Palgrave Macmillan.)

(Trent, S. A. (1995). Voice quality: Listener identification of African-American versus Caucasian speakers. Journal of the Acoustical Society of America 98 (5), 29-36.)

(Trudgill, P. (1974). Sociolinguistics: An introduction to language and society, England: Rowland Phototypesetting Ltd.)