Voice quality and variation: a pilot study of the Liverpool accent

Marion Coadou

Laboratoire Parole et Langage, CNRS UMR 6057, Université de Provence, Aix-en-Provence, France marion.coadou@lpl.univ-aix.fr

Abstract

Voice Quality is a concept which is quite difficult to define. This study proposes a description and a definition of it according to John Laver's model. Moreover, very few studies focused on the variation of voice quality according to accents. This pilot study aims at describing the voice quality of the accent of Liverpool thanks to a perceptual analysis called the *Vocal Profile Analysis* scheme. The results show that the VPA scheme is a useful tool in order to determine some features of the voice quality of the Liverpool accent.

Finally, to our knowledge, there is no study that compares voice quality across several accents of the British Isles. This is why, this study is part of a larger project that will try and identify the voice quality features of some British accents using the VPA scheme.

1. Introduction

Voice quality is a term which is frequently used by phoneticians. Defining the concept precisely is, however, quite difficult. The voice quality of a speaker is the result of the interaction between organic and phonetic factors ([1] and [5]). The organic factors may refer, for example, to the size or the shape of the vocal tract. The phonetic factors, which are studied here, can be due to muscular adjustments learnt by the speakers in their social environment.

First of all, this study proposes a definition of some keyconcepts in order to understand voice quality and a review of the literature on the subject. Then, a corpus is analyzed thanks to the Vocal Profile Analysis Scheme. This pilot study on four subjects from Liverpool shows that it is possible to observe variations of voice quality between various accents of the British Isles.

2. Definition of the Voice Quality System

Several studies have attempted to describe and define voice quality. Few however, have envisaged voice quality as a whole. John Laver was aware of this lack of classification and of an adequate descriptive model. Thus, his definition of voice quality integrates the two basic notions of "setting" and "phonation type". His description is based on auditory and on acoustic data. He established a system (see figure1) of relationships between the different settings of voice quality.

Figure 1 is a representation of the voice quality system according to Laver (1980), Laver (1991). Voice quality can be separated into two main categories, supralaryngeal settings and laryngeal settings, also called phonation types.



Figure 1: Main Settings of Voice Quality according to Laver's model

2.1. The Supralaryngeal Settings

These settings can be divided into three types of settings, each of them being further divided into "sub-settings":

The longitudinal settings are concerned with vertical modifications of the vocal apparatus. Changes can occur at two levels: at that of the lips (labial setting) resulting in labial protrusion or labiodentalization or at that of the larynx, (laryngeal settings) giving a raised or a lowered larynx voice.

The latitudinal or cross-sectional settings refer to horizontal modifications that can occur at several points along the vocal apparatus. For instance, the labial settings can result in lip rounding or spreading. The lingual settings can concern the tip or blade of the tongue and the lingual body or root. The latitudinal settings can also be located at the level of the jaw (mandibular setting) resulting in subsettings such as close jaw, open jaw or protruded jaw.

Finally, the velopharyngeal settings describe modifications at the level of the nasal cavity. The resulting voice qualities are simpler because the voice can be either nasal or non-nasal. Yet, in reality, a speaker's voice will probably have different degrees of nasality or denasality.

2.2. The Phonation Types

The laryngeal settings are also referred to as phonation types. Different types can combine according to specific rules. The laryngeal activity depends on a quite complex interaction between tension of the vocal folds and distribution of the airflow. Laver (1980) established three parameters in order to describe the different phonation types: adductive tension, medial compression and longitudinal tension. Thus modal voice has moderate values for the three parameters whereas falsetto has a high adductive tension, a large medial compression and also a high longitudinal tension. This is why

modal and falsetto cannot combine, since they require opposite values for the three parameters. Table 1 sums up the three parameters for each phonation type according to Laver (1980) and Laver (1991).

	Medial compression	Longitudinal tension	Adductive tension
modal voice	moderate	moderate	moderate
falsetto	high	high	high
whisper	moderate to high	moderate	low
creak	high	low	high
harshness	extreme	moderate?	extreme
breathiness	low	low	minimal
tense voice	high	high	high
lax voice	low	low	low

Table 1 : Description of the three parameters for each phonation types according to Laver's model

Thanks to the descriptive model introduced by Laver (1980) a classification of different types of voice quality is now available. Yet one must take into account the fact that the voice of an individual is the result of a complex network of supralaryngeal and laryngeal settings.

3. Literature Background.

Few studies have treated voice quality variations in English. They can be classified into two categories. They were either based on impressionistic remarks or on quantifiable data.

3.1. Impressionistic Remarks on Voice Quality

Knowles (1978) and Trudgill (1974) were two pioneer studies on the voice quality of dialects of England. The thesis of Knowles was a general description of the Scouse accent. He made interesting remarks on the voice quality of this accent. According to him, the voice quality of Scouse could probably be described as a lingual root (or body) articulation, a close jaw position and a certain degree of raised larynx voice accompanied with tense voice. This is why it is interesting to observe the voice quality of the Liverpool accent nowadays with the help of the Voice Profile Analysis protocol ([6]) and compare the results with the findings of Knowles (1978).

Trudgill (1974) also provided some interesting remarks on the voice quality of the Norwich accent. He compared the speech of working class to the one of middle class. He made a description of the typical voice quality of the working class. He also underlined the necessity of integrating voice quality in the description of a dialect as it seems to be a significant component.

The problem of these two studies is that they were based on impressions and not on quantifiable data. Moreover, the terms employed to describe voice quality did not refer to a rigorous and common classification of the settings. Yet these two studies were pioneer work in the domain of voice quality and variation in English.

3.2. Quantifiable Data on Voice Quality

The two following studies are different from the previous ones for several reasons. First, Esling (1978) and Stuart-Smith (1999) were focused on the study of voice quality, they were not a general description of the accent of Edinburgh and Glasgow. Another reason is that they both refer to the same descriptive model introduced by John Laver.

Esling (1978) focused on the social differentiation of voice quality in Edinburgh through two types of experiment : an auditory analysis and some objective analyses using a laryngoscope and a laryngograph. Differences between middle and working classes were observed, indeed some features of voice quality were predominant for each accent. For instance, working class Edinburgh accent was judged to be extremely harsh whereas standard Edinburgh English was described as having extreme creakiness and moderate nasality.

Stuart-Smith (1999) also described social differences of voice quality but in Glasgow. She used Laver's terminology which makes the comparison with the Edinburgh voice quality easier. This study is based on a perceptual analysis using the Vocal Profile Analysis protocol of Laver et al. (1981) in order to compare "Glasgow Standard English" (GSE) and "Glasgow Vernacular" (GV). Her results first showed that voice quality varied according to age, gender but also class. As far as class is concerned, working class's predominant voice quality features are the use of open jaw, raised and backed tongue body and whispery voice. Middle class speech is mostly characterized by the absence GV voice quality features. She also found typical features according to gender and age which will not be discussed here.

4. Using the Vocal Profile Analysis Scheme

The Vocal Profile Analysis protocol was originally developed by researchers to analyze pathological voices. The entire project is described in several articles ([8], [9] and [10]). This tool helps speech therapists and phoneticians determining and evaluating the vocal profile of a speaker. One has to bear in mind that in order to use the VPA scheme, a training session at the Queen Margaret University is usually required.

The use of the VPA scheme relies on two notions: the neutral setting and the relations between settings and segments. The neutral setting is precisely defined in Laver (1980). It is an important concept because all the other settings are considered as deviations from the neutral setting. This setting must not be compared to the "normal" voice as it is not usually observed in non-pathological voices. Analyzing the corpus it is also very important to keep in mind the relations between the settings and the segments. Indeed, the segments are more or less influenced by the different settings. For instance, the lip spreading will have a greater impact on rounded vowels than on /i/ sounds. This is why, in order to judge on a setting one must observe throughout the speech the segments which are easily influenced by the setting.

The first task when completing the analysis is to judge whether the setting is neutral or not. If the setting is not neutral, the second task consists in determining what kind of deviation it is and its extent with the help of the scalar degrees. If a compound phonation type such as whispery creaky voice has to be annoted, then the two boxes "whispery voice" and "creaky voice" will be ticked. Finally, in order to signal the intermittent presence of a phonation type, the scalar degree is signaled with an "i".

Although the VPA protocol was first created in order to evaluate pathological voices, this analysis can be used to determine the vocal profile of non-pathological subjects. Indeed, as mentioned in section 3, Stuart-Smith (1999) used the protocol in order to observe the variation of voice quality in Glasgow. This is why the VPA protocol was chosen, in this study, in order to observe the accent of Liverpool.

5. A pilot study of the accent of Liverpool

5.1. Corpus

The recordings used for this study are all extracted from The Intonational Variation in English corpus (IviE) created by Grabe, E., Low, L. and Nolan, F. in 1997. In order to neutralize any variation, the reading of a text has been chosen. Indeed, it seems that this style of speech has little influence on the voice quality of a speaker, although no study has tested this hypothesis. The accent of Liverpool was chosen mostly because its voice quality is quite well-known. Knowles (1978) refers to it as "Adenoïdal voice quality" Thus the corpus of this pilot study is made of four speakers (two males and two females) all of 16 years old. They were all born in Liverpool and come from the same neighborhood since they attend the same school. Thanks to the use of this corpus several variables such as age or social class background can be controlled. In this way, only the variation between accents can be observed.

5.2. Results

For this preliminary study, only one judge (the author) analyzed the corpus. Once this task is completed, it is possible to describe the voice quality of the four speakers.

The results can be treated in two different ways thanks to the two different stages of the analysis. The first type of results is in term of neutral versus non-neutral settings. Indeed the four speakers were all judged neutral for the pharyngeal settings. None of them appeared to constrict or expand their pharynx. On the contrary, the four speakers tend to use a non-neutral position of the tip and blade of the tongue. The same remark can be made for the velopharyngeal settings. Concerning the phonation types, all the speakers used whispery voice but sometimes with another phonation type added to those two.

The second type of results is more precise because the degrees of the non-neutral settings are being observed. The results confirm the hypothesis according to which the scalar degrees (SD) obtained by non-pathological voices are usually under four. Concerning the tip and blade of the tongue, three speakers out of four used a slightly advanced position (SD one or two). Nasality is also present in the voice quality of the four speakers, two of them obtained a SD1 and the remaining speakers scored respectively SD2 and SD3, meaning that all of them scored under four which is typical of a British accent. At the level of the vocal tract tension, a moderately tensed vocal tract was observed for three speakers out of four.

The results found for the phonation types were also all under the scalar degree four which means that the voice component was predominant in their voice quality. All the speakers used a small amount of the whispery component with a degree one. Another interesting result is that three speakers out of four used the creaky voice, but two of them (the female speakers) intermittently. One of the speakers seems to use creaky voice at the end of the breath group as it is often mentioned in the literature but the other one tend to use it at other places such as "a girl *called* Cinderella" or "Lily *and* Rosa". Of course this phenomenon is observed for one speaker out of four, so one explanation could be that it is an idiosyncratic variation, however it could be interesting to develop this research on other speakers.

Thus, the common features of the speakers can then enlighten some aspects of the voice quality of the Liverpool accent. To sum up, the voice quality of the Liverpool accent tends to show an advanced position of the tip/blade of the tongue according to three speakers out of four. Nasality is also present but moderately. On the laryngeal level, compound phonation types such as whispery voice and also whispery creaky voice are commonly used.

6. Conclusion

This pilot study succeeded in showing that the VPA protocol can be used in order to describe non-pathological voices. The scheme also permitted to establish the different settings that make the voice quality of those four speakers from Liverpool. Of course, in order to obtain reliable results, the corpus should contain more speakers of the same speech community. Moreover since only one judge analyzed the corpus, another way to increase the objectivity of this research would be that several judges analyze the same corpus. Then in order to confirm the results some statistical tests should be done along with an acoustic analysis using for example the Long Term Average Spectrum.

This is why this study is part of a larger project that aims at describing voice quality variations between several accents of the British Isles.

References

[1] Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh : Edinburgh University Press.

[2] Esling, J. 1978. *The identification of features of voice quality in social groups*. Journal of the International Phonetic Association, 7, 18-23.

[3] Kent, R.D. and Ball, M.J. (2000). *Voice Quality Measurement*. San Diego: Singular Publishing Group.

[4] Knowles, G. (1978). *Scouse, the Urban Dialect of Liverpool.* Unpublished PhD dissertation, University of Leeds, U.K.

[5] Laver, J. 1980. *The Phonetic Description of Voice Quality*. Cambridge University Press : Cambridge.

[6] Laver, J. 1981. The analysis of voice quality : from classical period to the twentieth century. Part 3.

[7] Laver, J. 1991. *The gift of speech : readings and analysis of speech and voice.* Edinburgh University Press, Edinburgh. (chapter 9, 14)

[8] Laver, J., Wirz, S., Mackenzie, J., and Hiller, S.M. 1985. Vocal profile analysis in the description of voice quality. In V. Lawrence (Ed.), *Transactions of the 14th Symposium on the Care of the Professional Voice*. 184-192. New York : The Voice Foundation.

[9] Laver, J., Wirz, S., Mackenzie, J., and Hiller, S.M. (1991). A perceptual protocol for the analysis of vocal profiles. In J. Laver (Ed.), *The Gift of Speech : Readings in the analysis of speech and voice*. 265-280. Edinburgh : Edinburgh University Press.

[10] Mackenzie Beck, J.M. 1988. *Organic variation and voice quality*. Unpublished thesis, University of Edinburgh.

[11] Stuart-Smith, J. 1999. Glasgow: Accent and voice quality. in P. Foulkes and G. Doherty (eds), *Urban voices: accent studies in the British Isles*, (Edward Arnold: London, 1999), 201-220.

[12] Trudgill, P. 1974. *The Sociolinguistic differentiation of English in Norwich*. In Coupland and Jarowski. Cambridge University Press, Cambridge.