Proceedings of the 1st Conference on Speech Sciences, Seoul, Korea 2002
pp. 128 — 131

Aspects of Voice Quality in Women

Evelyn Abberton

Department of Phonetics and Linguistics
University College London, England
e.abberton@ucl.ac.uk

Abstract

Voice quality is a multi-dimensional component of fluent speech determined both culturally and physiologically. This paper presents some results of a study of the voices of forty four 20-year-old female speakers of British English. Anechoic acoustic and electro-laryngograph recordings were made of a 2-minute read text, and analysed using Laryngograph® Speech Studio software to provide information on vocal fold contact ratios on the basis of ELG/EGG transconductance variation. These analyses can be related to the perceived breathiness of a voice. Fewer than half of the speakers (43%) had lower contact ratios at the upper end of their fundamental frequency ranges than at the lower end (correlating with a breathier voice at higher pitches); 25% of the speakers showed the opposite pattern, and for the remaining 32%, contact ratios varied symmetrically across the speaker's voice pitch range.

1. Introduction

Voice quality plays an essential role in speech communication. It can help or hinder intelligibility [1], and is a rich source of indexical information with linguistic, cultural and family determinants. An understanding of the physiological and acoustic correlates of perceived voice qualities supplements phonetic and impressionistic labels and descriptions, and can lead to improvements in synthetic and prosthetic speech systems; greater physical knowledge can also provide insights to surgeons and therapists in the Voice Clinic to help the diagnosis and management of dysphonic speakers.

It is hoped that the preliminary results of the ongoing work described in this paper will contribute to providing normative references, and increase our understanding of the voice quality characteristics of young women. Both physical and anecdotal evidence suggests that young female speakers have breathier voices than men, with a tendency for increased breathiness at higher pitches[2,3]. However, voice quality varies with social situation and speaking task, and the correlation of breathiness with fundamental frequency may be task-dependent. In particular, findings for sustained vowels may not hold when connected speech is examined. The present study analyses the variation of vocal fold contact ratios in a read text allowing for examination of relationships occurring in the natural pitch changes of intonation. Shorter contact ratios correlate with breathy voice quality.

2. Participants and Method

DAT recordings were made in an anechoic chamber, using both microphone and electro-laryngograph (EGG) inputs, of the speech of a group of 44 female Speech Science students, average age 20 years. A read text of 2 minutes duration was analysed using the Laryngograph Speech Studio system [4]. The participants were all native speakers of British English but there were no other exclusion criteria (such as smoking or medication) and the sample was thus considered to represent a group of normal young

Proceedings of the 1st Conference on Speech Sciences, Seoul, Korea 2002
pp. 128 — 131

women. However, it is intended to examine sub-groups based on such criteria in a further phase of the work.

2.1. Analyses

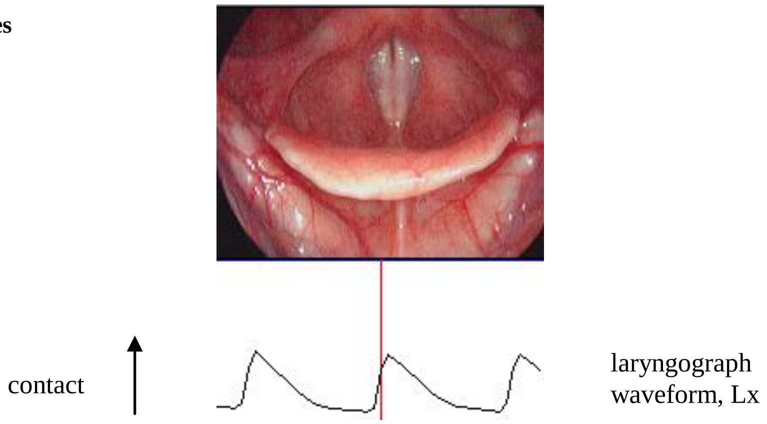

Figure 1 Stroboscopic view for near maximum contact – still showing a glottal chink

A perceptually breathy voice can be produced in more than one way. The vocal folds may not close completely during the vibratory cycle. Figure 1 gives a view from a stroboscopic examination of the vocal folds of a normal young woman phonating at 264Hz. The vertical marker on the accompanying laryngograph waveform (Lx) shows the moment in the vibratory cycle illustrated in the stroboscopic image. A posterior chink is evident throughout the cycle. Breathy voice can also be produced when there is a relatively long open phase in each cycle, allowing lung air to pass into the vocal tract with audible friction. Both mechanisms may be present at the same time.

Speech Studio provides a number of analyses, displays and statistical treatments, all of which can be explicitly related to aspects of hearing and auditory analysis [4]. Those relevant to the present paper relate to measures of Fx - fundamental frequency (the major physical correlate of perceived pitch) - in terms of range and central tendency (mean, mode, median); to irregularity of vocal fold vibration in connected speech (CFx); and to an indicator of the duration of the closed phase in each vocal fold vibratory cycle, Qx.

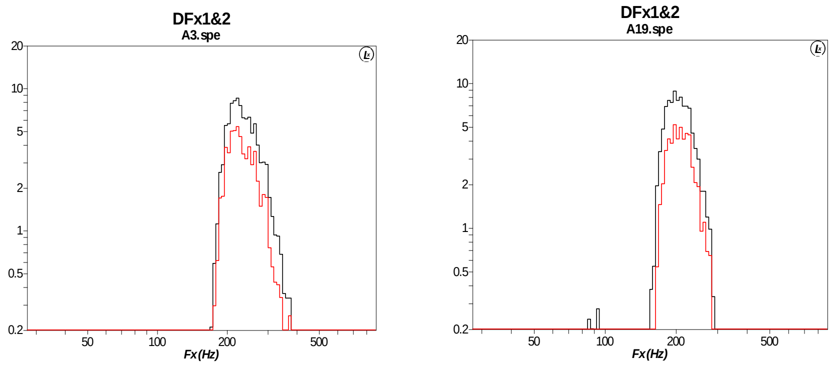

Figure 2 First and second order DFx plots for two female speakers A on left & B on right

Figure 2 shows DFx plots for two of the speakers in the study. In this figure fundamental frequency is plotted horizontally on a logarithmic scale, and the vertical axis is probability of occurrence of a particular frequency. For each speaker the slightly larger, first order distribution takes every fundamental frequency period of the 2-minute passage into account, and the smaller, second order distribution is based on pairs of adjacent periods falling into the same statistical bin. For these normal speakers the distributions are well defined with little difference between first and second order distributions.

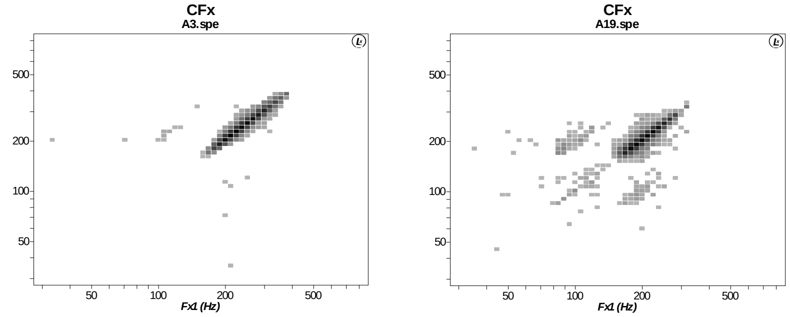

Figure 3 CFx – period by period crossplots for speaker A on the left and speaker B on the right

Figure 3 shows the crossplots, CFx, for the same speakers based on an analysis of all vocal fold periods in the spoken sample. Each axis represents log Fx. Although both speakers are normal, Speaker B's voice has more irregularity than Speaker A, shown by the outlying points in the lower part of the fundamental frequency range.

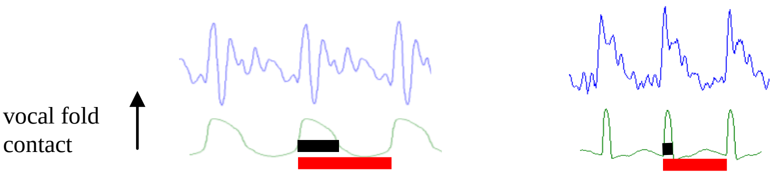

Figure 4 Illustration to show how Qx is calculated using the Lx waveform

Figure 4 shows the calculation of Qx, closed quotient ratio, for modal voice (with a relatively long closed phase) and for the type of breathy voice with a long open phase. The percentage of time the vocal folds are in contact during each cycle is calculated by measuring the width of the peaks in the Lx waveform at a point 70% down from maximum closure, using transglottal electrical conductance. Qx can be plotted against fundamental frequency using first and second order analyses.

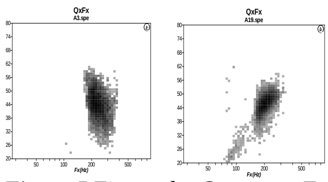

Figure 5 First order Qx versus Fx

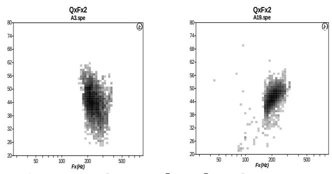

Figure 6 Second order Qx versus Fx

Figures 5 and 6 show this relationship for the two speakers A & B of Figures 2 and 3: Speaker A shows decreasing closed quotient with increasing fundamental frequency, whereas Speaker B shows the opposite trend.

The relationship between Fx and Qx was investigated for the 44 young female speakers described above to examine the hypothesis that their voices would be breathier at the higher end of their fundamental frequency ranges in connected speech: i.e. that there is an inverse relationship between fundamental frequency and closed quotient ratio.

3. Results

43% of the speakers showed the hypothesised inverse relationship between contact ratio and fundamental frequency, as represented by Speaker A: shorter contact ratios with increasing fundamental frequency. 25% showed the pattern of Speaker B: increasing contact ratios with increasing fundamental frequency, and 32% had contact ratios varying symmetrically across the speaker's voice pitch range. Figure 7 is a summary of Qx plots for each of the three categories.

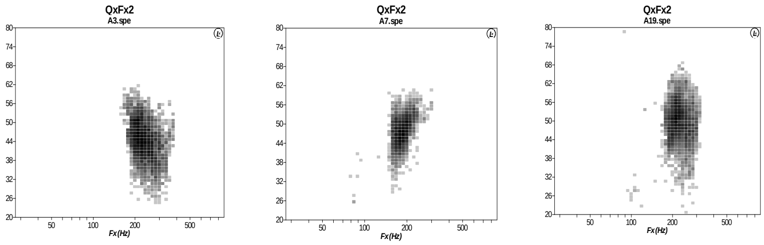

Figure 7 Summary of the three main voice quality types

4. Discussion and conclusions

Appropriately analysed fluent connected speech gives insights into structures of voice control. It is not obvious that a breathy voice quality is equally perceptually salient at high and low voice pitches in connected speech even if vocal fold contact ratios are smaller at the higher or lower ends of a speaker's pitch range. Loudness of the voice will also have an effect on perceived voice quality [3]. However, the results of this study clearly show that there are three main modes of contact ratio and fundamental frequency variation in this group of young female speakers, regardless of the possible existence of a permanent glottal chink (a factor that has not yet been studied for these speakers), and that young women do not necessarily have breathier voices at higher pitches. Physiological, cultural and linguistic factors may operate independently, and voice quality can be modified to meet social needs.

5. References

[1] Markham,D. and Hazan,V., "The Perception of Speaker Variability by Adults and Children." in press ICSLP 2002

[2] Linville,S.E. (2000) "The Aging Voice", in *Voice Quality Measurement*, R.D.Kent and M.J. Ball, Eds., San Diego: Singular-Thomson Learning

[3] Sodersten,M. and Lindestad,A. (1990) Glottal Closure and Perceived Breathiness During Phonation in Normally Speaking Subjects. *JSHR, 33(3): 601-611*.

[4] Fourcin,A. 2002, "Patterns, Structure and Measurement in Voice Pathology"–these Proceedings