The role of the formant intensity in the analysis of emotionally sung vowels

Siniša Djordjević, Gordon Troup, Charles Don

Department of Physics, Monash University, Clayton VIC 3168, Australia

Abstract: The formant intensity could reflect the emotional energy involved in singing. A spectrographic analysis of emotionally sung vowels [a] and [i] is used to investigate this hypothesis. The emotional effects on the formant intensity of the vowel [a] of a tenor, [i] of a bass and alto, and [i] of male and female singers are presented. The results show the emotional difference among individual singers as well as between groups of singers. The emotivity, defined as the ratio of the emotional and neutral formant intensities, of a bass is compared to that of an alto.

1. INTRODUCTION

It appears that different intentions concerning the emotional character of a piece of music will affect practically all variables of the performance (1). If emotions in singing are supposed to be manifested as a combination of different vocal efforts, there should be some differences among them, like among the soft, light, neutral, free, straight, vibrato and loud modes of singing (2). The spectral content of sung vowels is affected by different mood (3) and emotions (4). It was found that the fundamental frequency of the sung vowel [a], for instance, remains practically unchanged with an increase in total energy of the tone (5), the formant frequencies of a vowel are exclusively determined by the articulation (6), and the first three formants for well-known male singers show great similarity as far as frequency position is concerned (7). The sound intensity, and consequently formant intensity, are related to the sound pressure level (8) and can be used to distinguish emotions connected to it. The level of the singer's formant is found to increase with increasing vocal intensity, for the fundamental frequency constant (9).

How the formant intensity of other formants and the total intensity of a sung vowel represent emotional effects is the subject of this work. The formant intensity of a single vowel was studied first, then two voices, a bass and an alto, were compared, and finally, male and female singers were considered as different groups of voices. A quantitative measure, the 'emotivity' is introduced as $E = G_e/G_n$, where $G_e$-emotional formant intensity and $G_n$-neutral formant intensity, to represent personal or group emotional efforts.
3. EXPERIMENTAL RESULTS

The LPC (autocorrelation) method involving a KAY's Computer Speech Lab software was used for all measurements, with a frame size 10 ms, pre-emphasis 1 and sampling rate 16 kHz. For the emotionally neutral, sustained sung vowel [a], the sustainability is presented by measuring the intensity of the first three formants and their sum (Fig. 1a). The [a] was sung with fundamental frequency of 154 Hz and constant sound pressure level of 56 dB. The formant frequencies were F₁=660 Hz, F₂=1140 Hz, and F₃=2660 Hz. The formant intensities were measured at eleven successive points on the vowel waveform signal and the deviation determined to be: G₁=19.7±0.5 dB, G₂=21.6±0.4 dB, and G₃=10.2±0.6 dB. As all curves exhibit the same trend, the deviation appears to be due to the singer rather than uncertainty of measurement.

![Formant Frequency and Intensity Graphs](image)

The formant frequency and formant intensity values of the [a], sung with increasing intensity from 0 to 77 dB (SPL) during 1.35 seconds at the fundamental frequency of 154 Hz, were measured at three different times and are presented in Fig. 1b and 1c. Numerically, the formant frequencies were F₁=630 Hz and F₂=1160 Hz in all three moments, and F₃ varies between 2590 and 2650 Hz (Fig. 1b). The effect of sad, happy, and angry emotions, represented by the total formant intensities (G₅, G₆, and G₇ respectively), on the vowel's spectra are presented in Fig. 1d.

Comparative results for the bass and alto's formant frequency, formant intensity and total formant intensity for the emotional and neutral singing are given in Fig. 2a, 2b, and 2c.

![Comparative Graphs](image)
2. METHOD

The vowel [a] of a tenor was chosen to study the effect of emotions on the formant intensity of emotionally sung vowels. Firstly, an emotionally neutral [a] was extracted out of the sung sequence of notes (C4-E4-G4-C5-G4-E4-C4). If there are no other influences, the formant intensity, as well as the total intensity is expected to remain at a constant level. The size of the fluctuations in formant intensity indicates how well the voice is sustained. Secondly, a tenor sang the vowel [a] at the same pitch, changing only the voice intensity. The corresponding change of formant intensity is expected to follow the overall intensity trend. Thirdly, the [a] was taken out of the same song sung with the emotions sad, happy, and angry and the emotional effects in the spectra measured.

A comparison was made between a bass and an alto singer, both of whom sang the same vocal phrase ('O cessate di piangarmi' by Alessandro Scarlatti (1659-1725)) with the three emotions, sad, happy, and angry, as well as singing the neutral sequence. The formant frequencies and formant intensities of the emotional and neutral vowel [i] for the two voices were compared, as well as total formant intensity. In this study, the total formant intensity G is calculated as the sum of formant intensities of the first three formants, G1, G2, and G3.

In order to investigate the effect of emotions from a group of singers, the same vowel [i] was extracted from the singing of seven male and six female choir singers, and the formant intensities analyzed statistically.

As the quantitative measure of the emotional effects, the emotivities of the bass and alto were compared.
3. EXPERIMENTAL RESULTS

The LPC (autocorrelation) method involving a KAY’s Computer Speech Lab software was used for all measurements, with a frame size 10 ms, pre-emphasis 1 and sampling rate 16 kHz. For the emotionally neutral, sustained sung vowel [a], the sustainability is presented by measuring the intensity of the first three formants and their sum (Fig. 1a). The [a] was sung with fundamental frequency of 154 Hz and constant sound pressure level of 56 dB. The formant frequencies were $F_1=660$ Hz, $F_2=1140$ Hz, and $F_3=2660$ Hz. The formant intensities were measured at eleven successive points on the vowel waveform signal and the deviation determined to be: $G_1=19.7\pm0.5$ dB, $G_2=21.6\pm0.4$ dB, and $G_3=10.2\pm0.6$ dB. As all curves exhibit the same trend, the deviation appears to be due to the singer rather than uncertainty of measurement.

![Graphs showing formant frequency and intensity](image)

Fig. 2 - (a) The formant frequency, where the 'F' is for alto and 'Fb' for bass, (b) formant intensity ('G'-alto, 'Gb'-bass), (c) total formant intensity of a bass and alto (dotted line), formant intensity of (d) male and (e) female group, and (f) emotivity of a bass and alto for the vowel [i]

The formant frequency and formant intensity values of the [a], sung with increasing intensity from 0 to 77 dB (SPL) during 1.35 seconds at the fundamental frequency of 154 Hz, were measured at three different times and are presented in Fig. 1b and 1c. Numerically, the formant frequencies were $F_1=630$ Hz and $F_2=1160$ Hz in all three moments, and $F_3$ varies between 2590 and 2650 Hz (Fig. 1b). The effect of sad, happy, and angry emotions, represented by the total formant intensities ($G_5$, $G_{10}$, and $G_{15}$ respectively), on the vowel's spectra are presented in Fig. 1d.

Comparative results for the bass and alto's formant frequency, formant intensity and total formant intensity for the emotional and neutral singing are given in Fig. 2a, 2b, and 2c.
These figures show only a small change in the formant frequencies as the emotions changed with a marginally greater effect in the alto's F3 (Fig. 2a). There is a much greater variation in the formant intensities (Fig. 2b) the most marked being in the alto's G3. Overall, there is a higher total intensity of alto's than bass's [i] formants (Fig. 2c), which was already noticed (9).

Statistical results of [i]-male and [i]-female formant intensity are given in Fig. 2d and 2e, and emotivity of the bass and alto in Fig. 2f.

4. DISCUSSION

The data of Fig. 1a and 1b indicate that a singer can maintain the intensity of the first three formants of a sung vowel to within a dB and that the frequency of three formants is independent of the level of the emitted sound. However, the intensity of the various formants increases (Fig. 1c) as the singer's output level increases and the total formant intensity follows the same overall trend. When the total formant intensity is used to represent the emotional effects on the same vowel, spectral differences are apparent between sad, happy and angry emotions (Fig. 1d). These emotions have similar formant frequencies but different formant intensity levels. Some formants exhibit a frequency shift and additional formant appear for more intense emotions (eg. F3-happy is displaced higher than the F3-sad and becomes two formants in F3- and F4-angry), indicating that the vocal tract also slightly changes its shape in the singer's effort to express a stronger emotion.

When comparing the formant frequencies present when singing with different emotions (Fig. 2a), there is again little change with the greatest effect being at the higher frequencies, especially the alto's third formant. However, the intensities do show significant changes with emotion, sad being lower in all formants for the bass, although this trend was not so apparent in the alto voice. The group of male singers also indicate a lower intensity for a sad sung [i], however this was not apparent in the female group where there is much less sign of emotional effects.

Using the above results and the hypothesis that the formant intensity does represent emotional energy involved in sung vowels, the 'emotivity' can be used as a quantitative measure of the overall personal or group emotional efforts in singing (Fig. 2f) and, more generally, in speech.

BIBLIOGRAPHY

Czech Acoustical Society

Proceedings of the

32nd Czech Conference on Acoustics

SPEECH - MUSIC - HEARING

Prague, September 23 - 26, 1995
Czech Acoustical Society

in collaboration with

Faculty of Music of the Academy of Performing Arts
Czech Society for Phoniatrics and Pedaudiology
Postgraduate Medical School

under the auspices of
International Association of Logopedics and Phoniatrics

Proceedings of the

32nd Czech Conference on Acoustics

SPEECH - MUSIC - HEARING

Organizing Committee:

Alois Melka - President
Jan Štěpánek - Secretary
Petr Bažant
Mojmír Lejska
Josef Nosek
Alexej Novák
Zdena Palková
Jaroslava Přihodová
Josef Syka
Václav Syrový
František Šram

Prague, September 23 - 26, 1995
Proceedings of the 32nd Czech Conference on Acoustics
SPEECH - MUSIC - HEARING
held in Prague, September 23-26, 1995

Edited in VUZORT by: Alois Melka, Zdeněk Otčenášek, and Jan Štěpánek.
The contents and translation of the contributions is fully under the responsibility of their authors.

Printed in Prague, Czech Republic.
Copies of the Proceedings can be ordered from: VUZORT a.s.
Plzeňská 66
151 24 Prague 5
Czech Republic