Listening evaluation and classification of female singing voice categories

Marek Frič and Angelika Pavlechová
Musical Acoustics Research Centre, Academy of Performing Arts in Prague, Prague, Czech Republic

ABSTRACT
Objective: Classification of voices into types depends on several factors: physiological — the size of the laryngeal and vocal tracts; acoustic — musically acceptable vocal range; position of formants; and properties of timbre. The aim of the study is to verify whether a group of experienced voice pedagogues and singers can determine the vocal type of the artist based on listening to a part of the aria better than a group of musicians can, and to determine what acoustic properties of the recordings are linked with the perceptual results of their evaluation.

Methods: Freely available recordings of 11 females of different vocal categories of Rossini’s aria “Una voce poco fa” from the opera Il Barbiere di Seviglia were selected for listening tests performed on examples of recitatives and coloraturas. Seven voice teachers (experienced group) and seven musicians (laypeople group) evaluated the properties of the vocal category, timbre (dark-bright), resonance, vowel placement, suitability of vibrato, aesthetic impression, and voice flexibility.

Results: The results showed a significantly higher inter-judge reliability in the experienced group. The highest reliability was achieved in timbre and vocal category evaluation, the least consistent was the evaluation of resonance. Factor analysis of the assessment variability showed dependent ratings of the vocal category, brightness and vowel placement for both groups in recitative. The experienced group similarly evaluated the brightness and the vocal category in coloratura. Assessment of the vocal category correlated with the reported categories of singers only in the experienced group. The categories mezzo-soprano and soprano were differentiated by spectral levels (based on FFT analysis of whole stimuli) in the 3.5–4.1 kHz spectral band in the recitative and in the 1.3–2.1 kHz and around 2.5 kHz bands in coloratura, and by the position of the local minimum after the fifth maximum for both kinds of stimuli.

Conclusions: By means of correlation with ratings by experienced listeners, it is demonstrated that the voice category is mainly connected with the levels of specific spectral peaks, while brightness is correlated with the frequency positions of spectral maxima.

Introduction

The traditional classification of classical opera singing voice types divides male voices into 3 basic types (tenor, baritone, bass) (some literature also mentions the fourth type: contratenor); while female voices are divided into 3 types (soprano, mezzo-soprano, contralto). Among professional singers, the so-called “Fach System” is also applied in Europe, the original function of which was to divide singers into groups to describe their repertoire abilities.

In practice, classifying voices is a highly significant and crucial point in the vocal education process, especially of classical singers, as incorrect classification can cause voice problems later in their professional careers.

The division into voice categories is based on three fundamental criteria: range requirements and capabilities, voice timbre, and agility of the individual [1]. Other criteria for determining the voice category are features such as the position of transition of the registers, and the average fundamental frequency (f0) of the spoken voice [2].

Size (body and larynx), vocal length, and length of the vocal tract objectively describe the criteria for dividing the voice [3]. These properties have a significant influence on both the average f0 and the pitch range of the voice, the frequency positions and levels of formants, and thus the voice’s timbre.

The perceptual criteria for distinguishing the voice types based on vocal timbre are derived from two basic properties: the most commonly described is the bright-dark dimension [1], a property which is also called weight; the other feature discussed is the voice’s flexibility (the ability to sing coloratura). Objective studies of voice classification are based on the physiological properties of subjects and on the acoustic properties of voices, or they are related to perceptual evaluation of individual categories.

In the literature, it is possible to find differentiation of voices based on the prescribed ranges, in which the voice should be produced with a suitable voice quality (tessitura). Reported ranges, however, vary considerably among the authors [4]. The voice range itself cannot be considered as essential for distinguishing the voice category [2].

It is also possible to assume differences in the range whether it is a solo singer or choral singer. The agreement
band, which indicated the general frequency shift of the high female voices formants in this area.

A global relation among the ratings of brightness, voice category and vowel placement was shown, which could have been assumed based on the spectral connection with the positions of specific formants. To distinguish these specific effects, this study proposed observation, connected to evaluation by an expert group. The evaluation of the voice category was primarily connected with the energy in the band of 2–4 kHz and its spectral distribution, in the recitative represented by a negative correlation with the level of the third and fourth spectral maxima, and in coloratura by the negative correlation with SPR. In both kinds of stimuli, the position of the spectral minimum in the band above 3.5 kHz played an important role. In the recitative, a higher-pitched vocal category and brightness were connected with lower energy in the first spectral maximum. Brightness was more connected with the frequency position of spectral maxima in the recitative with the fourth and the fifth, and in coloratura more with the first and the position of the spectral maximum in the band below 2 kHz.

Correlations of vowel placement with spectral parameters were more similar to the bright-dark rating results than for the voice category. Only in coloratura was the vowel placement connected with the position of the first spectral maximum or with the position of the spectral maximum below 2 kHz.

The average pitch played a role only in evaluation of coloratura, with the experienced group linking the ability of longer-duration singing at a higher pitch with brightness, front vowel placement, and a more ringing voice. Average SPL was highly connected with resonance/ring and negatively with a higher-pitched category. Interestingly, resonance correlated negatively with the position of the spectral maximum in the band below 2 kHz and the position of the minimum after the second maximum in the recitative, while in coloratura the resonance/ring was positively correlated with the position of the second spectral maximum. This discrepancy could be explained by the formant-tuning typical for higher pitches [43], which was expected mainly in coloratura.

Conclusion

The present study confirms the assumption of a generally more consistent assessment of the voice properties by a experienced group than by an layperson one. The study confirmed the better ability to more accurately define a voice category in the experienced group’s listening evaluation. The bright-dark timbre rating was generally highly consistent, while resonance/ring was the least consistently rated. Factor analysis has shown that voice category rating is generally related to the brightness of the voice and in the recitative also with the vowel placement.

The listed categories of soprano and mezzo-soprano voices were basically different regarding the energy distribution and spectral level in the 2–5 kHz band. In the recitative, the differences were caused mainly by the position of the fifth spectral maximum, by the minimum after this maximum, and by the end of spectra dropping below 40 dB, while in coloratura it was conditioned by the whole spectral level in the 2–5 kHz range, by singing power ratio, and by the position of the local minimum after the fifth maximum in the FFT spectrum.

Based on the perceptual evaluation of the experienced group, a higher-pitched vocal category was mainly connected with the levels of specific spectral peaks, negatively with the first, third and the fourth spectral maxima in the recitative, and negatively as well with the second maximum and with the specific maximum around 2 kHz in coloratura. A negative correlation with the singing power ratio reveals the connection of relative levels in the 2–4 kHz band with voice category in coloratura.

On the other hand, brightness was correlated mainly with the frequency positions of spectral maxima, in the recitative with the fourth and the fifth, while in coloratura mainly with the first. Vowel placement revealed very similar correlations with spectral parameters, as did the bright-dark scale. The resonance/ring evaluation was connected mainly with the frequency position of the maximum below 2 kHz, especially with the second maximum in coloratura, which could be explained by the formant tuning.

Acknowledgements

This publication was written at the Academy of Performing Arts in Prague as part of the project “Sound Quality” with the support of the Institutional Endowment for the Long Term Conceptual Development of Research Institutes, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2018.

Disclosure statement

The authors report no declarations of interest.

ORCID

Marek Frič http://orcid.org/0000-0003-1583-4721

References


