Acoustic Analysis of Voice Quality in Iron Maiden’s Songs

Alexsandro R. Meireles[1] and Beatriz R. de Medeiros[2]

[1] Phonetics Laboratory, Federal University of Espírito Santo, Brazil

[2] University of São Paulo, Brazil

Introduction

- Aiming to investigate the interplay between men’s high-pitched voices (above C5 = 523.25 Hz) and complex modes of phonation, Meireles (2016) started to develop a research program with focus on voice production, in order to analyze the voice quality in high-pitched heavy metal singing. The fundamental frequencies of male singers in that study ranged from 366 to 666 Hz.
- Meireles (2016), on the other hand, described the complex interactions of voice quality settings in singing with perceptual and acoustic analysis, so as to contribute to a scientific investigation of voice in heavy metal. Also, his study contributed for the correlation between acoustic and perceptual data on singing, since there are very few studies on the field, and added the heavy metal style to the possibilities of research.
- The aim of this paper is to analyze Bruce Dickinson’s voice quality (*Iron maiden* singer) in 3 different *Iron Maiden*’s songs, in order to observe whether Dickinson uses a consistent voice quality throughout the songs. Therefore, this study aims at continuing the exploration of voice quality in high-pitched singing, so as to contribute to a scientific investigation of voice in heavy metal. We also hope that this study stimulates other researchers to work with this exciting field of research.

Methods

Singing Data:

• Excerpts from Iron Maiden’s songs (pre-scream and scream): Flight of Icarus (FOI), Run to the Hills (RTH), and The Number of the Beast (NOB).

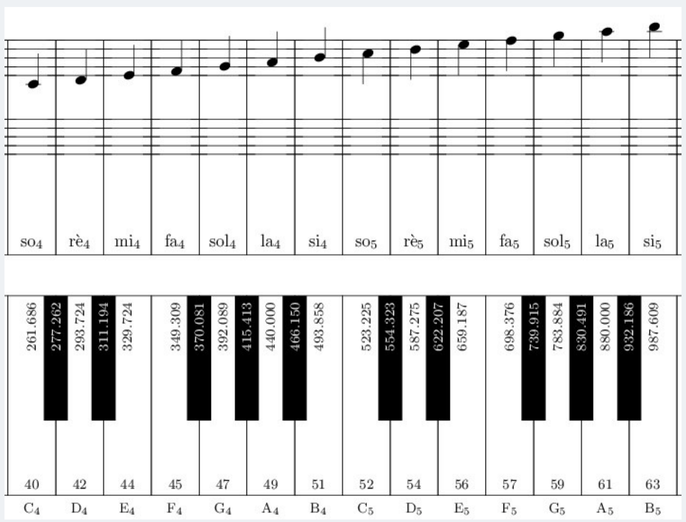
•The pre-scream part is closer to modal voice, but the scream part adds lots of air escape and vocal fold tension in order to produce the higher f0s.

•The total note range in the scores goes from F#4 (370 Hz) to A5 (880 Hz).

•The audio with the vocal tracks extracted from the original songs was freely obtained at www.youtube.com.

Acoustic Analysis: Voice Sauce (Shue, 2010)

Parameters: H1* (1st harmonic amplitude), H1*-H2*, H1*-A3* (3rd formant magnitude), CPP (cepstral peak prominence), Energy, HNR05 (0-500Hz), HNR15, HNR25, HNR35, F1, F2, B1, B2.



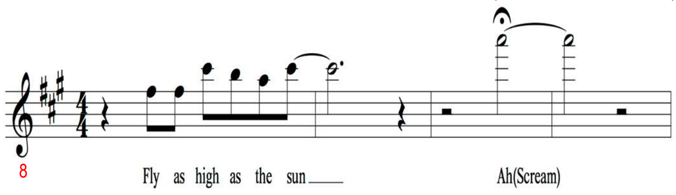


Figure 1: Excerpt of Flight of Icarus (3:34-3:47).

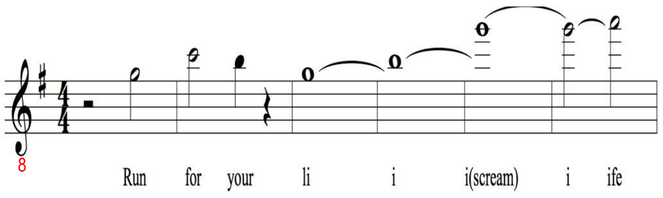
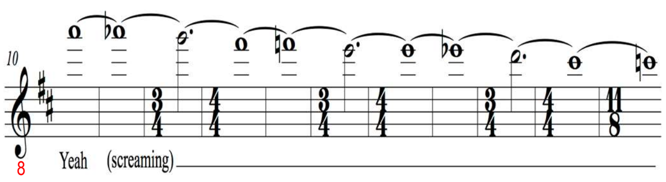


Figure 2: Excerpt of Run to the Hills (pre-scream and scream) (3:35-3:49).

Figure 3: Excerpt of The Number of the Beast (pre-scream) (1:04-1:16).



Parameter	Definition	Voice Quality
H1*	Relative amplitude of the 1st harmonic (* = corrected for formants)	Higher values -> Breathy Voice (Hillenbrand et al., 1994)
H1*H2*	Difference in amplitude between the 1st and 2nd harmonics	Lower values -> creaky and tense phonations (Keating et al., 2010)
H1*A3*	Difference between the first harmonic amplitude and the amplitude of the peak harmonic in F3 (spectral tilt)	Associated with tenser vocal folds (Shue, 2010)
CPP	Cepstral peak prominence	Higher values for modal phonation (Hillenbrand et al., 1994)
Energy	Voice intensity	Higher values for greater vocal effort (Shue, 2010)
HNR5	Harmonic-to-noise ratio (0-500 Hz)	Greater values for modal voice (Yumoto et al., 1982)
HNR15	Harmonic-to-noise ratio (0-1500 Hz)	Greater values for modal voice (Yumoto et al., 1982)
HNR25	Harmonic-to-noise ratio (0-2500 Hz)	Greater values for modal voice (Yumoto et al., 1982)
HNR35	Harmonic-to-noise ratio (0-3500 Hz)	Greater values for modal voice (Yumoto et al., 1982)
F1	Peak frequency of the 1st formant	Correlated with vowel height
F2	Peak frequency of the 2nd formant	Correlated with vowel frontness
B2	Bandwidth of the 1st formant	Expected lower bandwidth for modal voice
B2	Bandwidth of the 2nd formant	Expected lower bandwidth for modal voice

Acoustic Results

• A t-test with the 13 voice quality measures as a function of voice type (pre-scream and scream) revealed that these categories of voice qualities were statistically different from each other for 12 parameters.

• In the pre-scream part, 11 parameters were statistically different for all songs, except for H1* ((FOI = NOB) ≠ RTH) and HNR25 ((FOI = RTH) ≠ NOB).

• In the scream part, the songs behaved differently for 11 parameters, except H1*H2* ((FOI = RTH) ≠ NOB) and F1 ((FOI = NOB) ≠ RTH).

• The **H1* hypothesis** that higher values (in modulus) would be found for the “scream” part of the songs was fully corroborated for all songs .

• The **H1*H2* hypothesis** that a negative increase of this value would be found for the “scream” part of the songs was corroborated for 2 of the songs (FOI and RTH).

• The **H1*A3* hypothesis** that a negative increase of this measure would be found from the “pre-scream” part to the “scream” part was fully corroborated for all songs .

• The **CPP hypothesis** that we would find higher values for the “pre-scream” part in comparison to the “scream” part was corroborated for 2 of the songs (FOI and RTH).

• The **Energy hypothesis** that we would find a greater value of this parameter for the “scream” part was fully corroborated for all songs.

• The **HNR hypothesis** that we would find a greater value for the “pre-scream” part in comparison to the “scream” part was partially corroborated (8 out of 12 possibilities).

• The **F1 hypothesis** that higher F1s would be found in the “scream” part was fully corroborated for all songs.

• The **F2 hypothesis** that higher F2s would be found in the “scream” part was fully corroborated for 2 songs (RTH and NOB).

• The **B1 hypothesis** that higher B1s would be found in the “scream” part was not supported by the data.

• The **B2 hypothesis** that higher F1s would be found in the “scream” part was corroborated only for RTH

*HNR mean difference (pre-scream - scream). S stands for Song; N, HNR; n.s., non-significant; ***, highly significant.*

S	N5	N15	N25	N35
FOI	-2***	-1***	0 (n.s.)	1***
RTH	4***	6***	8***	9***
NOB	-10***	-3***	1***	4***

Conclusion

- This study is a further development of the new methodology presented in our previous study (Meireles, 2016). Our promising results have shown that this methodology was robust enough to analyze this kind of vocal performance.
- In the future developments of our method, we will complement the acoustic data with articulatory analysis such EGG, ultrasound and MRI that may refine the understanding of the strategies used by professionals to sing extreme high notes in heavy metal or other singing style.
- Moreover, we will continue to investigate the relationship between perceptual and production data, by development a new method for singing analysis.

Acknowledgements

The authors would like to thank the *Espírito Santo Research Foundation* (FAPES grant 80706630 to the first author) and the *São Paulo Research Foundation* (FAPESP grant 2015/06283-0 to the second author) for supporting this research, and Pablo Arantes and Plínio Barbosa for helping with Praat analysis.

References

A. R. Meireles. “Perceptual and Acoustic Study of Voice Quality in High-Pitched Heavy Metal Singing,” in Proceedings of the 8th International Conference on Speech Prosody, Boston, USA, pp. 1245-1249, 2016.

J. Hillenbrand, R.A. Cleveland, and R.L. Erickson. “Acoustic correlates of breathy vocal quality.” J. Speech and Hearing Research, 37:769–778, 1994.

P. Keating, C. M. Esposito, M. Garellek, S. u. D. Khan, J. Kuang. “Phonation Contrast across languages”. *UCLA Working Papers in Phonetics*, No. 108, pp. 188-202, 2010.

Y.-L. Shue. *The Voice Source in Speech Production: Data, Analysis and Models,* PhD Thesis, UCLA, 2010.

E. Yumoto, W. Gould and T. Baer, T. “Harmonics-to-noise ratio as an index of the degree of hoarseness”. J. Acoust. Soc. Am. 71, 1544–1550, 1982.