Anais do VI Colóquio Brasileiro de Prosódia da Fala, v.4, p. 79 – 82 / 2017
http://www.periodicos.letras.ufmg.br/index.php/anais_coloquio

Acoustic Analysis of Voice Quality in Iron Maiden's Songs

Alexsandro Meireles

Universidade Federal do Espírito Santo, Brasil

meirelesalex@gmail.com

Abstract

This paper studies the voice quality in high-pitched registers of Iron Maiden's songs. The f0 range varied from 298 to 998 Hz. Three Iron Maiden's songs were selected for analysis: *Flight of Icarus*, *Run to the Hills*, and *The Number of the Beast*. Two very high-register excerpts were selected from these songs, so as to verify Bruce Dickinson's vocal strategies to sing. The acoustic analyses were run with the software VoiceSauce [13] that automatically extracted thirteen parameters of long-term measures (H1*, H1*H2*, H1*A3*, CPP, Energy, HNR5, HNR15, HNR25, HNR35, F1, F2, B1, B2), and Praat [24]. Results indicate that twelve voice quality parameters were capable of discriminating two broad categories of voice quality: "pre-scream" and "scream". The only parameter that was not consistent for this discrimination was CPP.

Index Terms: singing; voice quality; heavy metal; Iron Maiden; acoustic analysis

1. Introduction

Classic heavy metal vocals (eg. Judas Priest, Aerosmith, Iron Maiden) are known to have a wide vocal range and a complex timbre variation. It is characterized by a combination of vocal settings such as pharyngeal constriction, raised larynx, tense vocal tract and larynx, and complex modes of phonation (falsetto, creaky voice, harsh voice, whispery voice). The scientific study of the voice quality in singing is important, since adoption of these settings requires special attention in clinical voice analysis as it may result in future speech pathologies.

Aiming to investigate the interplay between men's high-pitched voices (above C5 = 523.25 Hz) and complex modes of phonation, Meireles; Cavalcante [1] 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.

Voice production is understood here as stated by Laver [2], who defines voice quality as all the vocal characteristics that are related to speech, including laryngeal, supralaryngeal, and muscular tension and voice dynamical features. Pecoraro et al. [3], for example, has shown that Metal singers use many types of voice adjustments technically known as vocal drives, which can be physiologically produced with different vocal tract configurations.

Even though rare, some studies have focused on rock singing, such as the study by Oliveira and Behlau [4], Thalen and Sundberg [5], and Gonsalves and colleagues [6]. However, the style of singing used there is not vocally related to the vocal style of this paper. Despite drawing inspiration from 1960s' rock [7], as declared by members of bands that gave rise to heavy metal such as Black Sabbath and Motörhead, heavy metal is a much more aggressive variation of classical rock and was explored to an even lesser degree in academia.

Meireles and Cavalcante [1], 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. As this paper is a continuation of Meireles and Cavalcante's study [1], we will present in the next section a short description of its main results.

2. The heavy metal voice analysis of our previous study

Four singers (2 professionals and 2 amateurs) sang two excerpts of *Iron Maiden*'s *Aces High* that presents very high notes (starting at 366 Hz). Perceptual analyses were run using the VPAS protocol [8, 9, 10, 11]. Acoustic analysis was run based on the VoiceSauce software [12, 13] that automatically extracted thirteen parameters of long-term measures (H1*[1], H1*H2*, H1*A3*, CPP, Energy, HNR5, HNR15, HNR25, HNR35, F1, F2, B1, B2 [12, 13].

Regarding the perceptual analysis, the amateur singers (I, A2) varied less their voice qualities while singing at extremely high notes (above tenor C note, C5 = 523.25 Hz). On the other hand, the professional singers (A1, J), while also using muscular tension settings, demonstrated higher levels of these settings. Furthermore, we detected the following settings: open jaw (A1: grade 2[2]; J, grade 4), raised larynx (A1, grade 4; J, grade 5), minimized tongue body range (J, grade 2), and backed tongue body (J, grade 2). As for phonation settings, modal voice with creak (creaky voice) predominated in both singers. This type of setting corresponds to what is expected in

[1] The asterisk (*) means that spectral magnitudes H1, H2 and A3 have been corrected for the effects of formants [12, 13].
[2] The scale varies from 1 to 6.

heavy metal style.

Regarding the acoustic analysis, two excerpts were analyzed. H1*H2* is higher for the professional speakers in the first excerpt, suggesting a more breathy voice for the professionals ([14] have shown that higher levels of H1*H2* are usually correlated with breathy voice). Nevertheless a stronger Energy level is added to the quality of this professional breathy voice. Also, we found, in the first excerpt, higher values of HNR for the professional speakers, suggesting a more modal phonation than the amateur speakers ([15] suggest higher values for modal phonation). Nevertheless, in the second excerpt, an inverted pattern occurred, the professional singers showed lower HNR levels suggesting higher degree of breathiness than the other singer group. In addition, as in the first excerpt, the professional singers showed in the second excerpt higher Energy levels. Yet, the spectral slope (H1*A3*) and H1* were smaller for the professional speakers, what suggests an addition of air escape in the vocal folds but with the addition of high acoustic energy.

Summing up, the professional singers were most significantly distinct in their continuous maintenance of high tension in the vocal tract and the vocal folds, which was found intermittently and to a lesser degree among the amateur singers. Additionally, the experienced singers held an open jaw position and raised larynx vocal posture to reach the high notes. These settings were not observed in the amateur singers.

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.

3. Methodology

Three Iron Maiden's songs with extremely high f0 were selected for analysis: *Flight of Icarus* (Bruce Dickinson and Adrian Smith, henceforth FOI), *Run to the Hills* (Steve Harris, henceforth RTH), and *The Number of the Beast* (Steve Harris, henceforth NOB). The excerpts of the songs are shown in figures 1, 2, 3, and 4. As can be seen in the pictures, 2 parts called "pre-scream" and "scream" were chosen for each song. For figures 1 and 2, "Scream" is indicated in the score, what comes before that is the "pre-scream" part. By listening to the songs, we can perceive that these two parts have very different voice qualities. 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).

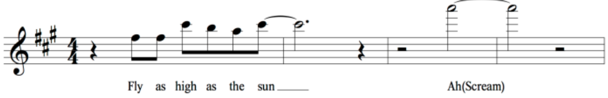

Figure 1: Excerpt of *Flight of Icarus* (3:34-3:47).

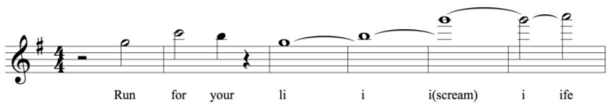

Figure 2: Excerpt of *Run to the Hills* (pre-scream and scream) (3:35-3:49).

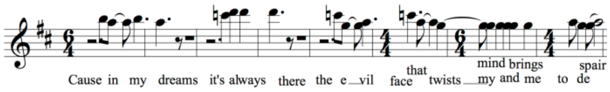

Figure 3: Excerpt of *The Number of the Beast* (pre-scream) (1:04-1:16).

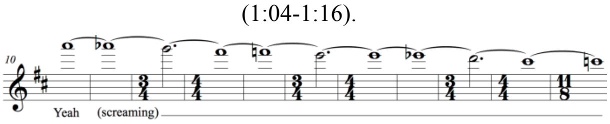

Figure 4: Excerpt of *The Number of the Beast* (scream) (1:18-1:30).

The audio with the vocal tracks extracted from the original songs was freely obtained at www.youtube.com, and downloaded to the computer desktop using the YouTube Audio and Video Downloader plug-in from Firefox 51.0.1. All files were downloaded as MPEG-4 Audio (stereo) with a sampling rate of 44.1 KHz, converted to WAV (mono) in Praat [24], and then annotated as "pre-scream" and "scream" in Praat. For the acoustic analysis we used the software VoiceSauce [13] that automatically extracted thirteen parameters of long-term measures (H1*, H1*H2*, H1*A3*, CPP, Energy, HNR5, HNR15, HNR25, HNR35, F1, F2, B1, B2) [12, 13]. For f0 extraction we used the function *To Pitch (ac)* in Praat [24], so as to easily manipulate the pitch floor and pitch ceiling.

H1* is the relative amplitude of the first harmonic corrected for the effects of formants. Higher values are usually associated in the literature with breathy voice [16, 17], so we hypothesize that higher values (in modulus) will be found for the "scream" part of the songs.

H1*H2* is the difference in amplitude between the first and the second harmonic. According to Keating and colleagues [14], higher values are associated with breathy and lax phonations [see also 16, 17, 19, 20, 21], and lower values with creaky and tense phonations. As for H1*, we hypothesize a negative increase of this value for the "scream" part of the songs. We have to be cautious though since some interaction between tense vocal folds and breathiness is usually found in singing, what can difficult the interpretation of this measure.

H1*A3* is the difference between the first harmonic amplitude and the amplitude of the peak harmonic in the F3 region. This is one of the ways of measuring spectral tilt. According to Gordon and Ladefoged [18], spectral tilt is "the degree to which intensity drops off as frequency increases" (p. 15) and is "is characteristically most steeply positive for creaky vowels and most steeply negative for breathy vowels" (p. 15). Many studies have associated this measure as a correlate for stress [22, 23], and according to Shue [12, p. 19], "words with more stress or emphasis will lead to tenser vocal folds which contain more high spectral frequency components during phonation". We hypothesize, therefore, a negative increase of this measure from the "pre-scream" part to the "scream" part.

CPP is the cepstral peak prominence. According to Hillenbrand and colleagues [17, p. 772], "the idea behind the CPP measure is that a highly periodic signal should show a well defined harmonic structure and, consequently, a more prominent cepstral peak than a less periodic signal". Our hypothesis is then that CPP should be higher for the "pre-scream" part in comparison to the other part.

Energy is a measure of voice intensity, which, according to Shue [12, p. 61-62], may be correlated with vocal effort. So, as greater vocal effort is expected for the "scream" part, we hypothesize a greater value of this parameter for this part.

HNR5, HNR15, HNR25, and HNR35 are the harmonics-to-noise ratios taken at the frequency ranges from 0-0.5 kHz (HNR5), 0-1.5 kHz (HNR15), 0-2.5 kHz (HNR25), and 0-3.5 kHz (HNR35). These measurements are made in VoiceSauce [12, p. 66] by "liftering the pitch component of the cepstrum and comparing the energy of the harmonics with the noise floor". Yumoto and Gould [15] found that the HNR for a healthy group varied between 7.0 and 17.0 dB with a mean of 11.9 dB. Based on previous studies, we hypothesize a greater HNR for the "pre-scream" part in comparison to the "scream" part.

F1 is the peak frequency of the first formant. It is correlated with vowel height. The higher the F1, the lower is the tongue position. As a common strategy of singers for reaching vowels with high fundamental frequencies is to widen the vocal tract [see, for example, 1], and, because the highest notes are on the "scream" part, we hypothesize higher F1s in the "scream" part.

F2 is the peak frequency of the second formant. It is correlated with vowel frontness. The higher the F2, the more anterior is the vowel. As widening the vocal tract helps the singers to reach extremely high notes, we hypothesize higher F2s in the "scream" part.

B1 and B2 correspond, respectively, to the bandwidth of the first and the second formants. Due to the increase of vocal fold tension, and a possible greater air escape, what may disturb the bandwidth measurements, we hypothesize greater bandwidths for the "scream" part for both parameters.

It is important to highlight that we intended to compare the voice productions using the same song used in Meireles and Cavalcante [1], but no good quality audio was available for the vocals-only of Aces High on the internet, nor could we extracted the vocals from the song file. That's why we chose songs with extremely high f0s that could be comparable to the voice quality presented in this song. Also, we didn't explore in our previous study the extreme vocal fold tensions (called "scream" here).

4. Results

All statistical analyses were run in the R language [25]. First, 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. Except for CPP as a function of voice type, all other comparisons were highly significant ($p < 2e^{-16}$). We highlight that the huge amount of data contributed for this significance (pre-scream: N=26405; scream: N=24499). Also, it may support a long-term acoustic analysis of voice quality.

Second, we separated the data in 2 groups (pre-scream and scream), and, for each group, ran an ANOVA with 13 parameters as a function of the song (FOI, RTH, NOB), in order to observe voice quality similarities among the songs (see tables 1 and 2). In the pre-scream part, we had a highly statistical difference for all parameters ($p < 2e^{-16}$). In order to observe whether all the songs were statistically different from each other, a subsequent Tukey HSD post-hoc was run. 11 parameters were statistically different for all songs, except for H1* ((FOI = NOB) ≠ RTH) and HNR25 ((FOI = RTH) ≠ NOB). In the scream part, we had also a highly statistical difference for all parameters ($p < 2e^{-16}$). The Tukey HSD test showed the songs behaved differently for 11 parameters, except H1*H2* ((FOI = RTH) ≠ NOB) and F1 ((FOI = NOB) ≠ RTH).

Finally, we separated the songs, and, for each song, ran a t-test with the 13 voice quality measures as function of voice type (pre-scream, scream), so as to test the hypotheses described in the previous section. See tables 1 and 2 for a reference to the measures.

The H1* hypothesis that higher values (in modulus) would be found for the "scream" part of the songs was fully corroborated for all songs (FOI, $p < 2e^{-16}$; RTH, $p<2e^{-16}$; NOB, $p < 2e^{-16}$).

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, $p < 2e^{-16}$; RTH, $p < 2e^{-16}$) Although the song NOB was statistically significant ($p < 2e^{-16}$), the value increased positively from the "pre-scream" to the "scream" part.

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 (FOI, $p < 2e^{-16}$; RTH, $p < 2e^{-16}$; NOB, $p < 2e^{-16}$).

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, $p< 2e^{-16}$;RTH, $p<2e^{-16}$). Although the song NOB was statistically significant ($p<2e^{-16}$), the value increased positively from the "pre-scream" to the "scream" part.

The Energy hypothesis that we would find a greater value of this parameter for the "scream" part was fully corroborated for all songs (FOI, $p < 2e^{-16}$; RTH, $p < 0.02$; NOB, $p < 2e^{-16}$).

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, cf. table 3).

The F1 hypothesis that higher F1s would be found in the "scream" part was fully corroborated for all songs (FOI, $p < 2e^{-16}$; RTH, $p < 4.88e^{-10}$; NOB, $p < 2e^{-16}$).

The F2 hypothesis that higher F2s would be found in the "scream" part was fully corroborated for 2 songs (RTH, $p < 2e^{-16}$; NOB, $p < 2e^{-16}$), and marginally corroborated for FOI ($p < 0.059$).

The B1 hypothesis that higher B1s would be found in the "scream" part was not supported by the data. Instead of increasing bandwidth from the "pre-scream" part to the "scream part", there was a significant decrease of this parameter in this direction (FOI, $p < 2e^{-16}$; RTH, $p < 4.88e^{-10}$; NOB, $p < 2e^{-16}$).

The B2 hypothesis that higher F1s would be found in the "scream" part was corroborated only for RTH ($p < 2e^{-16}$). For the other songs the hypothesis was not supported, since there was a significant decrease of this parameter from the "pre-scream" part to the "scream" part (FOI, $p < 2e^{-16}$; NOB, $p < 2e^{-16}$).

Table 1: *Voice quality measures (mean) for the pre-scream part. S stands for Song; N, HNR.*

S	H1	H1H2	H1A3	CPP	Energy
FOI	-5.27	0.47	-2.52	19.1	0.62
RTH	-0.26	2.77	0.53	17.8	1.94
NOB	-5.33	-0.26	-0.38	14.8	1.79

S	F1 (B1)	F2 (B2)	N5	N15	N25	N35
FOI	628 (275)	1383 (197)	36	24	21	22
RTH	589 (176)	1241 (183)	34	22	22	22
NOB	566 (184)	1294 (244)	23	16	16	18

Table 2: *Voice quality measures (mean) for the scream part. S stands for Song; N, HNR.*

S	H1	H1H2	H1A3	CPP	Energy
FOI	-15.79	-0.44	-13.45	16.7	1.41
RTH	-10.08	-0.57	-10.69	15.1	1.98
NOB	-8.94	2.62	-3.48	17.4	3.83

S	F1 (B1)	F2 (B2)	N5	N15	N25	N35
FOI	776 (61)	1406 (103)	38	25	21	21
RTH	632 (101)	1317 (269)	30	16	14	13
NOB	774 (137)	1446 (206)	33	19	15	14

Table 3: *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***

Finally, we would like to comment that all measures were made based on the singing of very high notes, what generates very high f0s compared to the normal male voice, which is known to be around 100 Hz (see table 4). Also, these f0s match the expected note frequencies represented in the musical scores (figures 1-4).

Table 4. *f0 mean, standard deviation (SD), minimum value (min), and maximum value (max). S stands for Song; SC, scream; PS, for pre-scream.*

S	Mean	SD	Min	Max
FOI (SC)	665.0	225.7	298.0	997.7
FOI (PS)	518.0	67.2	308.8	613.6
RTH (SC)	702.2	132.5	310.6	871.6
RTH (PS)	448.2	57.1	294.9	640.2
NOB (SC)	625.8	140.0	319.0	927.1
NOB (PS)	462.4	65.4	308.1	656.2

5. Discussion

This study shows that twelve voice quality parameters chosen for analysis were capable of discriminating two broad categories of voice quality: “pre-scream” and “scream”. The only parameter that was not consistent for this discrimination was CPP.

Another point of investigation here was whether there were similarities among the songs because of the heavy metal singing style. In the “pre-scream” part, our results reveal for H1* that *Flight of Icarus* (FOI) was similar to *The Number of the Beast* (NOB), and that both were different from *Run to the Hills* (RTH). In addition, FOI is similar to RTH, and that both songs are different from NOB, regarding the HNR25 parameter. In the “scream” part, FOI was similar to RTH, and both songs different from NOB, considering the H1*H2* parameter. Moreover, FOI is similar to NOB, and both songs different from RTH, for the F1 parameter. Therefore, we found evidence of similarities among the songs for some voice quality parameters.

Our results have also shown that the most robust parameters for differentiating the two different singing strategies (“pre-scream” x “scream”) for singing at extremely high registers of the male range were H1*, H1*A3*, Energy, F1, F2, HNR25, and HNR35. At least for these parameters, the hypotheses were fully corroborated. Therefore, in our study, these parameters were the more relevant for distinguishing two different voice qualities for singing the same song.

The other parameters that not validate the hypotheses may be due to the way we consider voice quality in this study. As presented in the introduction, voice quality is considered as a long-term analysis, so we analyzed a long stretch of speech without considering minor variabilities of voice quality within the speech signal. This is a matter that needs to be taken into consideration for future developments of our singing analysis. As an example, we checked the H1*H2* values for NOB and realized that in some parts of the signal we had evidence in the direction predicted by the hypothesis. The counterevidence of the hypothesis may be related to the greater standard deviation of the “pre-scream” part (7.50) in comparison to the scream part (3.05). Similarly, for CPP in NOB, greater standard deviation was found for the “scream” part (5.2) in comparison to the “pre-scream” part (3.89).

6. Conclusion

This study is a further development of the new methodology presented in our previous study [1]. Here we opted for working with pre-recorded heavy metal songs by Iron Maiden, so as to verify the validity of using the VoiceSauce analysis [13], allied to Praat analysis [24], for studying extremely high-pitched singing. Our promising results have shown that this methodology was robust enough to analyze this kind of vocal performance.

The next step of the research is to compare Bruce Dickinson’s voice quality in this study with the voice quality of the professional speakers in our previous study [1].

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 adapting the VPAS model [8, 9, 10, 11] for singing analysis.

7. Acknowledgements

The authors would like to thank 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.

8. References

[1] A. Meireles. F. G. Cavalcante. Qualidade de voz no estilo de canto heavy metal. *Per musi* (UFMG), v. 32, p. 197-218, 2015.

[2] J. Laver. *The phonetic description of voice quality*. Cambridge: Cambridge: Cambridge University Press, 1980.

[3] G. Pecoraro, A. Duprat, S. Bannwarth and M. Andrada e Silva. "Cantores de rock: ajustes dinâmicos de trato vocal, análise perceptivo-auditiva e acústica das vozes ao longo de cinco décadas". In: *Anais do 18o Congresso Brasileiro de Fonoaudiologia*. Curitiba, 2010.

[4] L. Oliveira and M. Behlau. *Perfil vocal de cantores amadores de banda de roque*. [monografia]. São Paulo(SP): Centro de Estudos da Voz, 2004.

[5] M. Thalen and J. Sundberg. "Describing different styles of singing- a comparison of a female singer's voice source in ''Classical'', ''Pop'', ''Jazz", and ''Blues". Log Phon Vocol. v.26, p.82-93, 2001.

[6] A. Gonsalves, E. Amin and M. Behlau. "Análise do grau global e tensão da voz em cantores de roque". Pró-Fono R. Atual. Cient. [online]. vol.22, n.3, p.195-200, 2010.

[7] G. Bayer. *Heavy metal music in Britain*. London: Ashgate popular and folk music series, 2009.

[8] J. Laver. "Phonetic evaluation of voice quality". In: *Voice quality measurement.* R.D Kent, Ball M.J. (ed). San Diego: Singular Publishing, p.37-48, 2000.

[9] J. Laver, S. Wirs, J. Mackenzie and S. M. Hiller. *A perceptual protocol for the analysis of vocal profiles*. Edinburgh; Edinburg University. Department of Linguistics; p.139-55. [Workin Progress, 14], 1981.

[10] J. Laver and J. Mackenzie-Beck. *Vocal Profile Analysis Scheme – VPAS*. Queen Margareth University College-QMUC, Speech Science Research Centre, Edinburgh, 2007.

[11] J. Mackenzie-Beck. "Perceptual analysis of voice quality: the place of vocal profile analysis". In: *A figure of speech: a festschrift for John Laver*. W.J. Hardcastle, J. Mackenzie-Beck (ed). Mahwah: Lawrence Erlbrum, p.285-322, 2005.

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

[13] Y.-L. Shue, P. Keating, C. Vicenik and K. Yu. "VoiceSauce: A program for voice analysis", *Proceedings of the ICPhS XVII*, 1846-1849, 2011.

[14] 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.

[15] 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.

[16] D. Klatt and L. Klatt, L. "Analysis, synthesis, and perception of voice quality variations among female and male talkers". *J. Acoustic. Soc. Amer.*, Vol. 87: 820-857, 1990.

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

[18] M. Gordon and P. Ladefoged. "Phonation types: a cross-linguistic overview". *J. of Phonetics* 29, 383-406, 2001.

[19] M. K. Huffman. "Measures of phonation type in Hmong". *J. Acoust. Soc. Am*., 81(2):495–504, February 1987.

[20] E. Fischer-Jorgensen. "Phonetic analysis of breathy (murmured) vowels in Gujarati". *Indian Linguist*, 28:71–139, 1967.

[21] M. Södersten and P.-A. Lindestad. "Glottal closure and perceived breathiness during phonation in normally speaking subjects." *J. Speech and Hearing Research*, 33:601–611, 1990.

[22] A.M.C. Sluijter and V.J. Van Heuven. "Spectral balance as an acoustic correlate of linguistic stress." *J. Acoust. Soc. Am.*, 100(4):2471–2485, 1996.

[23] M. Iseli, Y.-L. Shue, M. Epstein, P. Keating, J. Kreiman, and A. Alwan. "Voice source correlates of prosodic features in American English: A pilot study." In *Proceedings of Interspeech*, pp. 2226–2229, Pittsburgh, PA, September 2006.

[24] P. Boersma and D. Weenink. "Praat: Doing phonetics by computer (version 4.5.06)," http://www.praat.org/ (Last viewed December 8, 2010), 2006.

[25] R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/, 2013.