

Dysphonia is beautiful: A perceptual and acoustic analysis of vocal roughness

Melissa Barkat-Defradas, Camille Fauth, Ivana Didirkova, Benoit Amy de La Bretèque, Fabrice Hirsch, Christelle Dodane, Jérémi Sauvage

▶ **To cite this version:**

Melissa Barkat-Defradas, Camille Fauth, Ivana Didirkova, Benoit Amy de La Bretèque, Fabrice Hirsch, et al.. Dysphonia is beautiful: A perceptual and acoustic analysis of vocal roughness. 18th International Congress of Phonetic Sciences (ICPhS-18), Aug 2015, Glasgow, United Kingdom. hal-01820458

HAL Id: hal-01820458
https://hal.science/hal-01820458

Submitted on 21 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DYSPHONIA IS BEAUTIFUL
A PERCEPTUAL AND ACOUSTIC ANALYSIS OF VOCAL ROUGHNESS

[1,2]Melissa Barkat-Defradas, [3]Camille Fauth, [1]Ivana Didirkova, [4]Benoit Amy de la Bretèque,
[1]Fabrice Hirsch, [1]Christelle Dodane, [1]Jérémi Sauvage

[1]Praxiling UMR 5267 CNRS - Université Paul-Valéry Montpellier
[2]Institut des Sciences de l'Evolution, CNRS-UMR 5554
[3]ER 1339 LiLPa - Equipe Parole et Cognition Université de Strasbourg
[4]Laboratoire Parole et Langage UMR 7309 CNRS - Université Aix-Marseille

ABSTRACT

Researchers as well as speech therapists are interested in the determination of reliable acoustic cues that may be useful for the evaluation of vocal quality as well as for the diagnosis of vocal pathologies and remediation. In this way, experimental phonetics can be useful to clinical practice. This work which tries to connect phonetics and logopedic science, deals with the esthetic quality of dysphonic voices. The aim of this work is to evaluate if women can judge as attractive some masculine voices in spite of their pathological dimension. The results show that voices that are slightly rough (i.e. R1 on the GRBAS scale) are evaluated as the most attractive among a set of dysphonic and non-dysphonic voices. An acoustic study was carried out to quantify the acoustic characteristics of each type of pathological voices and to examine the acoustic correlates of voices that were perceived as the most attractive.

Keywords: voice quality, dysphonia, perceptive evaluation, attractiveness.

1. INTRODUCTION

The question of esthetics in speech has given rise to increasing interest from both the community of researchers in speech and speech professionals - implicated in the processes of vocal rehabilitation, as shown in [6] and [13]. This can be explained by the fact that speech is, among others, a tool for communication and social interactions, which strongly influences the image an individual can have on his or her interlocutor. Indeed, most studies, which have addressed the perceptive evaluation of the voice, have shown that it is possible - from vocal characteristics alone - to infer a certain number of physiological characteristics [10] such as: gender [12]; age [17] or height and weight [11]. Other research studies have shown a correlation between, the perceived vocal quality on the one hand and, facial attractiveness on the other hand [7]; sexual tendencies and/or behaviour [14], certain personality traits [3] or even characteristics related to the socioeconomic status of speakers [8]. This effect related to the vocal attractiveness stereotype refers one back to the idea that "*what sounds beautiful is good*" [23] which suggests that vocal pathology is perceived as being negative. However, we think that certain pathological voices can be perceived as attractive. Indeed, even if the way we judge a voice is based on quite personal factors, some of them, quite frankly somehow pathological, seem to be appreciated by a large number of people, such as for instance the voices of such artists as the singers Joe Cocker and Garou or the actresses Lindsay Lohan and Demi Moore. Although, apart from a study conducted on the topic of singing voice [19], the esthetic dimension of pathological voices is a question, which as far as we know, has not yet been tackled in the literature. The objective of the present study is quite precisely to evaluate, through a perceptive experience, the esthetic dimension of dysphonic voices, and to determine the acoustic indexes, which have an effect on judges' hedonic judgment.

2. RESEARCH HYPOTHESIS

The first objective of this research is to find out if certain dysphonic voices can be qualified as attractive. Here, we focused on the question of man's voices potential for seduction in the context of voice pathology. This is why the group of judges is exclusively composed of female subjects judging masculine voices.

Our second objective is to identify some of the acoustic characteristics of the most attractive pathological voices and to observe what makes them different from the dysphonic voices perceived as less attractive. As shown in [5] or [7], when they have to choose between several non-pathological masculine voices, women tend to perceive as most attractive the deepest voices. Likewise, according to [20], a certain number of dysphonic voices are themselves deeper than the average.

Consequently, our main hypothesis is that mild dysphonic voices, that are regularly perceived as deeper than healthy voices, should be perceived as more attractive than some normal (i.e. non pathological) voices. Our second hypothesis is that some vocal characteristics, such as roughness, could give to mild dysphonic voices a certain esthetic dimension.

3. MATERIAL AND METHOD

3.1 The acquisition of voice stimuli

In order to test our hypotheses, auditory stimuli produced by dysphonic male speakers (n = 31) and non-dysphonic speakers (n= 31) were selected (see Table 1). The pathological voices stimuli were collected from the EVA database of the Hospital and Research Centre La Timone in Marseille (France). Note here that voices were categorized as healthy and/or pathological on the basis of Hirano's GRBAS scale [9], which gives scores of 0, 1, 2, or 3 for the Grade of hoarseness; Roughness, Breathiness, Asthenia, and Strain, where 0 is normal, 1 is a slight degree, 2 is a medium degree, and 3 is a high degree. For each pathological voice, we used a recording of the first paragraph of the text "*La chèvre de Monsieur Seguin*" as well as recordings of a 3second sustained [a]. We recorded the healthy voices ourselves using a semi-professional equipment (Zoom H2, 22 kHz, 16 bits, mono). Subjects were placed in a standing position with a microphone placed at 15 cm from the mouth. They were requested to read out as naturally as possible the text, with their habitual voice.

Table 1: Distribution of voices according to grade and age

*****Inclusion criteria/age: 20-60 years (to avoid the effects of molting or presbyphonia)

Grade	N° of Subjects/62	Mean age*
G0	n=31	39.3
G1	n=14	43.7
G2	n=17	40.2

Subjects were free to become familiar with the text as much as they wished before the recording started. Each subject was recorded three times and we selected the best recording (i.e. no background noise, no hesitation, etc.). Before being submitted to the naive jury's judgment, healthy and dysphonic voices were subjectively evaluated by a jury of experts made of speech therapists and a laryngologist. This jury determined, through consensus and for all voices, the components G, R and B from Hirano's GRBAS scale. Components related to asthenia (A) and to voice strain (S) were not studied here. Inter-judge's reliability showed no significant difference. Table 2 gives the detailed proportion of voices that were evaluated as belonging to different degrees of roughness (R0, R1, R2) by the jury of experts.

Table 2: Distribution of voices according to grade

Grade	R0	R1	R2
G0	0%	0%	0%
G1	42.86%	50%	7.14%
G2	17.65%	17.65%	64.7%

The pathologies represented in this study are listed in Table 3. Only cases of functional or benign organic dysphonia were included. Malignant lesions were not included as they are too serious diseases to be included in a study concerning vocal esthetics. In the same way, as shown in [1] or [15], severe dysphonic voices (i.e. G3), known to be regularly perceived as unpleasant, were discarded. This brings the number of pathological voices stimuli used in our study to n=28.

Table 3: Causes and subjects in vocal pathologies

Lesions	Subjects/31
Polyps	n=12
Vocal fold paralysis	n=4
Granuloma	n=1
Kyst	n=1
Sulcus	n=1
Nodule	n=1
Paresis + atrophy	n=1
Reinke's Edema	n=1
Oval shaped leak	n=1
Normal larynx	n=5

3.2 Perceptive experience

In order to reduce the duration of the perceptive experiment, 34 out of 62 available voice stimuli were randomly selected for a perceptive experiment. Thus, 17 G0 voices, 10 G1 voices and 7 G2 voices were retained. Additionally, 6 extracts from these 34 samples were presented twice to measure each judge's response consistency. The 40 extracts (mean duration = 20 seconds) were presented to a naive jury of 92

women aged between 20 and 60 years (mean age = 35 years ± 12.6). In order to avoid the influence of perceived age on hearers' hedonic judgment, we matched as much as possible the judges' characteristics to that of the speakers. Then a computer graphic interface was designed using Perceval© (http://www.lpl-aix.fr/~lpldev/perceval/). The test duration was approximately 40 minutes. Once subjects had answered a sociological questionnaire giving information about their age, profession and qualifications, they were asked to judge the stimuli according to their pleasantness, on a 6-point Likert scale, where 1 corresponds to "*absolutely not pleasant*", 2 "*not pleasant at all* ", 3 "*not pleasant*", 4 "*rather pleasant*", 5 "*very pleasant*" and 6 "*extremely pleasant*". Listening occurred individually, in a comfortable and quiet environment using a PC-computer and professional headphones (eyerdynamic DT-770 Pro). Listeners were not aware they were judging normal as well as pathological voice.

4. RESULTS

4.1 Perception

Table 4, 5 and 6 below shows the mean evaluations (out of 6 points) which were attributed to each grade of dysphonic voices and for each of the GRBAS parameters under study (*i.e.* R and B). It is important to note that the analysis of intra-subject response consistency has not led to significant differences during the evaluation of the 6 repeated vocal stimuli (p=0.8).

Table 4: Mean evaluation according to grades (G)

Grade	Mean evaluation/6	sd
G0	3.52	1.11
G1	3.24	1.13
G2	2.76	1.24

Table 5: Mean evaluation according to the degree of roughness (R)

Grade	Mean evaluation/6	sd
R0	3.17	1.15
R1	3.43	1.13
R2	2.70	1.23

Table 6: Mean evaluation according to the degree of breath (B)

Grade	Mean evaluation/6	sd
B0	3.44	1.15
B1	3.20	1.13
B2	2.81	1.23

We carried out a statistical analysis using multiple tests (repeated measures ANOVA and post hoc Tuckey tests) to see which voices lead to the best evaluations. Our results show that dysphonia grade (G), roughness (R) and breath (B) have a significant effect on evaluations: basically we can say that the more dysphonic the voices, the more negative the judges' scores (p<.05).).

With regard to the grade, scores given to G0 voices are significantly higher than those given to dysphonic voices of G1 and G2 type (p<.0001 and p<.0001 respectively). Similarly, a comparison of evaluations attributed to dysphonic voices as a function of the degree of pathological severity indicates that G1 voices are significantly judged as more attractive than G2 voices (p<.0001).

Results concerning Roughness parameter reveal that medium rough voices are perceived as more attractive than R0 voices (p=.0001) and R2 voices (p=.0026). Thus, on the basis of these criterion alone, healthy voices do not yield to better evaluations than pathological ones. Results concerning breath show that B0 voices are better evaluated than B1 and B2 voices (p<.0001 and p<.0001 respectively), mild breathy voices (B1) being significantly better judged than B2 voices (p=0.0016).

2.4. Acoustics

In order to understand better the underlying acoustic characteristics of the judges' preferred voices, we performed an acoustic analysis using the Praat software [4]. Measurements of fundamental frequency (f_0) were carried out. With this measure, it was possible to find out if the judges were more attracted to deep voice pitch (a result regularly observed in the literature) or if their judgment was related to some other vocal characteristics. The Harmonics-to-Noise Ratio (HNR) was also quantified to observe if this parameter, which gives an indication of the noise/harmonics ratio in a voice, was a reliable feature to qualify a voice reported as attractive. Let's remind here that according to [22], the lower the ratio is, the more the signal is invaded by noise, and consequently, the rougher is the voice. The *jitter*, (i.e. a measure of short term disturbances in the fundamental frequency [18]), was also quantified.

Even though a great number of vocal measures are possible, we chose only three here to see whether female listeners were more attracted to the pitch of a voice or to some other characteristics. In other words, we seek to find out if the preference observed for mild rough voices can be explained by some particular acoustic parameters and, if so, which one. Measurement of vocal dysperiodicities and signal instability (HNR and jitter) were conducted on sustained vowels whereas measurements of F0 were performed on the text reading condition.

The fundamental frequency measures did not reveal any significant difference between the two groups (Table 7). Indeed, the mean F0 reaches 123 Hz (SD= 20 Hz) for the R1 voices whereas mean R2 F0 is around 127 Hz (SD = 35 Hz). Therefore, voice pitch does not seem essential for the judges' preferences, and it is thus interesting to observe the results which more directly concern the timbre of a voice.

Table 7: HNR, jitter and f_0 mean measures in terms of Roughness

Roughness	HNR (sd)	Jitter (sd)	f_0 V midpoint (sd)
R1	22 (3)	0,65 (0,14)	123 (20)
R2	14 (4)	1,07 (0,28)	127 (35)

As the fundamental frequency does not help for the discrimination of R1 and R2 voices, this is not true of the *jitter*. Indeed, this parameter was evaluated at 0.65% for R1 voices whereas it reaches 1.07% for R2 voices (note that the threshold generally reported to distinguish a normal voice from a pathological one equals 1.04% in the literature, e.g. [2]). To summarize, the *jitter*, which informs on the degree of f_0 disturbance, is slightly above the normal limit for R2 voices, whereas the R1 group show values within the norm for this parameter.

Turning to the HNR, our results show differences between the groups R1 and R2. This ratio reaches 22 (SD = 3) for speakers R1 voices and 14 (SD= 4) for R2 voices. Thus, the noise quantity in the voice is higher for the second group, suggesting more vocal roughness in R2 voices. From the clinical point of view, roughness indicates the inability of all, or a part of, the vocal fold mucus to be distorted during the vibration. For therapists, roughness is evidence for rigidity. Nevertheless, as far as the naive judge is concerned, our results suggest that this parameter is perceived – up to a certain point – as attractive and less esthetic above these values. It would therefore be justified to study further the question of a threshold for attractive vs. unattractive roughness.

5. CONCLUSION

Phonation is a complex biomechanical phenomenon, which is still not entirely understood. Voice alterations can lead to important disturbances for communication, but also for patient's identity and social representation. From the point of view of vocal characteristics alone, the degree of pathological severity seems to have a definite impact on the attractiveness dimension of the speaker's voice, at least with respect to the grade (G) and breathiness (B) components. As for roughness (R), our study shows that this parameter represents, up to a point, a significant factor for vocal attractiveness. Although we strongly support the view that roughness must be considered as an important symptom of illness, it seems that slightly hoarse voices have been, at the beginning of this 21st century, promoted to the rank of new vocal esthetics.

6. REFERENCES

[1] Altenberg, E., Ferrand, C. 2006. Perception of individuals with voice disorders by monolingual English, bilingual Cantonese-English, and bilingual Russian-English women. *Journal of Speech, Language and Hearing Research*, 49, 879-887.

[2] Baken, R.J., Orlikoff, R.F. 2000. *Clinical measurements of speech and voice* (2nd ed.), San Diego, CA: Singular Publishing Group.

[3] Barkat-Defradas, M., Raymond, M. 2011. Speech tempo: an honest signal for selecting mate for reproduction? Paper presented at the *23rd Annual Conference of Human Behavior & Evolution Society*, June 29 – July 3, 2011, Montpellier (France).

[4] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345.

[5] Bruckert, L., Lienard, J.S., Lacroix, A., Kreutzer, M., Lebourcher, G. 2006. Women use voice parameters to assess men's characteristics. In: *Proceedings of Biological sciences*, Vol. 273(1582), pp. 83–293. The Royal Society of London

[6] Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G, Kawahara, H., Belin, P. 2010. Vocal Attractiveness Increases by Averaging, *Current Biology* 20, 116–120.

[7] Collins, S. A. 2000. Men's voices and women's choices. *Animal Behaviour*, 60, 773–780.

[8] Harms, L. S. 1963. Listener comprehension of speakers of three status groups. *Language Speech*, 4, 109–112.

[9] Hirano, M. 1981. *Clinical examination of voice*, Springer Verlag, New York.

[10] Krauss, R.M., Freyberg, R., Morsella, E. 2002. Inferring speakers' physical attributes from their voices. *Journal of Experimental Social Psychology* 38, 618–625.
[11] Lass, N.J., Phillip, J.K., Bruchey, C.A., 1980. The effect of filtered speech on speaker height and weight identification, *Journal of Phonetics* 8, 91– 100.
[12] Lieberman, P., 1984. *The biology of and evolution of language*. Cambridge, Massachusetts: Harvard University Press.
[13] Liu, X., Xu, Y. 2011. What makes a female voice attractive? *Proceedings of XVIIth ICPhS*, Hong Kong, 17-21 August 2011, 1274–1277.
[14] Pierrehumbert J. B., Bent T., Munson B., Bradlow A. R., Bailey J. M. 2004. The influence of sexual orientation on vowel production, *Journal of Acoustical Society of America* 116 (4), 1905–1908.
[15] Raymond, F. 2010. *L'image sociale véhiculée par la dysphonie*, Mémoire de fin d'études en orthophonie, Université de Marseille II, Faculté de Médecine.
[16] Révis, J. 2004. *L'analyse perceptive des dysphonies : approche phonétique de l'évaluation vocale*, Thèse de doctorat Nouveau régime en Sciences du Langage, Université de Provence.
[17] Ryan, W.J., Burk, K.W. 1974. Perceptual and acoustic correlates of aging in the speech of males. *Journal of Communication Disorder*, 7, 181–192.
[18] Schoentgen, J. 2001. Stochastic models of jitter. *The Journal of the Acoustical Society of America*, 109(4), 1631-1650.
[19] Seidner, W., Büttner, M. 1999. Le caractère esthétique des voix chantantes rauques, *Bulletin d'audiophonologie*, Annales scientifiques de l'Université de Franche-Comté Médecine & Pharmacie, Vol. 15(1), 85–89.
[20] Teston, B., (document électronique). *L'évaluation objective des dysphonies: Etat actuel et perspectives d'évolution*, Laboratoire Parole et langage Université de Provence et CNRS [http://hal.arhivesouvertes.fr/docs/00/17/35/53/PDF/2075.pdf] - consulté le 22/03/2012.
[21] Vaissière, J. 2006. *La phonétique*. Collection QueSais-Je? N° 637 Presses Universitaires de France.
[22] Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *The Journal of the Acoustical Society of America*, 71(6), 1544-1549.
[23] Zuckerman, M., Driver, R. 1989. What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior*, 13, 67– 82.