

# Auditory-Perceptual Evaluation of Dysphonia: A Comparison Between Narrow and Broad Terminology Systems

\*Jenny Iwarsson, †Anne Bingen-Jakobsen, †Ditte Søbæk Johansen, ‡Inge Ernst Kølle, ‡Solveig Gunvor Pedersen, §Stine Løvind Thorsen, and \*Niels Reinhold Petersen, \*§Copenhagen, †Roskilde, and ‡Køge, Denmark

**Summary: Objective.** In auditory-perceptual voice analysis, a multiparameter approach and a more reductionist approach may be compared with narrow and broad phonetic transcription and used interchangeably, depending on the purpose. The aim of this study was to investigate the perspectives of a translation of the terminology used in the multiparameter Danish Dysphonia Assessment (DDA) approach into the five-parameter GRBAS system.

**Methods.** Voice samples illustrating type and grade of the voice qualities included in DDA were rated by five speech language pathologists using the GRBAS system with the aim of estimating inter- and intrarater reliability. The same samples were then rated using the DDA terminology.

**Results.** Both inter- and intrarater reliability were found to be very high for the GRBAS parameters grade, rough, and breathy, but somewhat lower for asthenic and strained. Further, strong and clear associations were found between the DDA and GRBAS rating for grade, rough, breathy, and strained, whereas the relation between DDA ratings and asthenic was weaker and less clear.

**Conclusion.** The data strongly support that the DDA system can be translated into the GRBAS system for auditory-perceptual voice analysis. The consensus discussion prior to the listening test is believed to have contributed to the high degree of inter- and intrarater reliability. We suggest for future use of the GRBAS system that rater reliability for asthenic and strained can increase, if these parameters are defined as behavioral terms and antagonists, reflecting muscular hypo- and hyperfunction.

**Key Words:** Auditory-perceptual voice analysis–Dysphonia–GRBAS–Listening test–Voice ratings.

## INTRODUCTION

Together with visual, acoustic, and physical examination, auditory-perceptual evaluation is a standard method in the analysis of voice disorders.<sup>1–3</sup> Basically, perceptual evaluation of voice means that the human ear is used in a systematic way to rate voice quality, using a set of terms representing specific features of the voice, and rated on scales representing the severity of each feature.<sup>1–5</sup>

Auditory-perceptual evaluation can serve different purposes. It can be used for assessment of treatment outcome comparing recordings from before and after treatment for the same individual(s). The method also has a long tradition among speech language pathologists (SLP) as a tool used for the first-visit assessment of a patient seeking help in a voice clinic, in which case the aim is to gain information relevant to diagnostics and treatment.

A number of systems have been developed for auditory-perceptual evaluation of the voice. The oldest and internationally most commonly used is the GRBAS system, which makes use of five parameters, defined by Hirano in 1981: Grade (G) represents the overall degree of hoarseness or voice abnormality.

Rough (R) represents a psychoacoustic impression of the irregularity of vocal fold vibrations. It corresponds to the irregular fluctuations in the fundamental frequency or the amplitude of the glottal source sound. Breathy (B) represents a psychoacoustic impression of the extent of air leakage through the glottis. It is related to turbulence. Asthenic (A) denotes weakness or lack of power in the voice. It is related to a weak intensity of the glottal source sound or a lack of higher harmonics. Strained (S) represents a psychoacoustic impression of a hyperfunctional state of phonation. It is related to an abnormally high fundamental frequency, noise in the high frequency range, or richness in high-frequency harmonics.<sup>6</sup> A 4-point (0–3) equal-appearing interval (EAI) scale or a visual analogue scale of, for example, 10 cm is typically used to rate the severity of each parameter. The GRBAS system has the great advantage of being simple and fast to use.<sup>1,7–9</sup> GRBAS has contributed a body of research allowing comparison of treatment outcomes as well as data on the reliability of the system.<sup>7,9–12</sup>

It could be argued that, as a clinical assessment tool, the GRBAS system has a shortcoming in being somewhat reductionistic. With only four specific parameters (the fifth, grade, referring to overall abnormality), important perceptual information relevant to both diagnostics and therapy may be lost. One example of reductionism in GRBAS is the broad rough parameter, which covers a number of voice qualities that are perceptually distinct; for example, gratings and diplophonia reflecting irregular vocal fold vibrations typically resulting from vocal fold pathology; coarse voice quality, which is noise accompanied by a very low pitch, often resulting from mucosal changes related to smoking; and vocal fry, which may be pathologic, but also

Accepted for publication July 10, 2017.

From the \*Department of Scandinavian Studies and Linguistics, Copenhagen University, Copenhagen, Denmark; †Department of Speech and Hearing Disorders, Roskilde, Denmark; ‡Department of Otorhinolaryngology and Maxillofacial Surgery, Zealand University Hospital, Køge, Denmark; and the §Department of Speech and Hearing Disorders, Copenhagen, Denmark.

Address correspondence and reprint requests to Jenny Iwarsson, Department of Scandinavian Studies and Linguistics, Copenhagen University, DK-23 00 Copenhagen, Denmark. E-mail: [jivarsson@hum.ku.dk](mailto:jivarsson@hum.ku.dk)

Journal of Voice, Vol. ■■■, No. ■■■, pp. ■■■–■■■  
0892-1997

© 2017 The Voice Foundation. Published by Elsevier Inc. All rights reserved.

<http://dx.doi.org/10.1016/j.jvoice.2017.07.006>

occur as a normal phenomenon of the speaking voice associated with sociolinguistic patterns.<sup>2,13,14</sup> It is unclear whether abnormal fundamental frequency fluctuation as in register breaks,<sup>9</sup> typically related to unstable muscle activity or lack of breath support, should also be registered under the rough parameter. Another example of information loss when GRBAS is used as a first-visit assessment tool is that aphonia is not registered separately. The occurrence of such intermittent or total lack of voicing may contribute valuable cues to underlying processes, apart from those associated with breathiness.<sup>15</sup> Moreover, if some listeners tacitly include a given perceptual feature (eg, aphonia) in one of the four specific parameters (eg, breathy) while other listeners exclude that feature from their ratings, the overall interrater reliability will suffer.

In contrast to GRBAS, a number of multiparameter systems for auditory evaluation have been developed in different countries over the years. Examples are the Australian *Perceptual Voice Profile*<sup>16</sup> and the Swedish *Stockholm Voice Evaluation Approach*.<sup>2,17–19</sup> It may be argued that a larger number of well-defined parameters make better use of the great potentials of human perception to detect auditory details and patterns relevant to voice function and thereby to meet the clinical demands of the SLP in diagnostics, goal-setting, and treatment planning.

In Denmark, a reference material for auditory-perceptual training has been developed and used for educational purposes since 2003.<sup>20–23</sup> The material contains voice samples illustrating type and grade of the voice qualities included in the Danish Dysphonia Assessment (DDA) protocol. The effect of auditory consensus training using this reference material was investigated in a study by Iwarsson and Petersen,<sup>24</sup> which showed very high values of both intra- and interrater reliability. The fact that values of interrater reliability were fairly high already before training (Cronbach's alpha across parameters being 0.913) supports the assumption that the set of parameters included in DDA is clearly distinct and allows robust ratings.

Summarizing, auditory-perceptual evaluation can serve different aims. If the aim is to measure treatment outcome from voice recordings of the same individual before and after treatment, the five-parameter GRBAS system may be sufficient. For this aim, it seems less important if a specific noise from the glottal voice source is categorized as, for example, gratings or vocal fry, but the GRBAS term rough may be sufficient. For a first-visit clinical assessment, on the other hand, where the goal is to obtain information relevant to diagnostics and choice of voice therapy content, the distinction between gratings and vocal fry may be important, and a more differentiated system may be more adequate for information-gathering and hypothesis testing.<sup>25</sup> The reductionist approach and the multiparameter approach may thus be comparable to broad and narrow phonetic transcription, and the choice between the two may be decided by the purpose. The aim of the present study was to compare how a group of listeners, trained with DDA but inexperienced with GRBAS, would use the two systems in rating the voice samples of the DDA reference material. Such a comparison should investigate the perspectives of a translation of the terms included in a multiparameter, “narrow” approach into the more reductionist “broad” approach. It was assumed that a translation from DDA ratings

into GRBAS would be possible if the DDA voice samples could be rated in a consistent manner with the GRBAS system and with fairly high inter- and intrarater reliability.

## METHODS

### Stimuli

The stimuli used in all three listening tests were 24 voice samples from the reference material for auditory-perceptual training with DDA mentioned above.<sup>22,24</sup> The voice samples were 15–22 seconds of reading of a standard text, the Danish version of “The North Wind and the Sun.” The samples included in the training material had originally been selected by four experienced SLPs from a large number of voice clinic recordings and validated by SLPs across Denmark. The selection criterion was that each of the samples should illustrate one voice quality alone, devoid of other features. The voice samples illustrating *hyperfunctional* and *breathy*, however, were an exception to this principle, included because this is a common combination among the voice disordered population. In the selection of the voice samples it was intended to have three grades of severity to each voice quality (1: mild, 2: intermediate, 3: severe). This proved impossible, however, for some voice qualities. [Appendix A](#) presents the definitions of the auditory-perceptual terms used in DDA, together with description of the voice samples included in the listening tests.

### Listeners

The listeners were five SLPs (A, B, C, D, E), with MA university degrees in speech pathology, all experienced users of DDA (>8 years) and well acquainted with the voice samples of the reference material. One of them (A) had also taken part in the original selection of the samples. None of the listeners had any experience with the GRBAS system for auditory-perceptual analysis.

### Procedure

Three listening tests were conducted, each with two sessions of 25 minutes with a break in between. In tests 1 and 2 (separated by 10 days), ratings were made using GRBAS with the purpose of investigating inter- and intrarater reliability. The definitions of the five GRBAS parameters by Hirano (1981) can be seen in [Appendix B](#). Before test 1, these were discussed in an oral consensus discussion, and notes from this discussion were read aloud before test 2. These notes can also be seen in [Appendix B](#). Except for this, neither ratings nor terminology was discussed between sessions or tests. Each voice sample was played three times for the listeners to make their rating of that voice, and the severity of each parameter was rated on a 4-point EAI scale (0–3). In test 3 (2 months after test 2), ratings were made using the DDA system. The aim of this was to investigate how the reference voice samples were rated using our familiar terminology and how these ratings related to the GRBAS ratings. As in tests 1 and 2, each parameter ([Appendix A](#)) was rated on a 4-point EAI interval scale (0–3). The terminology was not defined or discussed in association with listening test 3.

Download English Version:

<https://daneshyari.com/en/article/7533040>

Download Persian Version:

<https://daneshyari.com/article/7533040>

[Daneshyari.com](https://daneshyari.com)