2025 ACTA UNIVERSITATIS CAROLINAE PAG. 33–42
PHILOLOGICA 3

PERCEPTUAL EVALUATION OF THE EFFECT OF EXTERNAL RADIOTHERAPY IN THE NECK AREA ON CHANGES OF VOICE AND THE VOICE QUALITY OF CZECH PATIENTS

JAN ŠEBEK, TOMÁŠ BOŘIL

ABSTRACT
This research is focused on changes of voice and the voice quality before radiotherapy treatment and in time of one year after external radiotherapy (RT) in the neck area of Czech patients and to compare measurement for different subgroups of patients (by age, gender, surgical resection, aphonia, etc.). The perceptual test was performed on 16 patients undergoing external radiotherapy in the neck area and the changes of voice and the voice quality compared at before RT and 1, 6 and 12 months after RT. 2 clinicians and 1 trained voice specialist evaluated GIRBAS parameters of the voice quality and 96 lay listeners evaluated the scope of change of voice. The results of perceptual tests of lay listeners point to the difference of change in voice is the most pronounced in patients who had aphonia during the RT. Of the GIRBAS parameters, instability and roughness changed the most after RT treatment.

Keywords: changes of voice; voice quality; perceptual test; GIRBAS parameters; radiotherapy treatment; Czech patients

1. Introduction

In the last 20 years, there has been an increasing number of patients in the Czech Republic who had to undergo RT in the neck area as a part of their curative treatment of oncological diseases (Krejčí et al., 2022). One of the side effects of this treatment may be a temporary or permanent change of voice and the quality of the patients' voice and speech (Dršata et al., 2008).

Several research studies on this issue have been carried out internationally so far (Lazarus, 2009), but all of them were conducted especially from a medical perspective, focusing only on basic phonetic, especially acoustic, parameters (fundamental frequency, standard deviation of the fundamental frequency, jitter, shimmer, harmonicity, etc.) (Bagherzadeh et al., 2022; Bibby et al., 2007; Fung & Yoo, 2001), or questionnaire methods were used to examine the impact on the quality of life of patients after radiotherapy treatment (Mekiš et al., 2022; Šiupšinskienė et al., 2008) or some foreign studies focused on differences in voice quality when completing or not completing voice rehabilitation after radiotherapy treatment (Mei-jia et al., 2018; Tuomi et al., 2014).

https://doi.org/10.14712/24646830.2025.19

In addition to studying changes in acoustic parameters, it was also appropriate to focus research from a perceptual perspective. The voice quality assessment was performed using the GIRBAS method, which had become the most widely used voice scaling method by clinicians (Uloza et al., 2005; Karnell et al., 2007). The GIRBAS scale for subjective voice evaluation containing six voice-quality parameters, G (Grade of dysphonia), I (Instability), R (Roughness), B (Breathiness), A (Asthenia), and S (Strain), meet the criteria of reliability and relevance well, and the scale is easy to use (Olivares et al., 2023). Yamauchi et al. (2010) state that S parameter can be defined as extent to which strain or hyperfunctional use of phonation is heard, A parameter as degree of weakness heard in the voice, B parameter as degree to which air escaping from between the vocal folds can be heard by examiner, R parameter as impression of irregularity of the vibration of the vocal folds and I parameter as degree of change of the voice quality over time.

In the Czech environment, changes in the voice of patients before and after RT have not been yet evaluated by perceptual tests completed by lay listeners who, during the listening test, did not know that they were hearing the voices of patients before and after RT. From this perspective, it is beneficial to find out how the changes in the voice of these patients are perceived by public, whether the changes of voice are audible, and whether the lay listeners perceive the patient's voice before treatment compared to the voice after treatment as the voice of the same speaker.

The main aim of the present study was to investigate the effect of RT treatment on voice and evaluated the changes of voice and the voice quality perceptually. The changes of voice and the voice quality compared at before RT and 1, 6 and 12 months after RT. On these results analyse if they indicate any trend or are correlated to any properties of research group.

2. Method

2.1 Research group

The study was started after obtaining approval from the ethics committee of Bulovka University Hospital (No. 12.10.2021/10214/EK-Z). 16 respondents (4 women and 12 men) aged 19 to 76 years (average age 56 years) voluntarily participated in the research study and successfully completed the entire research study. All respondents were native Czech speakers, as required by the research study. All respondents at the Department of Radiation Oncology of Bulovka University Hospital in Prague underwent external radiotherapy treatment of the neck area as a part of their curative treatment of carcinoma in larynx area. The research group consisted of 2 smokers and 14 non-smokers, 5 of whom reported that they had been smokers in the distant past (more than 5 years ago). All respondents received a total radiation dose to the tumor site ranging from 66 to 77 Gray (Gy). 11 respondents also received curative radiation to the site of metastasis (neck node area), with a total dose of 54 to 66 Gy. 15 respondents underwent radiotherapy treatment in 33 fractions and 1 respondent in 35 fractions. All respondents received external radiotherapy to the neck area using the VMAT (Volumetric Modulated Arc Therapy) method

with IGRT (Image-Guided Radiation Therapy) on a linear accelerator with coverage of areas outside the planning target volume using a multi-leaf collimator in the normofractionation mode (5 fractions/week with a dose of up to 10 Gy/week). 4 respondents underwent surgical resection of part of the tongue before participating in this study, and 6 respondents had up to 2 weeks of aphonia during RT treatment.

2.2 Control group

The control group consisted of 4 Czech native speakers (2 women and 2 men) aged 38 to 52 years (average age 46 years) who were recruited at the start of this study. The control group was matched for age and gender. These speakers were representatives of a healthy population and had not any vocal abnormalities or diseases. Voice recordings of members of the control group were used to test the ability of lay and educated listeners to objectively evaluate the voice change and the voice quality of healthy and diseased voices of speakers.

2.3 Material

The recordings were obtained in the sound-treated recording studio of the Institute of Phonetics in Prague, using the high-quality AKG C4500 B-BC condenser microphone, with 48 kHz sampling frequency and 16-bit resolution. Four recording sessions were analysed in this study – first recording before RT, second in time of 1 month after RT, third at 6 months after RT and the last (fourth) at 12 months after RT. 20 tasks were recorded in each session with each respondent and once with each member of control group. Three of all tasks were used for perceptual test: spontaneous speech about some experiences of the last days (the first task, approx. 1 minute long), reading text of 5 sentences (the fourteenth task, see appendix) and sustained vowel /a/ (the nineth task).

2.4 Test of GIRBAS parameters

In order to evaluate the vocal competence of the patients before and after the RT treatment, the GIRBAS perceptual assessment scale were used to evaluate their voice quality. The evaluators evaluated voice quality of all respondents and all members of control group. Same set of three recordings (spontaneous speech about some experiences of the last days, reading text of 5 sentences and sustained vowel /a/) was used to evaluate each GIRBAS parameter.

The evaluation was made by 2 clinicians from Department of Phoniatrics at General Faculty Hospital in Prague and 1 trained voice specialist from Institute of Phonetics in Prague. The evaluation was made by giving a score from 0 to 3 (0 = normal, 1 = mildly affected, 2 = moderately affected, 3 = highly affected) including half grades for a finer scale of evaluation (0,5 = value between normal and mildly affected, 1,5 = value between mildly affected and moderately affected, 2,5 = value between moderately affected and highly affected).

Evaluators were allowed to replay any of the recordings. To minimize the order effect and also achieve the most objective assessment of the voice quality possible, evaluators did not know if the set of three recordings was recorded by member of research group or

member of control group and also did not know in which time of treatment the recordings were recorded. Sets were randomized and evaluators evaluated 4 sets weekly in half of year. Their ratings correlated with each other, so all data were used for analysis.

2.5 Listening test for lay listeners

For our perceptual test for lay listeners, we used recordings of reading same text of all respondents and all members of control group. In total, 96 recordings (48 pairs of recordings in 3 line-ups) of up to 20 seconds duration of each recording were used for the test. First 16 pairs of recordings (first line-up) were composed of recording of patient's voice before RT and recording of reading the same text by the same patient in time of 1 month after RT. Next 16 pairs of recordings (second line-up) were composed of recording of patient's voice before RT and recording of reading same text by same patient in time of 6 months after RT. And the last 16 pairs of recordings (third line-up) were composed of recording of patient's voice before RT and recording of reading same text by same patient in the time of 12 months after RT. The volume of the audio recording was unified to the same level for all recordings. The test was designed in PsyToolkit (Stoet, 2017) and it could be filled out at any time between March 15 and April 30, 2025. For each pair of recording, the listener had to decide, using a cursor on a scale, how much the speaker's voice had changed due to the disease, without the listener knowing what the disease was. On the scale, the minimum (value 0) was indicated as "no change in the voice of speaker" and the maximum (value 100) was indicated as "such a change in the voice that it is the voice of another speaker".

If a patient's recording was missing from the pair of recordings (there were 4 cases due to health reasons he was unable to attend the recording of the tasks), this recording was replaced with a recording of the same reading text spoken by the voice of a member of the control group of the same gender to test the ability of listeners to objectively evaluate the difference between healthy and diseased voices of speakers (we expected a rating close to 100 for these pairs of recordings).

Listeners were allowed to replay any of the recordings. In order to minimize the order effect, samples were randomized (in each line-up, as well as the line-ups themselves).

Besides the test, we gathered basic information from the listeners (gender, age, occupation). The test was successfully completed by 68 female and 30 male listeners, aged 20 to 79 (average age 45 years), all coming from the Czech Republic and from varied occupation (unemployed, hairdresser, social worker, courier, teacher, clerk, administrative worker, librarian, waiter, IT specialist, gardener, doctor, student, musician, physiotherapist, businessman, quality manager, HR specialist, designer, pensioner etc.). The test was intended and focused on lay listeners, precisely to get as close as possible to the perception of change of the voice from the perspective of the general public.

It was revealed later that two listeners had followed the test instructions in reverse (their responses reached negative values of the correlation coefficient against the other listeners), and their responses were eliminated. Therefore, 96 listeners' answers were analysed. The average time of the listeners finished the test (including the information questions) was 32 minutes.

2.6 Statistical analysis

Linear model and correlation analysis were conducted using stats package in R (R Core Team, 2024), data processing and visualisation was performed using tidyverse package (Wickham et al., 2019) and scatterpolar plots were plotted using plotly package (Sievert, 2020).

3. Results and discussion

External radiotherapy in the neck area has effect on changes of the parameters G (Grade of dysphonia), I (Instability) and R (Roughness) the most, while RT treatment has minimal effect on B (Breathiness), A (Asthenia) and S (Strain). Compared to the values of the GIRBAS parameters before RT treatment (marked as N1) and with the time interval after the end of RT (1 month after RT marked as N2, 6 months after RT marked as N3 and 12 months after RT marked as N4), the values of the I parameter increase (from value 0.66 to 1.17). The highest value of the G parameter was found in voices recorded with an interval of 6 months after RT treatment (value 1.31) and the highest value of the R parameter was found in voices recorded with an interval of 1 month after RT treatment (value 0.78) (see Table 1). Changes and dispersions of parameter values are depicted in the Figure 1.

Table 1 Average values of GIRBAS parameters.

Groups of respondents / time of recording	G	I	R	B	A	S
All members of research group / before RT (N1)	0.84	0.66	0.58	0.04	0.02	0.18
All members of research group / 1 month after RT (N2)	1.02	0.73	0.78	0.10	0.06	0.20
All members of research group / 6 months after RT (N3)	1.31	1.05	0.71	0.01	0.02	0.23
All members of research group / 12 months after RT (N4)	1.24	1.17	0.51	0.01	0.04	0.14
All members of control group / N/A	0.29	0.17	0.04	0.00	0.00	0.04

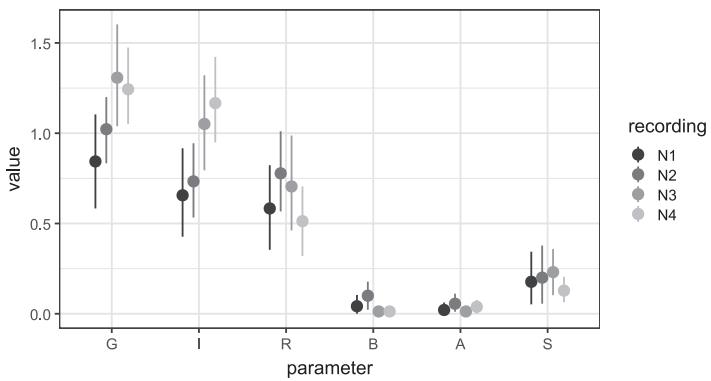

Figure 1 Changes of values of GIRBAS parameters of research group before (N1) and after RT (N2 = 1 month, N3 = 6 months, N4 = 12 months), point = mean value, error bars = 95% confidence interval estimated via a bootstrap method.

Comparing data by age, gender, aphonia, occurrence of metastases and surgical resection, no associations and correlation with the value and its change were found.

If we were to display the IRBAS parameters data using radar graphs, we can compare the graphs to see whether the overall voice impairment is worsening (distance from the origin is increasing) or improving (distance from the origin is decreasing). At the same time, we can determine from the shift of the pentagon shape defined by the IRBAS parameter values whether this shift is, for example, typical for the disease in question or is related to the treatment and its success (see Figure 2).

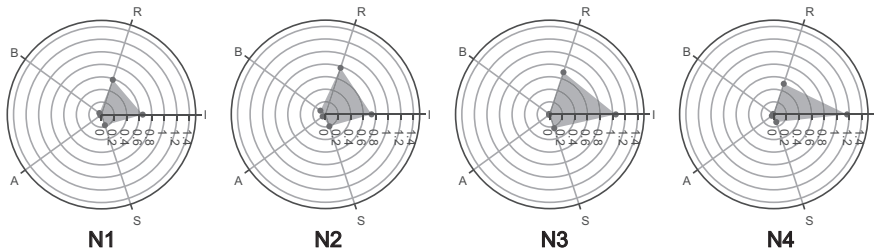

Figure 2 Radar plots of IRBAS pentagons of research group (scale zoomed).

Using different representations of changed values led us to the idea of a model that would most accurately determine the value of the G parameter from the values of other IRBAS parameters. If we assume that all other IRBAS components contribute to the total voice damage (parameter G), then we could determine that the relationship for the value of the total voice damage can be given by the linear model formula

$$G = \alpha + a.I + b.R + c.B + d.A + e.S + \varepsilon$$

where *a, b, c, d, e* are linear coefficients, α is the intercept value and ε is the error therm, the parameter G thus becomes a linear combination of the other parameters. From the actual IRBAS parameter values of this study, we obtained the formula as the most accurate relation for the G parameter:

$$G = 0.629I + 0.467R - 0.056B + 0.385A + 0.030S + \varepsilon$$

The found formula shows that the I, R, and A parameters contribute the most to the G parameter and the B and S parameters the least. Pearson correlation between actual values and the model predictions is 0.896.

Another approach to estimation of the G parameter is an Euclidean distance,

$$G = \sqrt{I^2 + R^2 + B^2 + A^2 + S^2} + \varepsilon$$

where ε is the error therm. Pearson correlation between actual values and the model predictions is 0.940.

The perception of lay listeners also confirmed that voice changes occur as a result of RT treatment. Based on data from a listening test for lay listeners, it was found that apho-

nia determines how much a change in voice will be perceived by the listener. The voices of patients who did not experience treatment-related aphonia were rated as having less change, while the voices of patients who experienced treatment-related aphonia were perceived as having more change and, in extreme cases, were rated as voices of different speakers (see Figure 3).

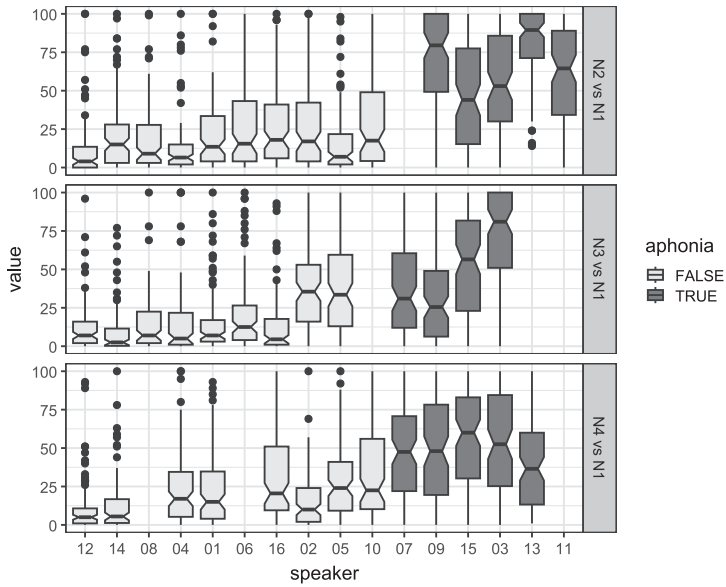

Figure 3 Evaluation of voice changes by patients according to aphonia.

Comparing results of listening test for lay listeners by age, gender, occurrence of metastases and surgical resection, no associations and correlation with the value and its change were found.

During the statistical processing of the data, we found that for the question of how much the listener perceives the change of the voice in a pair of recordings, it was not necessary to use such a fine scale of values as we used (0 to 100) because for the voices of patients

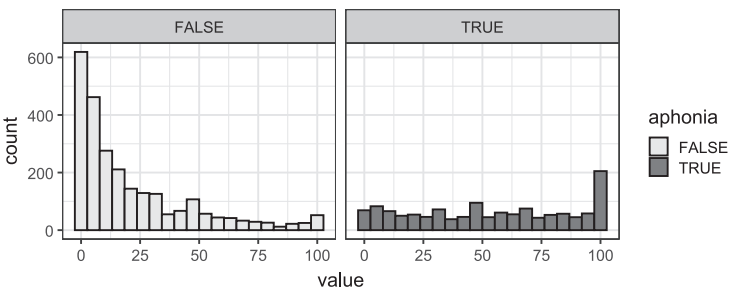

Figure 4 Histogram of listeners' values of voice changes (speakers without and with aphonia). The peaks around 0, 50 and 100 raise the question of whether listeners tend to use the scale in a rather discrete way.

in which there were small changes, listeners tended to give values close to zero, while for the voices of patients in which there were large changes, listeners also tended to give extreme values (close to 100), see Figure 4. Clearly, listeners more often preferred the exact midpoint of the scale (value 50) than the nearby values in both cases (with and without aphonia). It is likely that a discrete scale might be sufficient and more convenient to raters.

4. General discussion and conclusion

In our study, we examined the effect of external radiotherapy in the neck area on changes of voice and the voice quality of Czech patients by perceptual evaluation. It has been perceptually verified that educated and lay listeners perceive changes in the voice and voice quality of patients who have undergone RT treatment. The RT treatment thus most affects voice instability and temporarily roughness. The occurrence of aphonia will affect the perception of the voice change in these patients. In those who had aphonia during the RT treatment, the public perceives the voice change as very significant, audible, and identified as a completely different voice.

In conclusion, this study has shown that when comparing the results of voice recordings of 16 patients, it is not always possible to draw unambiguous conclusions for different groups according to age, gender, etc. However, despite the possible inhomogeneity of the research group, the occurrence of aphonia in these patients was shown to be the main factor of different perceptual evaluation.

Caution should be exercised when formulating conclusions from the analyses: different analysis methods may provide somewhat different results, and the interpretation may therefore be less unambiguous. A higher number of patients involved in the research study would contribute to greater statistical accuracy and more precise modelling.

Acknowledgements

This study was supported by Grant Agency of Charles University, project GAUK no. 26224.

REFERENCES

Bagherzadeh, S., Shahbazi-Gahrouei, D., Torabinezhad, F., Rabi Mahdavi, S., & Salmanian, S. (2022). The effects of (chemo) radiation therapy on the voice and quality of life in patients with non-laryngeal head and neck cancers: a subjective and objective assessment. *International Journal of Radiation Research*, *20*(2), 397–402.

Bibby, J. R., Cotton, S. M., Perry, A., & Corry, J. F. (2007). Voice outcomes after radiotherapy treatment for early glottic cancer: Assessment using multidimensional tools. *Head & Neck*, *30*(5), 600–610.

Dršata, J., Vokurka, J., Čelakovský, P., Hudíková, M., Růžička, J., & Kordač, P. (2008). Přehled foniatrických možností úpravy hlasu po onkologické léčbě nádorů oblasti hlavy a krku. *Onkologie*, *2*(2), 91–93.

Fung, K., & Yoo, J. (2001). Effects of head and neck radiation therapy on vocal function. *The Journal of otolaryngology*, *30*(3), 133–139.

Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., & Hoffman, H. T. (2007). Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. *Journal of Voice*, *21*(5), 576–590.

Krejčí, D., Mužík, J., & Dušek, L. (2022). *Novotvary 2019–2021 ČR – Cancer incidence 2019–2021 in the Czech Republic*. Ústav zdravotnických informací a statistiky ČR.

Lazarus, C. L. (2009). Effects of chemoradiotherapy on voice and swallowing. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *17*(3), 172–178.

Mei-jia, Z., Ji-wei, M., Xiang-ru, C., Xin, Z., & Chong, F. (2018). Effect of voice rehabilitation training on the patients with laryngeal cancer after radiotherapy. *Medicine*, *29*, 97–100.

Mekiš, J., Strojan, P., Mekiš, D., & Hočevar Boltežar, I. (2022). Change in Voice Quality after Radiotherapy for Early Glottic Cancer. *Cancers 2022*, *14*(12), 2993.

Olivares, A., Comini, L., Di Pietro, D. A., Vezzadini, G., Luisa, A., Boccali, E., Boccola, S., & Vitacca M. (2023). Perceptual and qualitative voice alterations detected by GIRBAS in patients with Parkinson's disease: Is there a relation with lung function and oxygenation. *Aging Clinical and Experimental Research*, *35*, 633–638.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Sievert, C. (2020). *Interactive Web-Based Data Visualization with R, plotly, and shiny.* Chapman and Hall/CRC Florida.

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31.

Šiupšinskienė, N., Vaitkus, S., Grėbliauskaitė, M., Engelmanaitė, L., & Šumskienė, J. (2008). Quality of life and voice in patients treated for early laryngeal cancer. *Medicina*, *44*(4), 288–295.

Tuomi, L., Andréll, P., & Finizia, C. (2014). Effects of voice rehabilitation after radiation therapy for laryngeal cancer: A randomized controlled study. *International Journal of Radiation Oncology, Biology & Physics*, *89*(5), 964–972.

Uloza, V., Saferis, V., & Uloziene, I. (2005). Perceptual and acoustic assessment of voice pathology and the efficacy of endolaryngeal phonomicrosurgery. *Journal of Voice*, *19*(1), 138–145.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Yamauchi, E. J., Imaizumi, S., Maruyama, H., & Haji, T. (2010). Perceptual evaluation of pathological voice quality: a comparative analysis between the RASATI and GRBASI scales. *Logopedics Phoniatrics Vocology*, *35*(3), 121–8.

APPENDIX

The assignment of the fourteenth task (reading text of 5 sentences), which were used for perceptual tests.

Prosím, přečtěte nahlas a zřetelně tento text:

Vítejte na naší frekvenci rádia F.F.U.K. Je tři čtvrtě na sedm a v půl deváté Vás budou čekat zprávy v českém jazyce. Po nich přiklopýtá kolegyně se zajímavostmi ze světa klokanů a kasuárů. Něco nám tu píská, šumí a ševelí. To budou určitě svišti na smetišti.

English translation:

Please read this text loud and clearly:

Welcome to our frequency of radio F.F.U.K. It is a quarter to seven and at half past eight, there will be news in Czech. After that, a colleague will shuffle in with interesting facts from the world of kangaroos and cassowaries. Something is whistling, rustling, and whispering here. It must be the groundhogs in the dumps.

RESUMÉ

Tento výzkum je zaměřen na změny hlasu a kvality hlasu před radioterapeutickou léčbou a v čase jednoho roku po zevní radioterapii (RT) v oblasti krku u českých pacientů a na porovnání změny hlasu a kvality hlasu u různých podskupin pacientů (podle věku, pohlaví, chirurgické resekce, afonie atd.). Percepční test byl proveden s nahrávkami hlasu 16 pacientů podstupujících RT a byly porovnány změny hlasu a kvalita hlasu před RT a 1, 6 a 12 měsíců po RT. Parametry kvality hlasu GIRBAS hodnotili 2 kliničtí lékaři a 1 vyškolený hlasový specialista a rozsah změn hlasu hodnotilo 96 laických posluchačů. Výsledky percepčního testu hodnoceného laickými posluchači ukazují na to, že změna hlasu je nejvýraznější u pacientů, u kterých se během RT vyskytla afonie. Z parametrů GIRBAS se vlivem RT léčby nejvíce změnila nestabilita a drsnost.

Jan Šebek
Institute of Phonetics
Faculty of Arts, Charles University
Prague, Czech Republic
jan.sebek@ff.cuni.cz

Tomáš Bořil
Institute of Phonetics
Faculty of Arts, Charles University
Prague, Czech Republic
tomas.boril@ff.cuni.cz