PERCEPTUAL AND ACOUSTIC ASSESSMENT OF A CHILD'S SPEECH BEFORE AND AFTER LARYNGEAL WEB SURGERY

GORDANA VAROŠANEC ŠKARIĆ[1], IVA BAŠIĆ[1], SINIŠA STEVANOVIĆ[2]

[1] University of Zagreb, Faculty of Humanities and Social Sciences, Ivana Lučića 3, Zagreb, Croatia, contact: gvarosan@ffzg.hr

[2] Sestre milosrdnice University Hospital Centre, Vinogradska cesta 29, Zagreb, Croatia

Received: 06.02.2023.
Accepted: 23.05.2023.

Original research article
UDK: 81'342.1:616.22-008.4
doi: 10.31299/hrri.59.1.5

Abstract: *The aim of this paper was to point to the importance of early diagnostics and surgery in patients with laryngeal web in order to achieve normal breathing, as well as to stress the need for an interdisciplinary approach to observing the quality of voice and prosodic features at an early age. The subject under consideration was a 6.5-year-old girl who had previously been diagnosed with irregular breathing (R06). An endoscopic exam revealed a laryngeal web between the vocal folds and the fact that the posterior intercartilaginous section of the glottis of the child's larynx was in order (normal). The child's speech had been recorded in the acoustic studio, both before and after the vocal-fold surgery (six and twelve months later). Due to severe dysphonia, difficulties with breathing, and frequent noisy breathing (stridor), we recorded only the phonation of the vowel [a], as well as spontaneous speech before the surgery. In addition, there was intense glottic and supraglottic strain before the surgery, which in phonetics corresponds to the term laryngeal and supralaryngeal strain and pathologically creaky whispery phonation (according to VPA protocol). This strain was visible in the area of the chest, neck, and head, as well as audible in the voice quality. Acoustic analysis showed that the average F_0 for the vowel [a] was remarkably high (442 Hz), and the pathological values were established using the following measures: local jitter (1.68%), local shimmer (0.7 dB), and the harmonic to noise ratio (17.6 dB). In contrast, six months after the surgery, the pitch for [a] was half the value of the preoperative one (220.5 Hz, $p < 0.001$), and the local jitter for all vowels (0.30-0.47%) and the harmonic to noise ratio (22.46 dB, $p = 0.05$) was within the normal range. There was also significant improvement in the F_0 values, standard deviation of F_0, and minimum and maximum F_0 values. The average and median F_0 values in spontaneous speech were also lower postoperatively. The voice quality showed a more balanced timbre (LTASS), particularly after one year. Some other prosodic features also showed improvement.*

Keywords: *laryngeal web, prosody, timbre, acoustic analysis, voice*

INTRODUCTION

The congenital category is one of the numerous aetiological factors of hoarseness (aetiology of dysphonia): it includes laryngomalacia, laryngeal web, and vocal fold cysts (Prstačić et al., 2020: 233).

The most common symptom that indicates the severity of hoarseness in patients with laryngeal web is, for example, stridor - a noisy, high-frequency sound that is heard during breathing, especially when inhaling, which affects the shortening of speech columns. Other common symptoms include breathing difficulties, hoarseness, lower voice, hoarse voice, and weak voice (e.g., Strakowski et al., 1988). Since vocal dysfunction is the most difficult symptom to treat, according to Cohen (1985), an interdisciplinary approach to solving the problem is rather important, as well as the cooperation between phono surgeons, speech therapists, and phoneticians, who continue to take care of the voice, pronunciation, intonation, and performance from the ancient predecessors of the *edomatores vocis*. Like them, today's phoneticians conduct exercises with different types of speakers, from professional speakers to those with pathological voices, because they have developed a different approach to improve voice aesthetics, and also for therapeutic purposes.

Laryngeal webs can be congenital or acquired. The category of congenital laryngeal webs is much rarer than the category of acquired webs (Yang et al., 2014). The laryngeal web - in the form of a membrane - is formed in the area of the *rima glottidis*, obstructing the passage of air towards the trachea, and can appear in different forms, from the less noticeable thin transparent membranes to thicker fibrous structures. They are most often formed on the anterior part of the glottis (98%): more specifically on the glottis (75%), the subglottis (12.5%), and the supraglottis (12.5%), while only 2% are formed on the back part of the glottis.

According to Cohen's classification of congenital glottic webs, there are four types of categories classified according to the degree of obstruction of the glottis and the thickness of the web (Cohen, 1985: 2): Type 1 – thin webs (anterior webs) that affect up to 35% of the glottis, mainly the front part of the glottis, and may extend slightly into the subglottic space. Vocal cords are always visible within the web, and the most common description of voice quality is mild hoarseness; Type 2 – thick or moderately thick webs, formed in the front part of the glottis and affecting 35-50% of the glottis, extending into the subglottic area, with the vocal cords visible within the web. The quality of the voice in type II is described as weak, quiet, whispery. The tone is not stabilised, and is generally described as a husky voice; Type 3 - laryngeal webs (anterior laryngeal webs) are tougher fibrous formations that affect between 50-75% of the glottis, spreading into the subglottic space, with the vocal cords potentially visible, although less often than in the first two types. Breathing is difficult, the voice is dysfunctional, very quiet, very weak, and whispery. These webs are usually thicker in the anterior portion than in the posterior one; Type 4 – in this type of the most severe form of laryngeal webs, thicker and firmer fibrous structures are formed, affecting 75-90% or more of the glottis (some sources state up to 99% of the glottis). They are equally thickened in both parts, i.e., in the front and the back, while the vocal cords are not visible. Airflow is severely reduced by 75-90%, stridor occurs regularly, as well as aphonia, dyspnoea, and asphyxia. This classification is generally accepted and used today when reporting on the laryngeal web in the paediatric (e.g., Lawlor et al., 2020: 234) and the younger population (Tiwari et al., 2018: 145-146). Type 3 is the most common type of congenital laryngeal web, and surgeries for types 3 and 4 are expected to be performed with a CO2 laser (LAVA). In countries with larger populations, around 50 cases of patients with laryngeal web are reported over several years (e.g., Strakowski et al., 1988). The rarity of these cases is evidenced by the fact that only two such procedures were performed at the UHC Zagreb in the course of seven years (Žirovčić Rajković, 2017).

Since the prosody of speech plays a key role in the development of verbal communication in children, this case report compares the prosodic features before and after the surgery. Prosody of speech is determined to a significant extent by temporal patterns and rhythm (Fletcher, 2010). A study conducted by Dilley et al. (2013) examined whether the manipulation of different speech parameters (prosodic and segmental) affects the perception of the speaker's age.

Just as the research by Logan et al. (2011), Redford (2014), and Kowal, O'Connell and Sabin (1975) showed, the prosodic competence of speakers (children) grows during speech-language development. The assessment of verbal communication competencies includes the assessment of prosodic factors that are an integral part of the phonetic assessment of typical and atypical speech. The speech tempo of younger speakers is expected to be greater than the speech tempo of older speakers (slower speech pace) (Linville, 2001). The temporal characteristics of speech are investigated more often in older people due to the decline in speech tempo. Huszár and Krepsz (2021) pointed out that there is not much information about this in relation to younger people.

Assessment of prosodic factors, such as F_0 value, fluency, and speech tempo in atypical speakers is of particular importance when it comes to diagnostics and, of course, therapeutic procedures. When analysing pathological voices, in addition to measuring time patterns (features), temporal features are also important (Škarić, 2007: see for columns, Linville, 1996).

AIM OF THE STUDY

The purpose of this study was to investigate the acoustic parameters of voice and voice quality before and after vocal fold surgery. The aim of this paper was to point to the importance of early diagnostics and surgery among patients with laryngeal web in order to achieve normal breathing. In addition, we stress the need for an interdisciplinary approach to observing the quality of voice and prosodic features at an early age. Interdisciplinarity in this case provides a holistic approach to the voice treatment and therapy, which ensures a good physiological base for proper (healthy) breathing and voice production. Through phonetic and speech therapy exercises, the patient mastered proper breathing, phonation, articulation, and voice resonance. In this way, the patient is ensured a high quality of life and optimal communication.

METHODS

Study participant

The subject under consideration was a six-year-old girl who had previously been diagnosed with irregular breathing (R06). After being admitted at the University Hospital Centre *Sestre milosrdnice,* an endoscopic exam revealed a laryngeal web between the vocal folds and an orderly intercartilaginous section of the laryngeal web. Surgical removal of the malformation was performed before the girl started school. Before the surgery, she was recorded in the acoustic studio.

The patient was taken to a doctor (otolaryngologist, phoniatrician) due to difficulties with the quality of her voice (hoarseness, quieter speech), vocal fatigue, breathing difficulties, and audible breaths, but the diagnosis of laryngeal web was not established. Therefore, she was referred to a speech therapist[1] for voice and pronunciation therapy. However, vocal therapy did not solve the problems with her voice and pronunciation. Fortunately, the parents insisted on further examination because they wanted the child's voice and breathing to improve before starting first grade of elementary school. During an additional second specialist examination, this time at the University Hospital Centre Sestre Milosrdnice in Zagreb, she was diagnosed with laryngeal web.

The girl, who was already diagnosed with laryngeal web, was recorded for the first time in the studio for acoustic recording at the Department of Phonetics (Faculty of Humanities and Social Sciences in Zagreb) in April 2021 before the LAVA vocal cord surgery and six months after the vocal cord surgery, i.e., in October 2021: the purpose of these recordings was to monitor voice quality, therapeutic voice protection, speech communication, expression, and longitudinal monitoring of voice quality improvement. The parents were given instructions that the girl should rest her voice for two weeks after the procedure, which is a routine procedure (e.g., Yang et al., 2014). They were also told that after that period, she should use her voice carefully and ensure that she follows the instructions provided by speech therapists and phoneticians. After each audio recording, the parents received an expert opinion from a phonetician about the recorded speech sample and the child's speech in general.

Surgical procedure

Multiple examinations (method: laryngomicroscopy, high-speed laryngoscopy) performed by a phoniatrician is considered to be a common diagnostic procedure after the surgical resection of the laryngeal web with a CO2 laser (LAVA[2]). It is performed multiple (2) times postoperatively at different intervals (the day after the surgery, one month after surgery, 6 months after the surgery, and one year after the surgery) for monitoring purposes, but also due to a possible recurrence of the web. In this case, surgery was necessary in order to create an unobstructed passage for the air flow and to form the vocal cords from the laryngeal web.

[1] The authors of the article do not have the information about frequency of speech therapy. The therapy was provided by another speech and language therapist in the place where the patient has been living.

[2] The LAVA laser surgical technique using CO2 was described by Degnello et al. (2010), and at that time, it was considered a new surgical technique for web removal by CO2.

Data collection

Endoscopy

Three endoscopic recordings of the vocal cords were conducted: the first recording in October 2020 (Figure 1), the second one in April 2021 (Figure 2), and the third one in May 2022 (Figure 3).

Figure 1. *Preoperative endoscopic image of the child's vocal cords showing the laryngeal anterior web Type II (October 2020)*

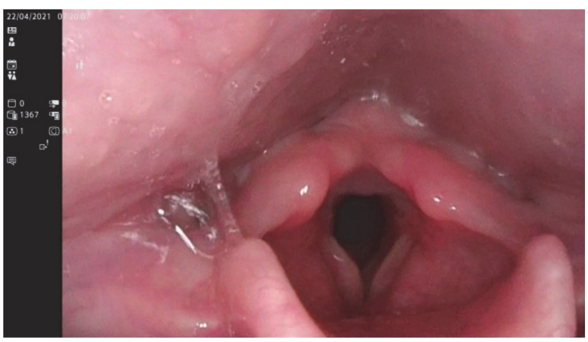


Figure 2. *Endoscopic image of the child's vocal cords taken on the day after the surgery (April 2021)*

Speech recording

The child was recorded in an acoustic recording studio at the Department of Phonetics at the Faculty of Humanities and Social Sciences in Zagreb. The studio is furnished with high quality acoustic equipment (AKG microphone - C414B-ULS; sound card: Fireface UFX; sampling rate: 44.1 kHz, 16 bits). The first recording was conducted in April 2021 (preoperative), the second in October 2021, and the third in May 2022. All recordings were conducted in almost the same methodological way and included two parts. Prior to the recording process, the child's parents underwent the protocol of filling out a demographic questionnaire, signing an informed consent form for the recording of their child and the use of the recorded data for research purposes, and for the purposes of voice quality therapy.

The recording process began in the form of an interview between a phonetician and the child, where the child was asked questions in order to prompt conversation (to present herself, to describe her family, pets, hobbies, (pre)school, and so on). This process enabled the gathering of approximately ten minutes of spontaneous speech. At the very end of the recording, the child phonated vowels in the Croatian language. It was not possible to record the phonation of all vowels five times in the first recording session due to voice fatigue. Additionally, during the first session, it was noticed that due to frequent stridor, the child was not able to prolong the phonation of vowels, other than the vowel [a] (recorded four times), for more than one second. In the first and second postoper-

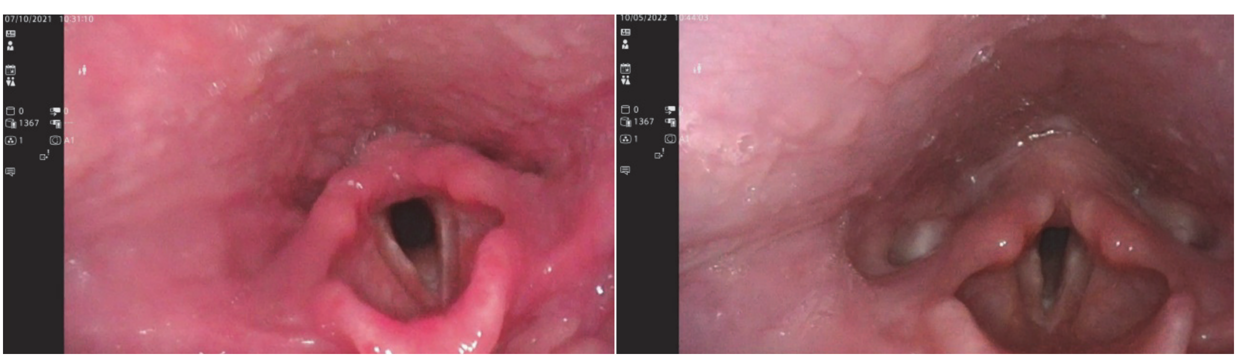


Figure 3. *Comparison of postoperative endoscopic images of the child's vocal cords (left: 6 months after surgery - taken in October 2021; right: one year after surgery - taken in May 2022)*

ative recording sessions (October 2021 and May 2022), the child phonated each of the vowels three times (in accordance with fundamental frequency measurements in acoustic phonetic research).

The total duration of the original voice recordings collected in three recording sessions (one preoperative and two postoperative sessions) was 56 minutes and 49 seconds. Unnecessary and distracting noise (the speech of the phonetician who conducted the interview, the sound of the child tapping her fingers on the table, and so on) were removed from the original recordings using software for acoustic processing and sound analysis (*Praat* and *Cool Edit Pro*). After editing the recordings and separating the part of the recordings with vowel phonations, spontaneous speech lasted 7.35 minutes preoperatively, 5.3 minutes in the first postoperative recording session, and 3.52 minutes in the second postoperative recording.

Voice assessment

We followed the perceptive descriptive phonetic protocol (Laver, 1980) for each recording (N = 3). The assessment of the child's voice quality included a perceptive phonetic evaluation carried out by two expert phoneticians using the phonetic descriptive protocol and the clinical Vocal Profile Analysis Protocol (VPA protocol) (Laver, 1980; 1996). According to the VPA protocol, timbre is described through the following parameters: vocal tract features, overall muscular tension, and phonation features. In the shorter version of the descriptive protocol, pitch has been described using a 1-5 scale, in which 1 corresponds to a very deep voice, number 2 to a deep voice, number 3 to a medium high voice, number 4 to a high voice, and number 5 to a very high voice. Speaking loudness has been described using a 1-3 scale (very quiet, quiet, and loud voice)[3].

[3] The method of subjective assessment using the VHI index (voice handicap index VHI-10 questionnaire) was not applied here, because it is a method of subjective self-assessment of the voice for adult patients that is carried out before surgery and 6 months postoperatively (e.g., it was also used by Yang et al. (2022) in a paper describing the treatment of acquired anterior laryngeal web in 32 patients with an average age of 38.2 years, within an age range from 11 to 80 years).

The parents also provided information on the child's speech difficulties, which was the reason why they sought the help of speech therapists and medical experts. In addition to the speech therapy (exercises) that the child attends at her place of residence, the phonetician shared advice on phonetic therapeutic exercises and exercises to warm up the voice and improve pronunciation (Varošanec-Škarić, 2010, Varošanec-Škarić et al., 2021).

Data analysis

Acoustical analysis

Long-term average speech spectrum (LTASS) analysis (*Praat*) was performed on the basis of spontaneous speech for all three recordings (one preoperative and two postoperative recordings) in order to compare the timbre of the voice.

Measures of fundamental frequency (F_0, minimum and maximum F_0, alternative baseline - Alt Fb), as well as baseline (Fb, F_0 range) and time organisation (duration in seconds) were calculated based on vowel phonation in the Praat script (Lindh, 2005) in Hz units and in semitones (ST). Alt Fb and Fb measures are commonly used in forensic phonetics because of their robustness (Varošanec-Škarić, 2019). Lindh and Eriksson (2007) suggested the use of these two long-time measures of fundamental frequency, since they have proven to be independent of speaking style, vocal effort, and recording quality. Bašić and Biočina (2020) mentioned that these measures are also minimally affected by acoustical recording conditions and speaker intra-variability (than other measures that are often used in acoustical research).

Furthermore, measures such as local jitter (%), shimmer (dB), and harmonic to noise ratio (HNR, dB) were also calculated. These measures are commonly used in clinical and scientific research of atypical voices. Based on these calculations, we estimated the variability of the fundamental frequency tone, its amplitude, and amount of additional noise in the voice, since laryngeal web prevents the production of a healthy (typical) voice.

Temporal and acoustic prosodic parameters

All analysed parameters were compared before and after the surgery (comparison of values in all three recordings) based on the child's spontaneous speech. For temporal prosodic factors, speech rate values were calculated based on the speech tempo parameters: speech rate - number of syllables per second (syl/sec)) and articulation tempo (TA, articulation rate) - number of syllables per second (analysed using the Praat programme; Boersma & Weenink, 2022). The differences in the values of speech and articulation tempo indicate the time spent on pauses (fillers or filled pauses, hesitations (duration in seconds), and frequency of pauses). The duration of speech units was also analysed, and descriptive statistical measures were calculated using all analysed measures: average values, median, minimum, maximum values, and standard deviation.

Acoustic prosodic parameters

In the category of acoustic prosodic parameters, various measures of fundamental frequency were analysed through the phonations of vowels: average, median, minimum and maximum values, as well as measures of alternative and basic fundamental frequencies (AltFb and Fb) in hertz and semitones. Standard deviations of F_0, jitter, shimmer, HNR, and phonation time were also measured. Acoustic analysis of sound parameters was performed using Praat 6.2.14 (Boersma & Weenink, 2022).

A fundamental frequency analysis was also performed based on the three recordings of spontaneous speech (preoperatively and postoperatively). Average values, median, minimum and maximum F_0 values, alternative and basic F_0 values, as well as duration and standard deviation were calculated. Measures of descriptive statistics and t-tests were used in the statistical analysis performed to examine significant differences in the preoperative and postoperative recordings with respect to the acoustic prosodic parameters.

RESULTS AND DISCUSSION

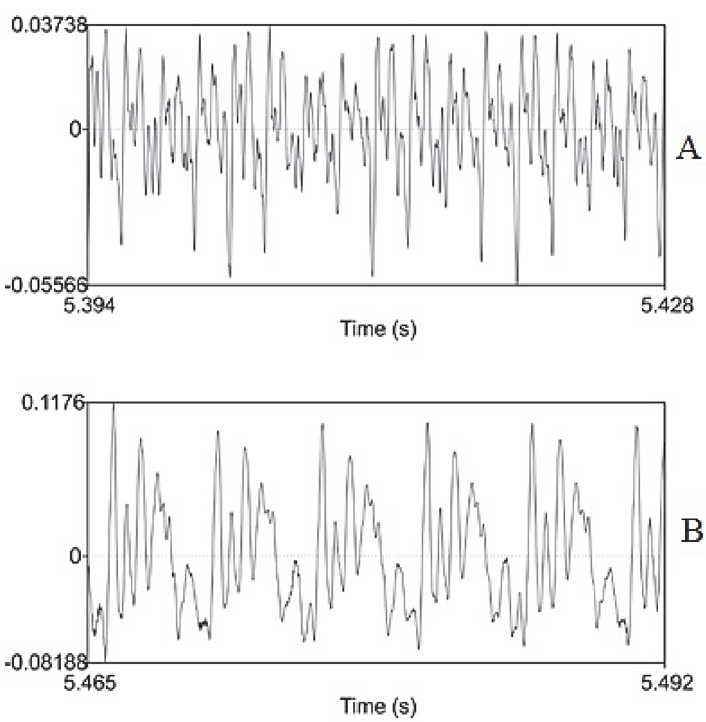
Comparison of auditory perceptual description of voice quality and prosodic parameters

Following the VPA protocol, before surgery, the child's timbre can be described according to the vocal tract features as follows: severe degree of supralaryngeal tension, severe degree of laryngeal tension, pathological squeaky whispering phonation (Figure 4A, based on oscillogram before surgery). In addition, according to the assessment of prosodic forms, the timbre can be described as high-pitched, very quiet voice, aperiodic, very shaky, with glottal fry, and a whispering voice present both in phonation and spontaneous speech. Due to frequent inhalations and aperiodic whispering phonation type, the pronunciation was also weaker, so in terms of auditory impression, her speech was insufficiently intelligible. According to the temporal organisation, her voice had on average faster speech tempo and tempo of articulation (compared to a group of children up to seven years old). The rhythm of speech was interrupted due to audible frequent breaths. The parameters of speech tempo and articulation tempo before the surgery were adequate: the average value of speech tempo in spontaneous speech was 3.31 syl/sec, while the tempo of articulation was 4.56 syl/sec, which is above the average values for children aged 5.9 to 7.9 years, which, according to Redford (2014), is 3.72 syl/sec for articulation tempo. Promising results in the other parameters can be explained by the fact that the child was being highly communicative.

After the surgery, the average values of speech tempo and articulation tempo were 3.26 syl/sec and 4.31 syl/sec, respectively. The assessment and the values of the temporal measures confirmed that the girl was very advanced for her age with respect to speech, despite the fact that she had difficulties breathing before the surgery. Preoperatively, shorter intonation patterns were separated by very short audible noise breaths (Figure 5), which can be detected in the image of the narrow-band spectrogram analysis and the accompanying oscillogram display (Figure 4B). The inspiratory audible stridor was very brief (about 1.3-1.5 ms; Figure 6), because the girl was focused only on the communication context. As she interestingly put it in a recoded interview one year after the surgery: “And now that I’ve grown a little, I was surprised before the surgery that I could hear myself normally, and other people couldn’t.”

Six months after the surgery, the voice was assessed as closer to a normal child’s voice, medium-high and medium-loud, without the presence of tension at the supralaryngeal level, but still with slight hoarseness, slight tension at the laryngeal level. Modal phonation is also present, in the part of phonation with a subtype of modal voice - a slight degree of roughness (Figure 4 - the oscillogram below), the voice is medium firm, with an appropriate speech pace and rhythm, without breathing disturbances, with fewer audible breaths and without stridor. Postoperatively, one year later, the voice was brighter in spontaneous speech, as is normal for children; there were no difficulties with breathing, and the pronunciation of vowels and consonants was excellent, with speech completely intelligible and comprehensible.

Intonational patterns in spontaneous speech are more connected in both postoperative recordings, with significantly smaller noise components: they are more harmonic, resonant, do not contain audible high-frequency stridors, which is evident in the images of the spectrogram analysis (Figure 7). For approximately the same duration (slightly more than 6 sec), four audible high-frequency stridors (Figure 5 and Figure 6) and several other breathing pauses can be detected on the images taken before the surgery. After the surgery, there are no stridors and normal breathing is observed. So, in addition to improved voice quality, the dynamic features of speech also improved postoperatively, including those characterising normal speech - appropriate tempo and rhythm, and connected intonation units of several connected spoken phrases. The average values of speech tempo and articulation rate before and after the surgery were similar, with insignificantly lower values recorded after the surgery. This can be interpreted as longer intonation patterns with normally distributed delimitation and processing pauses, including sound fillers that occur naturally in the sponta-



Before surgery

Mean F_0 (Hz)	230,81
S D F_0 (Hz)	169,48
Min F_0 (Hz)	73,63
Maks F_0 (Hz)	473,18
Jitter (%)	5,32%
Shimmer (dB)	1,556
HNR (dB)	7,09

After surgery

Mean F_0 (Hz)	207,33
S D F_0 (Hz)	1,56
Min F_0 (Hz)	203,19
Maks F_0 (Hz)	210,78
Jitter (%)	0,40%
Shimmer (dB)	0,438
HNR (dB)	17,94

Figure 4. *Comparison of oscillograms from the middle section of the most stable part of the phonation of the vowel [a] within 3 seconds: (A) preoperatively, and (B) six months postoperatively*

neous speech of speakers with normal voice quality. We observed that the little girl lengthened the vowels with conjunctive and prepositional functions at the points of demarcation of intonation units, such as the conjunction /i/ (and) - which is realised as [iː] (Figure 7). This was possible because the girl's speech was not interrupted by stridors and could flow naturally, since there were significantly fewer breathing pauses postoperatively than preoperatively (Figure 8 and Figure 9).

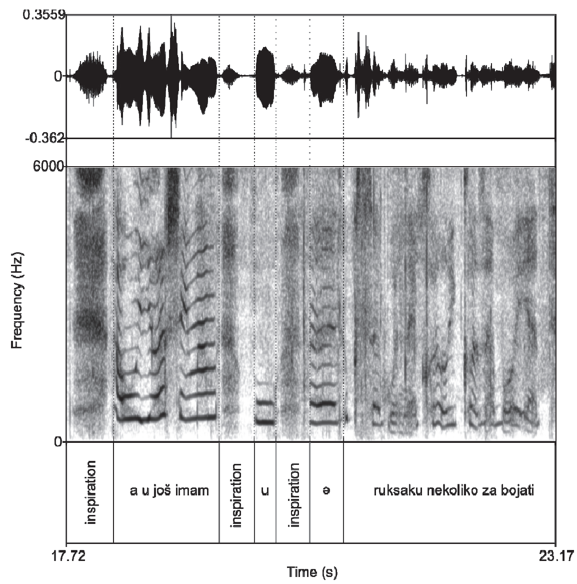


Figure 5. *Narrow-band spectrographic analysis of a speech segment enclosed by audible breaths and initial stridor (before the surgery)*

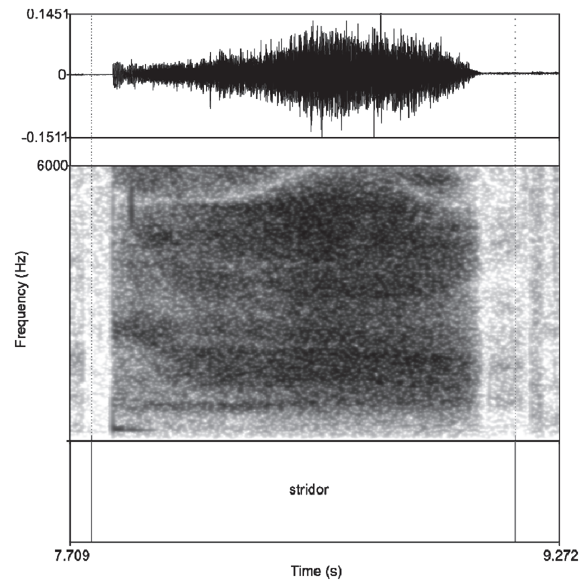


Figure 6. *Narrow-band spectrographic analysis of high-frequency audible stridor (preoperatively) during spontaneous speech*

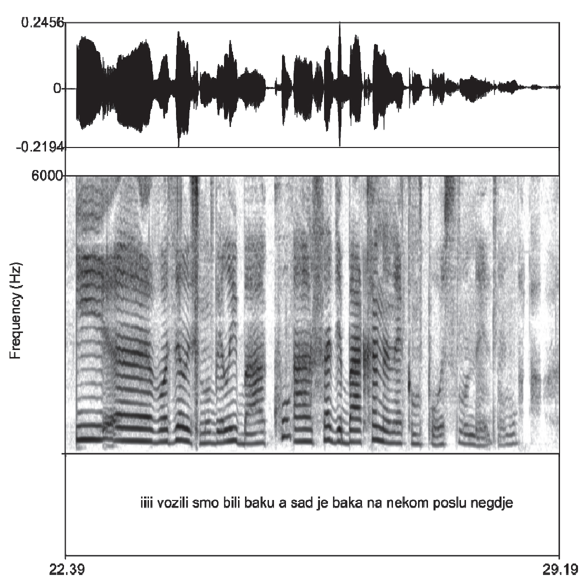


Figure 7. *Narrow-band spectrographic analysis of a clip containing spontaneous speech (6 months postoperatively)*

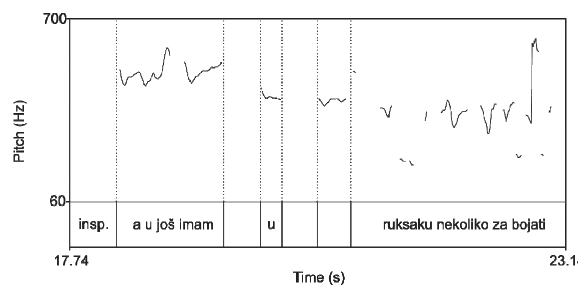


Figure 8. *Pitch contour (in Hz) during the interrupted intonation pattern in spontaneous speech (preoperatively)*

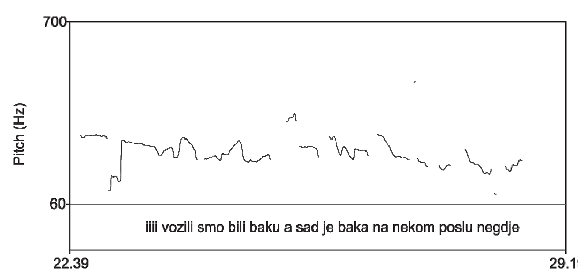


Figure 9. *Pitch contour in the connected intonation pattern of spontaneous speech (6 months postoperatively)*

Acoustic values before and after the procedure

Spontaneous speech

The overall permanent voice quality postoperatively showed a more balanced spectrum of the voice in longer speech (Figure 10) in its acoustic correlate. Preoperatively, it was noticeable that the average amplitude of F_0 in the voluminousness area (from the beginning of the spectrum to 400 Hz) was very weak, even 20 dB weaker than the strongest average amplitude of around 628 Hz and weaker than the area that also includes F2, which is a characteristic of a non-periodic squeaky noise voice. The range of stridency was amplified, peaking at 5 kHz. The spectrum was strongest in the F_0 area, as expected for the values of a normal child's voice, (i.e., above 250 Hz). Then it reduces slightly in the central area, while being more balanced, without a sudden increase in intensity around 5 kHz, in the area of brilliance and stridency. To maintain balance of the long-term average spectrum of the voice in speech, it is important that the higher resonant areas of the spectrum are not stronger in intensity than the lower parts, and that the F_0 area is the strongest and that the spectrum reduces slightly after it, so that it does not decrease too much in the area of brilliance and stridency. The difference was most evident in the F_0 area (Figure 10 and Figure 11). Average values for F_0 were indirectly obtained based on all vowels in stressed positions in spontaneous speech in the first and second postoperative speech samples (Table 1), which are approximations, i.e., 260.11 Hz (after 6 months) and 251.66 Hz (after a year). Yet the timbral quality of the voice was at its best one year after the surgery, as the voice was brighter, which is also reflected in the stronger spectral energy in the middle part of the spectrum up to 2 kHz (Figure 11). In the area of sonority in the part around 2 kHz, the spectrum was stronger by about 7 dB compared to the spectrum observed six months after the surgery, which is appropriate for a brighter child's voice.

It can be concluded that, one year after the surgery, the timbre was considered to be balanced for a child's voice (Figure 11).

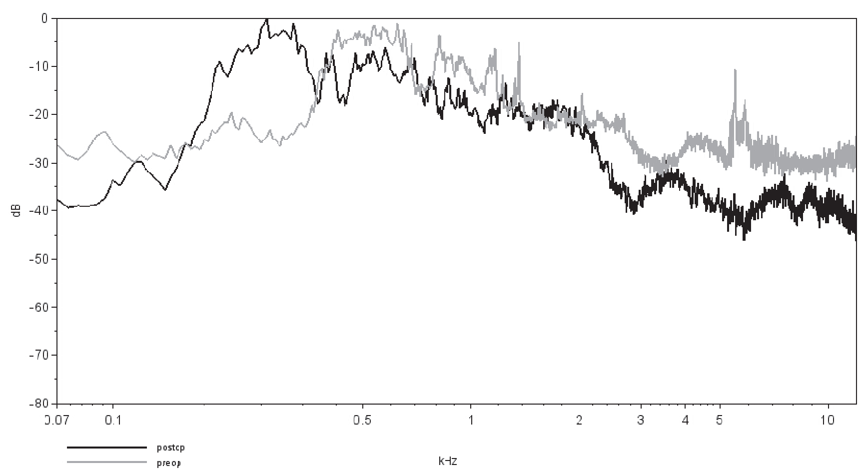


Figure 10. *Comparison of spectral curves of long-term average spectra (LTASS) based on spontaneous speech preoperatively (grey) and six months postoperatively (black)*

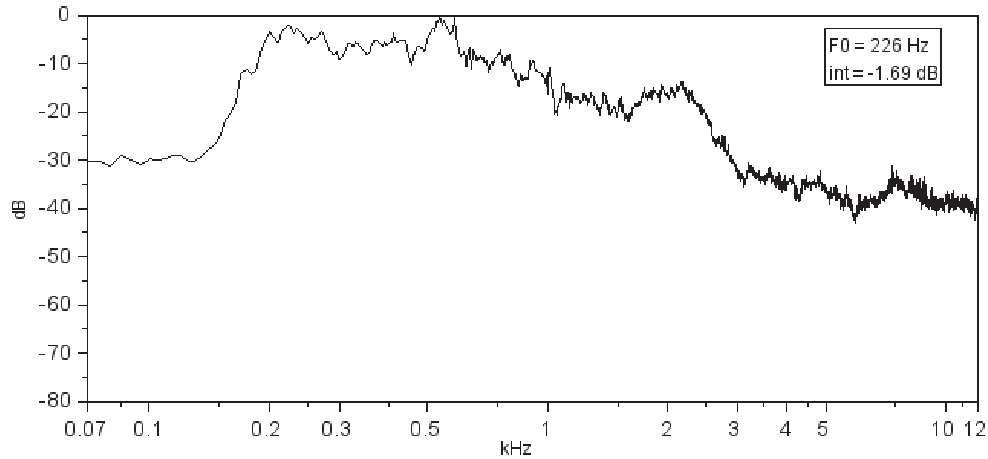


Figure 11. *LTASS based on spontaneous speech one year postoperatively*

Table 1. *Results of the acoustic analysis of fundamental frequency parameters measured preoperatively (1) and postoperatively (2 and 3) in spontaneous speech (in hertz and semitones)*

				Hertz			
	Mean F_0	**Median F_0**	**SD F_0**	**Min F_0**	**Max F_0**	**Alt_Fb**	**Fb**
Preop. rec. (1)	429.57	442.59	132.2	100.85	872.09	195.29	240.52
Postop. rec. (2)	260.11	251.95	61.37	100.45	872.14	195.73	172.35
Postop. rec. (3)	251.66	237.7	67.41	76.91	595.58	182.84	155.27
				Semitones			
	Mean F_0	**Median F_0**	**SD F_0**	**Min F_0**	**Max F_0**	**Alt_Fb**	**Fb**
Preop. rec. (1)	103.9	105.48	6.55	79.87	117.22	91.31	94.54
Postop. rec. (2)	95.84	95.72	3.85	79.8	117.22	91.35	90.34
Postop. rec. (3)	95.11	94.72	4.57	75.18	110.62	90.17	88.58

Vowel phonations

The average F_0 for the vowel [a] based on phonation was very high (442 Hz) preoperatively. Similarly, standard deviation F_0 was high (51.95 Hz) and pathological values were established by the following measures: local jitter (1.68%), local shimmer (0.7 dB), and HNR (17.6 dB; Table 1, Preop. rec. (1)). After the surgery, there was a significant decrease in the mean F_0 for [a] in comparison with the value of the preoperative one (220.5 Hz; $p < 0.001$), as well as in the maximum and minimum F_0 values ($p < 0.001$), and the standard deviation F_0 ($p = 0.01$). Furthermore, there was a statistically significant decrease in the values for local jitter ($p = 0.01$) and shimmer ($p = 0.02$) decreased significantly (Table 2, Postop. rec. (2)).

After the surgery, the local jitter for all vowels was within the normal range (0.30-0.47%). Similarly, the HNR was 22.46 dB at six months after surgery and 19.79 dB a year after surgery (for all vowels) (Table 2, Postop rec. 2 & 3). The span for the vowel [a] decreased significantly six months after the surgery. When the voice stabilised after a year, the span was in the expected range, with the difference between the maximum and minimum values being about one octave. The phonation time for all vowels was significantly longer ($p < 0.001$) one year after surgery than after six months, indicating that the voice stabilised over time (Tables 3a and 3b). Although a very short time had passed, the possibility that the phonation time was longer was lower, because of the increase in vital capacity due to the child's growth that occurred in the meantime.

Table 2. *Results of the acoustic analysis of fundamental frequency and duration preoperatively (1) and postoperatively (2 and 3) during vowel phonation*

		Mean F_0 (Hz)	SD F_0 (Hz)	Min F_0 (Hz)	Max F_0 (Hz)	Jitter (%)	Shimmer (dB)	HNR (dB)	Phon. time (sec)
PREOP. REC. (1)	**[a]**	442.33	51.95	378.83	520.29	1.68	0.70	17.58	15.28
POSTOP. REC. (2)	**[a]**	220.55	2.45	213.83	225.98	0.41	0.64	18.24	10.30
	[e]	196.84	12.82	179.68	213.44	0.48	0.35	18.51	11.02
	[i]	207.81	3.36	201.97	216.24	0.55	0.24	22.80	10.38
	[o]	211.12	19.35	196.25	269.43	0.44	0.32	24.23	10.83
	[u]	217.70	3.66	210.12	226.12	0.30	0.21	28.43	11.75
	All vowels	210.85	8.22	200.50	229.98	0.45	0.36	22.46	10.90
POSTOP. REC. (3)	**[a]**	141.33	42.03	80.01	196.99	0.92	0.70	12.19	15.03
	[e]	194.40	5.33	183.91	204.45	0.58	0.48	19.89	16.19
	[i]	185.90	3.55	177.39	193.98	0.49	0.16	20.35	14.44
	[o]	182.84	4.43	175.24	196.70	0.45	0.26	21.89	14.44
	[u]	190.14	7.19	173.29	213.09	0.50	0.34	24.62	16.39
	All vowels	178.92	12.51	157.97	201.04	0.59	0.39	19.79	15.30

The average and median F_0 values in spontaneous speech were also lower postoperatively. In particular, the high preoperative F_0 and tone dispersal median values of 442.59 Hz and 132 Hz, respectively, were notably lower (one year after surgery, median F_0 = 237.7 Hz; SD F_0 = 67.4 Hz). The average acoustic values of the fundamental frequency remained stable one year after the surgery compared to the values based on the first postoperative recording (six months after the surgery), but they were not significantly different.

Results of the t-test for acoustic variables based on preoperative and postoperative phonation

Table 3a. *T-test results for the preoperative recording and the 1st postoperative recording during vowel phonation*

	Mean F_0	SD F_0	Min F_0	Max F_0	Jitter	Shimmer	HNR	duration
t	8.63	2.65	4.61	10.37	2.80	2.47	2.04	4.30
p	0.00	0.01	0.00	0.00	0.01	0.02	0.05	0.00

Table 3b. *T-test results for the 1st and the 2nd postoperative recording*

	Mean F_0	SD F_0	Min F_0	Max F_0	Jitter	Shimmer	HNR	duration
t	2.09	0.43	1.58	2.12	1.38	0.73	1.88	6.75
p	0.07	0.68	0.15	0.06	0.20	0.48	0.09	0.00

These results show that the surgery was necessary in order to provide the child with the necessary physiological foundations for the production of a healthy voice. The success of the surgery was confirmed by the phonetic evaluation of her voice (in all analysed parameters). Furthermore, speech therapy and treatment administered by the phonetician helped improve voice production, resulting in a smooth and healthy voice.

CONCLUSION

Before the surgery, audible breaths (stridors) were often present along with very difficult breathing, severe laryngeal tension, and pathological squeaky whispering phonation. Before the procedure, the average fundamental frequency was very high (442 Hz) and the pathological values were determined based on the following measures: tone aperiodicity (1.67%), tone amplitude aperiodicity (0.7 dB), and the HNR in the voice (17.6 dB).

The surgery helped improve the larger vibrating part of the vocal fold and widen the air flow passage (between vocal folds). Therefore, the child's vocal folds could vibrate along their entire length, and the wider passage between the vocal folds could allow air to pass freely. After the surgery, the pitch was about an octave lower (220.5 Hz), closer to the target value for the girl's age group. Improvements are expected in the next recording, considering the narrow range of F_0 that has been observed so far. The measure for local F_0 aperiodicity is no longer pathological - it has the values of normal healthy voices in all five vowels (from 0.30-0.47%). Postoperatively, when we considered all vowels, the HNR measure showed values typical for a normal healthy voice (22.46 dB). Statistically significant postoperative improvement was revealed in fundamental frequency values, F_0 range, minimum and maximum F_0 values, frequency and amplitude aperiodicity of F_0, as well as in the values of the HNR in the voice.

The overall quality of the voice postoperatively showed a more balanced timbre on the images of the long-term average speech spectra. At one year after surgery, the voice was brighter and normal for children of a similar age group. Other dynamic features, which characterise normal speech, also showed improvement, and intonation units were longer and more connected. In addition, there were significantly fewer breathing pauses in the child's speech. The results of the present study could be applied to other papers and research on a similar topic. In fact, the application of phonetic voice assessment (perceptive and instrumental) can contribute to a more precise description and assessment of the voice before and after surgery and/or vocal therapy. The course and success of speech therapy can be monitored by comparing the values of the analysed parameters. Finally, this work aims to improve the early recognition and diagnosis of cases of laryngeal web in children in the work of paediatricians and otorhinolaryngologists.

Overall results after surgery indicate significant improvements in the child's speech in general. Breathing during speech was balanced and values of several F_0 measures were within the range for healthy voices. Therefore, we can conclude that surgery provided a better quality of life for the child and her parents.

REFERENCES

Bašić, I., & Biočina, Z. (2020). Metodologija akustičkih snimanja. *Zbornik radova Filozofskog fakulteta u Splitu, 13*, 19-40. https://doi.org/10.38003/zrffs.13.10.

Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by computer* (verzija 6.0.31), http://www.praat.org/.

Cohen, S. R. (1971). Congenital atresia of the larynx with endolaryngeal surgical correction: A case report. *Laryngoscope*, *81*, 1607-1616.

Cohen, S. R. (1985). Congenital glottic web in children. A retrospective review of 51 patients. *Annals of otology, rhinology and laryngology, supplement*, *121*, 2-16.

Dilley, L. C., Wieland, E. A., Gamache, J. L., McAuley, J. D., & Redford, M.A. (2012). Age-related changes to spectral voice characteristics affect judgments of prosodic, segmental, and talker attributes for child and adult speech. *Journal of Speech, Language and Hearing Research, 56*(1), 159-77. DOI: 10.1044/1092-4388(2012/11-0199).

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In: W. J. Hardacastle, J. Laver, and F. E. Gibbon (Eds.), *The Handbook of Phonetic Sciences* (pp. 521-602). Oxford: Willey-Blackwell.

Gay, I., Feinmesser, R., & Cohen, T. (1981). Laryngeal Web, Congenital Heart Disease, and Low stature: A Syndrome. *Arch Otolaryngol, 107,* 510-512.

Halle M. (1985). Speculations about the representation of words in memory. In: V. A. Fromkin (Ed). *Phonetic linguistics*. New York: Academic Press; 101-104.

Holland, B. W., Koufman, J. A., Postma, G. N., & McGuirt, W. F. (2002). Laryngopharyngeal Reflux and Laryngeal Web Formation in Patients With Pediatric Recurrent Respiratory Papillomas. *The Laryngoscope, 112,* 1926-1929.

Huszár, A., & Krepsz, V. (2021). The development of variability in pausing and articulation rate in Hungarian speakers ten years apart. *Govor*, *38*(2): 121-146. https://doi.org/10.22210/govor.2021.38.09.

Kowal, S., O'Connell, D., & Sabin, E. (1975). Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research*, *4*, 195-207. https://doi.org/10.1007/BF01066926.

Kuhl P. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50*(2): 93-107.

Laver, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Lawlor, C. M., Dombrowski, N. D., Nuss, R. C., Rahbar, R., & Choi, S. S. (2020). Laryngeal Web in the Pediatric Population: Evaluation and Management. *Otolaryngology – Head and Neck Surgery, 162*(2), 234–240. DOI: 10.1177/0194599819893985. http://otojournal.org (accessed 14 March 2022.)

Lindh, J. (2005-2006). *Praat script.* Last accessed on February 20th 2020. from https://sites.google.com/site/phoneticolindh/home/scripts-and-audio/scripts

Linville, S. E. (1996). The sound of senescence. *Journal of Voice*, *10*(2), 190-200.

Linville, S. E. (2001). *Vocal aging.* San Diego: Singular Thomson Learning.

Logan, K. J., Byrd, C. T., Mazzocchi, E. M., & Gillam, R. B. (2011). Speaking rate characteristics of elementary-school-aged children who do and do not stutter. *Journal of Communication Disorders*, *44*, 130–147. DOI: 10.1016/j.jcomdis.2010.08.001.

Prstačić, R., Penezić, A., Laksar Klarić, Ž, Maržić, D., & Vela Ljubić, J. (2020). Smjernice za promuklost: Guidelines for hoarseness. *Medica Jadertina*, *50*(3), 231-236.

Redford, M. A. (2014). The perceived clarity of children's speech varies as a function of their default articulation rate. *Journal of Acoustical Society of America*, *135*(5): 2952-2963. DOI: 10.1121/1.4869820. PMID: 24815275; PMCID: PMC4032421.

Stasney, R. C. (1995). Laryngeal Webs: A New Treatment for an Old Problem. *Journal of Voice*, *9*(1), 106-109.

Strakowski, S. M., Butler, M. G., Cheek, J. W., Moore, W. T., Netterville, J. L., & Phillips III, J. A. (1988). Familial laryngeal web in three generations with probable autosomal dominant transmission. *Dysmorphology And Clinical Genetics, 2*: 9-12.

Škarić, I. (2007). *Hrvatski izgovor.* Zagreb: Nakladni zavod Globus.

Tiwari, M., Fernandes, V.L.G., George, S., Sansgiri, V. B., & Khandolkar, P. (2018). Congenital Laryngeal Web: A Laryngology Rarity. *International Journal of Otolaryngology and Head & Neck Surgery, 7,* 143-147. https://doi.org/10.4236/ijohns.2018.74017.

Varošanec-Škarić, G. (2010). *Fonetska njega glasa i izgovora.* Zagreb: FF press.

Varošanec-Škarić, G. (2019). *Forenzična fonetika*. Zagreb: Ibis grafika.

Varošanec-Škarić, G., Stevanović, S., & Bašić, I. (2021). Comparative perceptual evaluation and acoustic voice analysis of a transgender client male to female before and after laser-assisted voice adjustment surgery. *Hrvatska revija za rehabilitacijska istraživanja*, *57*(1); 40-55. doi:10.31299/hrri.57.1.2.

Yang, X., Jun, W., Demin, H., Lijing, Y., & Wen, X. (2014). Vocal cord mucosal flap for the treatment of acquired anterior laryngeal web. *Chinese Medical Journal, 127*(1), 1294-397.

Žirovčić Rajković, S. (2017). Patološke promjene grkljana u dječjoj dobi. *Zbornik radova za medicinske sestre*, 123-129.