RETROSPECTIVE LONGITUDINAL ACOUSTIC AND PERCEPTIVE STUDY OF SUBSTITUTION VOICE AFTER PARTIAL LARYNGECTOMY

Lise Crevier Buchman[1&2], Emeline Roques[2], Nicolas Audibert[1]

[1] Phonetics and Phonology Lab (LPP), UMR7018 CNRS/Univ. Sorbonne Nouvelle Paris 3,
[2] Voice and Speech Lab, Head & Neck Department, Hospital HEGP, APHP, Univ. Paris 5
lise.buchman@numericable.fr, roques.emeline@yahoo.fr, nicolas.audibert@univ-paris3.fr

ABSTRACT

The goal of this longitudinal study in substitution voice after vertical and horizontal partial laryngectomy was i) to analyse perceptual and acoustic characteristics related to specific surgery, ii) to determine relevant acoustic measurements to classify voice quality. 30 male patients were recorded at 3, 6 and 12 months after surgery and 15 male controls. Two perceptual scales (GRB & IINFVo) were compared for relevance mean. Vowel [a, e, i, o, u] identification test was compared to acoustic analysis of the vocalic triangle to better understand the confusions. Long Term Average Spectrum (LTAS) could measure energy distribution in noisy voices. The first six months can be considered as an adaptation period for voicing features and improvement of breathy voice quality. Overlapping formant frequencies could explain the vocalic perceptual confusions. Voice and speech are better preserved with at least one vocal fold and the remaining of 2 arytenoids for neoglottic closure efficiency.

Keywords: partial laryngectomy, substitution voice, voice quality, vowels, perceptual evaluation.

1. INTRODUCTION

The partial laryngectomies require the conservation of at least one functional crico-arytenoid unit whose role is crucial for the maintenance of the respiratory tract and to allow a mobilization of these structures for the phonatory function [3]. Thus, various types of partial laryngectomies were developed according to the extent of the lesion to treat. Such laryngectomies consist of the removal of whole or part of the glottis (one or two vocal folds) of a variable part of the thyroid cartilage, and of a more or less great part of the supra glottal area (ventricular folds, arytenoid cartilage), sometimes extended to the neighbouring structures like the epiglottis and the root of the tongue. The partial laryngectomies are classified according to whether they are "vertical" with variable conservation of a hemilarynx like the frontolateral partial laryngectomies [1], or "horizontal" with removal of the glottis and the neighbouring structures. However all these partial laryngectomies with removal of at least one vocal fold will give a voice called "substitution voice". Substitution voice is voicing without two true vocal folds [9]. This modification of vocal quality can result from the modified laryngeal vibrator (called "neovibrator") and from the vocal tract, with acoustic and perceptual consequences depending on the type of surgery. Thus our first hypothesis is that the vertical partial laryngectomies, which respect the antero-posterior axis of the laryngeal vibrator, will have acoustic and perceptual parameters belonging to the voices known as "glottic". On the other hand, the horizontal partial laryngectomies, which modify the axis of the vibrator, will have acoustic and perceptual parameters belonging to the voices known as "laryngeal aglottic". A second assumption was the existence of acoustic and perceptual characteristics, which would make it possible to define the vocal consequences of a category of partial laryngectomy. More precisely, our objectives are both to define acoustic and perceptual indices relevant to follow vocal improvement after partial laryngectomy, and to identity distinctive indices between various surgeries.

To address these assumptions, two perceptual scales (GRB & IINFVo) [5,9] were compared for relevance. A perceptual vowel identification [a, e, i, o, u] test, using Praat [2] was compared to acoustic analysis of the vocalic triangle to better understand the confusions. The frequency distribution was measured and compared over time using Long Term Average Spectrum (LTAS) [6]. Taken together, those analyses outline an evolution profile in relationship with a set of relevant indices, beyond classical measures based on fundamental frequency (F0) and its variations caused by aperiodicity.

2. METHODOLOGY

2.1. Partial laryngectomy

We compared three techniques of external partial laryngectomies:

i) A vertical technique, frontolateral partial laryngectomy (FL), which removes a vocal fold, the

anterior commissure, and the anterior part of the controlateral vocal fold. The neoglottis remains in an antero-posterior orientation [1].

ii) Two horizontal techniques, the supracricoid partial laryngectomies reconstructed by cricohyoidoepiglottopexy (CHEP), or by cricohyoidopexy (CHP). In both techniques, the glottis (vocal folds and ventricular folds) is removed entirely, as well as the epiglottis for the CHP. The reconstruction is done by suturing the hyoid bone to the epiglottis for CHEP, and to the root of the tongue for CHP [3]. Moreover, for these two techniques, one arytenoid can be removed (hereafter CHEP1 and CHP1, cases with two arytenoids remaining being noted CHEP2 and CHP2). The neoglottis loses its anteroposterior orientation, voicing is then produced by an approximation between the two arytenoid cartilages on the back side, and either the root of the tongue (for the CHP) or the epiglottis and aryepiglottic folds (for the CHEP) on the front side. In these two cases the vocal tract is shortened by approximately 3 cm [4].

2.2. Population.

We recorded 10 male patients per type of laryngectomy (30 patients altogether) age ranging from 59 to 78 years and 15 male control speakers, aged 57 to 76 years to match the patients.

2.3. Recordings

All patients were recorded at 3, 6 and 12 months post-operative for a longitudinal follow-up of the perceptual and acoustic modifications of vocal quality. The equipment used was a one-way microphone Electret SHURE SM 48 placed 15 cm from the lips, connected to a Sony ES DTC DAT recorder covering the frequency range 0-48 kHz.

2.4. Perceptual evaluation

2.4.1 Dimensional voice quality ratings

We carried out a perceptual evaluation on a reading task by using the GRB scale [5] dedicated to voice quality evaluation and the IINFVo [9] scale specifically designed for substitution voices. Five naive listeners evaluated the voices at the three post-operative periods.

2.4.1 Vowel identification

We carried out a perceptual forced-choice identification test of the 5 vowels [a, e, i, o, u], produced by the 30 patients at the three post-operative times and by the 15 controls. Altogether, 504 vowels were evaluated by five native listeners, plus 50 vowels randomly selected and presented a second time for intra-judge reliability estimation, for a total duration of approximately 45 minutes.

2.5. Acoustic analysis

2.5.1. Vocalic triangles

Formant frequencies of vowels were extracted using Praat [2], for each patient at each post-operative time and for control speakers. Extracted frequencies were then plotted in the F1/F2 space to allow a qualitative comparison between the acoustic distinctiveness of vowels and confusions obtained in the listening test.

2.5.2. Long Term Average Spectrum (LTAS)

For each speaker and post-operative time, the LTAS was computed on 100 Hz bands from at least 40 seconds of read text, after energy normalization. This measure of spectral energy distribution is used to characterize the noise component frequently observed in such substitution voices.

2.6. Statistics

The evolution over 12 months post-operatively of the various parameters for each type of surgery, was tested by ANOVA with 2 factors (type of surgery and post-operative time). Pairwise comparisons are performed using Tukey's HSD.

3. RESULTS

3.1. Voice quality ratings

3.1.1 GRB scale

- The perceptual feature "G" reflects the general impression on the voice. We obtained a significant effect of the type of surgery on this feature ($F(4, 835) = 226.74$, $p<0.001$) and not according to the post-op period. The voices after FL are altered ($p<0.001$). The ablation of one arytenoid alters the global quality of the voice for the CHEP ($p<0.05$) and CHP ($p<0.001$).

- The feature "R" reflects the rough character of the voice. A significant difference exists between the various types of surgery ($F(4,835) = 227.10$, $p<0.001$), FL is the least rough ($p<0.001$). CHEP2 are rougher than the CHEP1 ($p<0.001$) and the CHP is the roughest substitution voice ($p<0.001$). Only CHEP2 gets worse between 3 and 6 months ($p<0.001$).

- The feature "B" reflects the presence of breath in the voice and varies statistically according to the type of surgery ($F(4,835) = 37,79$, $p<0.001$). FL has

a breathy voice over the one-year study. CHEP1 and CHP1 ($p<0.001$) are breathier than FL and than the CHP2 and CHEP2 ($p<0.001$). However, we noticed an improvement (decrease) of breathiness between 3-6 month for the CHP2 ($p<0.001$), CHP1 ($p<0.05$) and CHEP2 ($p<0.001$) that was prolonged to a lesser extent between 6-12 months ($p<0.05$) for CHEP1.

3.1.2. IINFVo scale

- The feature (Ia) "Overall Impression" reflects the general impression of the voice and varies according to the type of surgery, FL being less degraded than the CHEP and CHP ($p<0.001$). The CHEP1 and the CHP1 are worse than the CHEP2 and CHP2 ($p<0.05$ and $p<0.001$ respectively).

- The feature (Ib) "Impression of Intelligibility" characterizes the comprehensibility of the voice. The voices after FL are less altered than the 2 other laryngectomies ($p<0.001$) except the CHEP2. For each type of horizontal surgery, the comprehensibility is better when 2 arytenoids are preserved, for CHEP ($p<0.01$) and CHP ($p<0.001$). We noticed an improvement in time ($F(2,835) = 6.76$; $p<0.01$) for all the surgeries, especially between 6 and 12 months ($p<0.05$).

- The feature (N) "Noise" reflects the presence of noise added in the voice signal. It varies to a significant degree according to the type of surgery ($F(4,835) = 90.80$; $p<0.001$), FL being less deteriorated, and in addition according to post-op period ($F(2,835) = 13.32$; $p<0.001$) between 3-6 months ($p<0.001$) and 3-12 months ($p<0.05$). For the CHEP and CHP, there is more noise if one arytenoid only remains ($p<0.001$).

- The feature (F) "Fluency" reflects the speed and rhythm of speech. It varies according to the type of surgery ($F(4,835) = 53.92$; $p<0.001$). FL has a less modified fluency compared with the other laryngectomies ($p<0.001$). The CHP2 are less altered than the CHP1 ($p<0.001$). The improvement of the fluency is significant ($F(2,835) = 8.90$; $p<0.001$) between 3 and 6 months ($p<0.05$) for all the laryngectomies.

- The feature (V) "Voicing" reflects the quality of voicing. It varies to a significant degree according to the type of surgery ($F(4,835) = 22.74$; $p<0.01$), FL has a better quality of voicing ($p<0.001$) than CHEP1 and CHP1. The loss of an arytenoid alters voicing ($p<0.001$). The three surgeries improve ($F(2,835) = 34.20$; $p<0.001$) between 3-6 months and 6-12 months ($p<0.001$).

3.1.3. Reliability measurements

For both scales, intra and inter-judge reliability measured by Cronbach's alpha reach high values ($0.91 \leq \alpha \leq 0,97$ in all cases). This result indicates that judgements were as accurate and reproducible on the fine-grained scale IINFVo than on GRB.

3.2. Vowel identification

3.2.1. Perceptual analysis of vowels

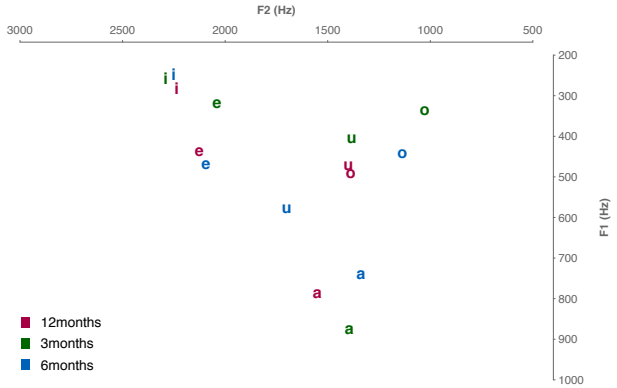
Since confusion matrices for each surgery and post-operative time cannot be reproduced here, only their most remarkable characteristics are reported in this section, focusing on confusions over chance level (20%).

For FL, we noticed a deterioration of the identification of the vowels over time except for [a]. [e] is confused with [i] (3 months 28%, 6 months 58%, 12 months 56%) and [o] with [u] (3 months 36%, 6 months 55%, 12 months 41%). For CHEP1 there is a relative stability in time and the confusions are between [e/i] and [o/u]. For CHEP2 we note a better distinctiveness over time. Conversely, for CHP1 confusions increase for [a] and [u]. Confusions persist between [e/i] and [o/u]. For CHP2 one notes an improvement of [o] and a degradation of [a] and [i] in the course of time.

In summary, major confusions were observed between pairs of vowels [e/i] and [o/u] for all surgeries, while the vowel [a] remains well identified.

3.2.2. Vocalic triangles

Figure 1: The vocalic triangle for FL at 3, 6, 12 months post-operative.



In order to better understand the confusions observed previously, we compared them with the acoustic analysis of the vocalic triangles. The partial laryngectomies modify the characteristics of the vocal tract and in particular the formants of certain vowels. However, whatever the type of surgery, the [a] vowel, in spite of a rise in the value of the first

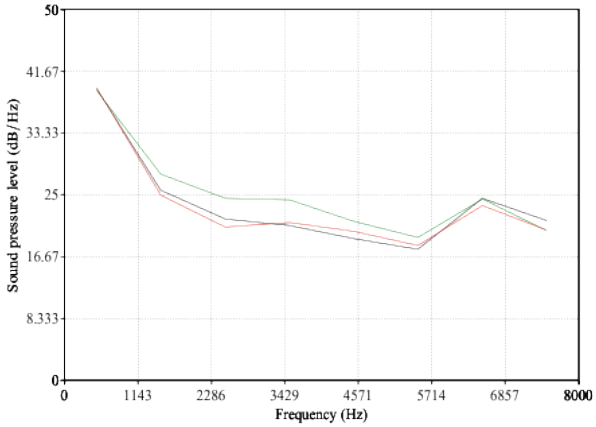
two formants (800 Hz and 1500Hz respectively in average) when compared with the values of our control group, remains well identified and does not vary in time. For FL, we note a degradation of the results during the 1st year except for [a]. At 12 months post op. [u] and [o] concentrate very close to each other (fig. 1). For CHEP1, [e] have formant values close to [i]. Vowels [i, o, u] have sub-normal formant values. For CHEP2, a convergence of formant values for [o] and [u] (with high F2) can be noticed from 6 months post-operative.

Overall, the examination of vocalic triangles indicates an improvement of vowels distinctiveness over time. However, for CHP1 we noted that vowels [a] and [u] were less well identified in the course of time. For CHP2, we note an improvement of [o] and a degradation of [i] whose formants approach those of [e]. In summary, we found the same trends as in the perceptual identification test. When confronting perceptual and acoustic results, we also observe that values of F2 and to a lesser extent of F1 increase in all surgeries compared to control speakers, without an impact on vowels identification.

3.4. Long Term Average Spectrum (LTAS)

The spectra (see fig. 2 for CHEP2) showed the presence of a peak of intensity between 5700 and 6800 HZ for all surgeries, at three post-operative period compared to those of the control group ($p<0.001$).

Figure 2: LTAS for CHEP2 at 3 months (green line), 6 months (black line) and 12 months (red line) post-op.



4. DISCUSSION & CONCLUSION

Our results lead to the following conclusions regarding our two assumptions.

Partial vertical laryngectomies FL respect the antero-posterior axis of the laryngeal vibrator, with its "V" shape in breathing and phonation [1]. As expected, the acoustic and perceptual parameters belong to the range of voices known as "glottal". The good quality of the voicing which improves during the first year, and keeps the characteristics close to a dysphonia with breathy voice but less severe and rough than in the other surgeries [3,4,8].

Concerning the identification of the vowels, we note a degradation of the results during the first year except for [a]. The progressive atrophy of the laryngeal scar could increase the neoglottic incompetence and consequently the breathy voice.

The voice after partial horizontal laryngectomies distinguishes between CHEP and CHP depending amongst arytenoid and epiglottis conservation for the sphincter phonatory function of the neolarynx. Surgeries in which the symmetry of the neoglottis (with two arytenoid cartilages) is preserved gives better vocal results in particular for the voicing feature, with a "T" shape closure compared with the "L" shape closure when only one arytenoid remains. The rough and breathy voice quality is more important in the CHP1 surgery associated to the least effective closure of the neoglottis. Roughness can be explained by the slow and irregular vibrations of the mucous membrane of the arytenoids [11].

The overall vocal results improve in the course of time without distinction between types of surgery. The breathiness reduces faster than roughness, around 6 months post-op. There is an improvement of quality of closure of the neoglottis but there is no real change in the vibratory pattern of the mucous membrane of the arytenoids [4]. Another specificity related to horizontal partial laryngectomies is the shortening of the vocal tract due to its reconstruction [4,10]. The acoustic consequence of this shortening is an increase in F2 for the vowels [a, o, u] [8].
Thus, with caution regarding the limited sample size, the results of an acoustic analysis can help to evaluate the vocal function and its evolution in time. The first six months after surgery would correspond to the establishment of the "new voice" with improvement of the breathiness and voicing, whereas the added noises, fluidity and comprehensibility improve only from 6 months post-operative. Concerning the LTAS our first results show the interest of this measurement, which enables to differentiate the substitution voices from the control voices.

ACKNOWLEDGMENTS

This work was supported by the French Investissements d'Avenir - Labex EFL program (ANR-10-LABX-0083) and by the European Union i-Treasures project (FP7-ICT-2011-9-600676-i-Treasures).

5. REFERENCES

[1] Biacabe, B., Crevier-Buchman, L., Laccourreye, O., Brasnu, B. 1998. Laryngectomie partielle verticale avec reconstruction glottique, résultats carcinologiques et fonctionnels. Annales d'Otolaryngologie de Chirurgie Cervicofaciale. 115, 189-95

[2] Boersma, P., Weenink, D. 2014. *Praat: doing phonetics by computer* [Computer program]. Version 5.4.04, retrieved 28 December 2014 from http://www.praat.org/

[3] Crevier-Buchman, L., Laccourreye, O., Weinstein, G., Garcia, D., Jouffre, V., Brasnu, D. 1995b Evaluation of speech and voice following supra-cricoïd partial laryngectomy. The journal of laryngology and otology,109, 410-413

[4] Crevier Buchman, L., Laccourreye, O., Wuyst, Fl., Monfrais-Pfauwadel, Mc., Pillot, C., Brasnu, D. 1998. Comparison and evolution of perceptual and acoustic characterisitics of voice after supracricoid partial laryngectomy with cricohyoidoepiglottopexy. Acta Otolaryngol. 118(4), 594-999.

[5] Hirano, M. 1989. Objective evaluation of the human voice clinical aspects. Folia phoniatrica. 41 : 89-144.

[6] Kitzing, P. 1986. LTAS Criteria pertinent to the measurement of voice quality, Journal of Phonetics, 14, 477-82

[7] Kreiman, J ; Gerratt, B.R ; Precoda, K. 1990. Listener Experience and Perception of Perceptual Evaluation of Voice Quality : Journal of Speech and Hearing Research, 33 :103-115

[8] Laccourreye, O. Crevier-Buchman, L., Weinstein, G., Biacabe, B., Laccourreye, H., Brasnu, D. 1995. Duration and frequency characteristics of speech and voice following supracricoïd laryngectomy.The Annals of Otolaryngology, Rhinology and Laryngology. 104 (7), 516-521

[9] Moerman, M., Martens, Jp., Crevier Buchman, L., Woisard, V., Dejonckere, P. 2005. Perceptive evaluation of substitution voices : the I(I)NFVo rating scale. Revue de Laryngologie Otologie Rhinologie. 126(5), 323-325

[10] Schlinder, A., Favero, E., Nudi, S., Spadola- Bisetti, M., Ottaviani, F., Schlinder, O. 2005.Voice after supracricoïd laryngectomy : subjective, objective an self assessment data. Logopedics phoniatrics Vocology, 30, 114-119.

[11] Yuce, I., Cagli, S ., Bayram, A., Karasu, F., Sati, I., Guney, E. 2009. The effect of arytenoid resection on functional results of cricohyoidopexy. Otolaryngology head and Neck Surgery, 141, 272-275.