AN ACOUSTIC STUDY OF SUSTAINED VOWELS PRODUCED BY PATIENTS AFTER THYROID SURGERY

Camille FAUTH[1], Béatrice VAXELAIRE[1], Jean-François RODIER[2], Pierre Philippe VOLKMAR[2], Rudolph SOCK[1&3]
[1]LiLPa & IPS – University of Strasbourg France
[2]Clinique Sainte Anne – Strasbourg France
[3]Université Pavol Jozef Šafárik, Košice, Slovakia

ABSTRACT

The aim of this study is to assess the consequences of thyroid surgery on the voice of patients, in order to identify various perturbations which this surgery may provoke, and also to reveal possible compensatory strategies or readjustments that the patient may develop. The assessment was based on the analysis of acoustic signals, from which a large amount of cues related to voice quality was extracted. The experiment deals with the spectral characteristics of the voice of patients who underwent thyroidectomy, with no laryngeal paralysis. Our interpretations, made from the acoustic data, reveal perturbations of gestures on the glottal level, with readjustment strategies varying according to patients. The concept of a "target" as a control space for execution of possible articulatory and acoustic entities, which are perceptually acceptable, seems particularly relevant in this study, since disordered speakers reorganise their productions according to their own physiological and anatomical constraints, caused by the disorder.

Keywords: speech production, speech disorder, clinical phonetics, perturbations, readjustments.

1. INTRODUCTION

Thyroid gland surgery is often carried out in cases of impaired function, if the thyroid is too large or if the patient has a tumor. Laryngeal immobility may occur as the result of this operation and is a complication frequently associated with such a highly codified surgery.[11] Recurrent nerve paralysis remains one of the most studied examples of complications that affect voice quality after thyroidectomy (see, for example [1]'[13]). Operation of the larynx is rarely anecdotal. When removing the thyroid gland, the surgeon must take great care to preserve the two recurrent nerves responsible for the mobility of the vocal cords. Even if the laryngeal nerves are preserved, thyroidectomy may impair a patient's voice[6] and impede swallowing.[14][15] In terms of the larynx, the voice and speech of the patient may be altered after surgery with the following consequences: a breathy or hoarse voice, irregular laryngeal vibrations or diplophonia, a modification of voice periodicity,[6] an augmentation of jitter, shimmer[18] and noise relative to harmonics, a reduction of intensity,[8] voice fatigue[12] and an alteration of airflow, etc.[10] Finally, dysphonia following endotracheal intubation (otherwise known as hoarseness) may occur even when there is no visible damage to the vocal cords.[21] Without laryngeal damage, dysphonia declines rapidly and spontaneously.[9] The principal aim of this study is to analyse some of the spectral characteristics of the voices of patients without diagnosed paralyses, but with alteration of the voice when producing sustained vowels. This longitudinal study falls within the perturbation and readjustment paradigm. It is primarily concerned with evaluating the flexibility of the speech production and perception system, and determining the range of linguistically tolerated deviations from speech "targets".

METHOD

Patients

For this specific piece of work, 7 patients have been recorded: 5 women (anonymised as follows: NPPGER, NPPHOE, NPPHOF, NPPKRE and NPPLEN) and 2 men (anonymised as follows: NPPENS and NPPKAU). All of the patients are native French speakers who are aged between 60 and 70 and underwent total thyroidectomy for benign nodules. In all cases, the pre-operative laryngeal examination showed no impairment of vocal cord mobility.

Data was obtained in 3 phases: 1.) a pre-surgery phase, the day before surgery, in order to obtain the patient's unaltered reference voice 2.) a first post-surgery phase, the day after surgery, when the patients' voices had altered to varying degrees 3.) a second post-surgery phase, 15 days after surgery, in order to observe probable voice and speech recuperation.

Corpus

The corpus consisted of the three extreme vowels, /i, a, u/, as these allowed us to explore the maximum vowel space of each speaker. The three vowels were presented on a card in a random order and the subject was required to pronounce each vowel 10 times and for about 5 seconds *per* vowel.

Measurements

The measurements were acquired with the software PRAAT ©.[4] For each vowel in the steady-state portion we measured: 1.) the fundamental frequency (F0 in Hertz) 2.) the harmonics-to-noise ratio (HNR in dB). HNR can be considered to be a quantitative index of the degree of vocal hoarseness, i.e. the extent to which noise replaces harmonic structure in the spectrogram of sustained vowels[3][22][23] 3)the Vowel Space Area (VSA) in kHz^2 using Heron's formula. This value (in kHz^2) provides information related to the space used in differentiating vowels.[2] We are aware of other methods to calculate vowel spaces (see, for example, [7],[20]), but medical constraints meant that we were unable to ask patients to perform more than ten repetitions of the three extreme vowels 4.) the jitter and the shimmer.[17],[16]

HYPOTHESES

It is hypothesised that in the first post-surgery phase: 1.) difficulty in controlling the voice could affect F0 values 2.) altered laryngeal activity could affect usual HNR target values 3.) changes in fundamental frequency may result in changes in the voice's periodicity, which could then lead to a modification in jitter and shimmer values 4.) perturbations of the voice at the source could have an impact on supraglottal resonances, i.e. on F1 and F2 values 5.) perturbations of formant values may affect the size and shape of vowel spaces 6.) in time and with speech therapy the aforementioned parameters could be improved and may even begin to resemble their reference values.

RESULTS

General Remarks

An analysis of variance (ANOVA, repeated measures) was carried out on the data of patients without paralysis (No Paralysis Patients or NPP). These analyses were made to suit all variables (F0, F1, F2, harmonics-to-noise ratio (HNR) and Vowel Space Area (VSA)). The aim was to determine if there were effects for the NPP group in the pre- and post-surgery phases.

The main effect of the recording phase was significant ($p<0.05$) for variables F0 and HNR (see the results of the statistical analyses below). As a result, post-hoc pair-wise comparisons (Tukey's HSD) were carried out on the mean values for these variables.

Fundamental Frequency (F0) Values

It should be noted that close analysis of the fundamental frequency values of the male speakers was carried out independently of those of the female speakers in order to ensure that our statistical measures were not distorted. The main effect of the recording phase was significant for the variable F0 ($p<0.05$) for the female speakers. These results indicate that, when all the settings are combined, fundamental frequency values are modified according to the recording phase. Overall, they suggest that a decrease in fundamental frequency is to be expected in the post-operative phase. Furthermore, these results demonstrate that the interaction of vowels and time factors was not significant (p=ns), indicating that fundamental frequency values will be modified over time whatever the vowel study.

Fundamental frequency values decrease for the female speakers between the pre-operative and first post-operative phases. When recording in the second post-operative phase, F0 values increase for all of the speakers. This phenomenon is observed regardless of the vowel produced (see Figure 1). Figure 1 shows that vowel /i/ varies from 207.75 Hz (std=23.45 Hz) in the pre-surgery phase to 172.47 Hz (std=38.64 Hz) in the first post-surgery phase. Reduction of F0 values is also observed for /a/ and /u/, whose values go from 202.20 Hz (std=17.74 Hz) to 156.82 Hz (std=26.87 Hz) and from 209.92 Hz (std=18.18 Hz) to 180.22 Hz (std=31.86 Hz), respectively. In the second post-surgery phase, all fundamental frequency values increase: F0 is at 179.95 Hz for /i/ (std=29.49 Hz), 167.66 Hz for /a/ (std=23.71 Hz) and 182.82 Hz (std=30.39 Hz) for /u/.

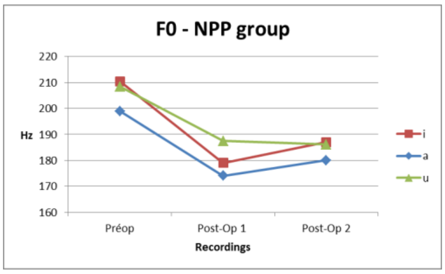

FIGURE 1 – F0 mean values

HNR Values

The main effect of the recording phase was not significant for the HNR variable (p=ns). However, the main effects of the speaker and vowel were significant (p<0.05), therefore implying some inter-speaker variability.

For the patients, HNR values are lower in the first post-operative phase than those measured in the pre-operative phase and in the second post-operative phase. Regardless of the recording phase, however, the measurements remain very close to the expected values for each of the vowels (see Figure 3). The first post-operative phase is also characterised by larger standard deviations.

Figure 3 shows that before surgery, /i/ was at 25.72dB (std=3.98), /a/ at 23.47dB (std=3.55) and /u/ at 30.28dB (std=3.60). In the first post-surgery phase, all of the values decrease. HNR is at 22.64dB for /i/ (std=8.35), at 17.31dB for /a/ (std=8.35 dB) and at 24 dB for /u/ (std=9.97 dB). In the second post-surgery phase, these ratios resemble those measured before surgery: /i/ is now at 23.94 dB (std=8.16), /a/ at 19.68dB (std=8.89), and /u/ at 25.65 dB (std=10.28).

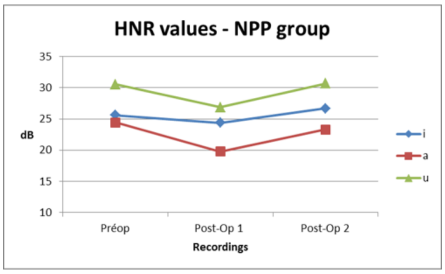

FIGURE 2 – HNR mean values

Jitter

An ANOVA was not able to identify a main effect of time for the jitter variable, indicating that this parameter was calculated at around the same value no matter what recording session was considered. The jitter study shows that this parameter increased between the pre-operative phase (0.42% std 0.17) and the first and second post-operative phases, where it is measured at 0.84% but does not exceed the normal/pathological threshold. Standard deviations are larger for the post-operative phases: they are 0.72 and 0.97 respectively, indicating that the output of some patients may exceed the normal/pathological threshold.

During the second post-operative phase, jitter values decrease for all of the speakers when compared with the previous recording phase.

The standard deviation of intra-speakers is low whatever the recording phase.

Shimmer

An ANOVA in the shimmer gave similar results to those observed in the jitter, namely that a main effect of time had not been reached (p=*ns*). However, it brought interaction between the speaker and time factors to light (p<0.05), thereby reflecting a degree of inter-individual variability.

Following the same trend as the jitter, the shimmer measurement increases in the first post-operative stage where it is measured at 5.45% (against 3.08% pre-operatively). The shimmer then decreases during the last recording phase where it is measured at 4.28%. Average measurements for the two recording phases are beyond the normal/pathological threshold.

The post-operative phases are characterised by significant standard deviations (4.03 and 3.74, respectively), which probably reflects a degree of individual variability.

Vowel Space Area (VSA) Values

Small changes in formant values had an impact on the area of the vowel space. An ANOVA indicated a main effect of the recording phase for the variable area of vowel space. The vowel space area was significantly reduced (p<0.05) for patients between the pre-operative and first post-operative stages. It then re-increased (p<0.05) in the second post-operative phase without reaching its reference value in the pre-operative phase. The vowel space is 0.34kHz², 0.21kHz² and 0.28kHz², respectively. Note that even if the vowel space is reduced, it remains geometrically conventional.

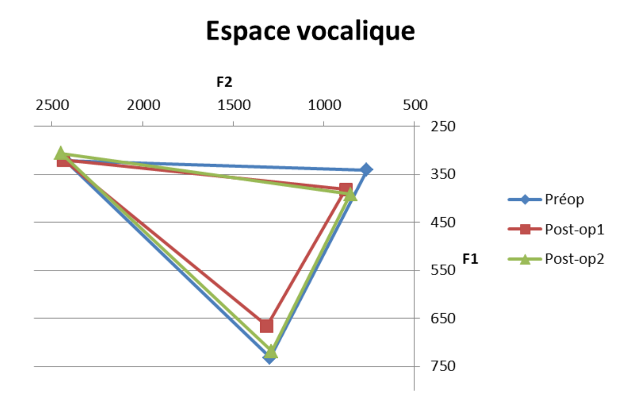

FIGURE 3 – Vowel Space area values

Table 1 shows individual changes for the VSA values. For four of the patients (NPPGER, NPPHOE, NPPHOF and NPPKAU), vowel spaces are reduced dramatically between the pre-operative phase and the first post-operative phase. For NPPHOF and NPPKRE, the reduction in vowel spaces is smaller. NPPENS and NPPLEN demonstrate the opposite of this trend as their vowel space is slightly larger in the first post-operative phase than the pre-operative phase. During the second post-operative phase, the vowel space area of all the speakers, except NPPENS and NPPLEN, increases again.

TABLE 1 – Vowel Space Area values according to the speaker

kHz²	Pre-Op	Post-Op 1	Post-Op 2
NPPENS	0,29	0,31	0,17
NPPGER	0,46	0,15	0,34
NPPHOE	0,22	0,03	0,16
NPPHOF	0,42	0,34	0,38
NPPKAU	0,20	0,07	0,17
NPPKRE	0,33	0,22	0,37
NPPLEN	0,38	0,40	0,33

DISCUSSION

In terms of tendencies, our initial hypotheses have been largely proven. 1.) F0 values were reduced. F0 values diminished in the post-operative phases and demonstrated spontaneous recuperation in the second post-operative phase (Hypothesis 1). 2.) Altered laryngeal activity affected the usual harmonics-to-noise ratio (Hypothesis 2). 3.) Jitter and shimmer values were higher in the post-operative phases due to modifications in voice periodicity. 4.) Formant values were perturbed for all of the patients (except for the vowel /i/). 5.) Vowel Space Areas were also perturbed for all of the patients. They either underwent reduction or geometric modifications, which implies a reorganisation of the extreme vowels in the maximal vowel space. This is probably due to an alteration at the voice source rather than to potential perturbations in the tongue's mobility. It is also plausible that the patient used gestural economic strategies in the post-surgery phases. 6.) Hypothesis 6 was related to the benefits of time. From the second post-surgery phase, patients without vocal cord paralysis recovered. Overall, all of the parameters measured resembled their initial reference values over time.

It should be noted that speaker variability is very high in the post-operative recording phases. It might therefore be useful to increase the number of speakers and place them into subgroups based on the impact of the surgery on their voices. Additional recordings are currently underway in order to increase the number of subjects in each group.

CONCLUSIONS

According to the literature and our data, voice quality can be modified without laryngeal immobility and timing recuperation is delayed for patients with diagnosed laryngeal immobility. Patients should be informed that there is a potential for modification in voice quality after thyroidectomy, even if the laryngeal nerves are preserved. This study reinforces the importance of pre- and post-operative laryngeal evaluations that assess the direct effects of thyroid surgery on voice quality (see, for example, [8] and [5]). Voice disorders can be measured after thyroidectomy, even in cases where the laryngeal nerves are preserved. This trend could be explained by the fact that thyroid surgery can affect the supraglottic musculature, which naturally impacts upon acoustics.[19] At least in the short term, laryngeal surgery appears to have direct consequences for the patient's voice.

7. REFERENCES

[1] Benninger, Gillen and Altaian 1998. Changing etiology of vocal fold immobility. *The Laryngoscope*. 108, 9, 1346–1350.

[2] Blomgren, M., Robb, M. and Chen, Y. A note on vowel centralization in stuttering and nonstuttering individuals. *Journal of speech and hearing research*. 41, 5, 1042–1051.

[3] Boersma, P. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *IFA Proceedings 17* (University of Amsterdam, 1993), 97–110.

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

[5] Debruyne, F., Delaere, P., Wouters, J. and Uwents, P. 1994. Acoustic Analysis of Tracheo-Oesophageal Versus Oesophageal Speech. *The Journal of Laryngology & Otology*. 108, 04, 325–328.

[6] Debruyne, F., Ostyn, F., Delaere, P. and Wellens, W. 1997. Acoustic analysis of the speaking voice after thyroidectomy. *Journal of voice: official journal of the Voice Foundation*. 11, 4, 479–482.

[7] Fox R. A., J.E. 2008. Analysis of total vowel space areas in three regional dialects of American English. *The Journal of the Acoustical Society of America*. 123, 5, 3068.

[8] Hong, K.H. and Kim, Y.K. 1997. Phonatory characteristics of patients undergoing thyroidectomy without laryngeal nerve injury. *Otolaryngology - Head and Neck Surgery*. 117, 4, 399–404.

[9] Jones, M.W., Catling, S., Evans, E., Green, D.H. and Green, J.R. 1992. Hoarseness after tracheal intubation. *Anaesthesia*. 47, 3, 213–216.

[10] Keilmann, A. and Hülse, M. 1992. Dysphonie nach Strumektomie bei ungestörter respiratorischer Beweglichkeit der Stimmlippen. *Folia Phoniatrica et Logopaedica*. 44, 6, 261–268.

[11] Laccourreye, O., Malinvaud, D., Ménard, M. and Bonfils, P. 2009. Immobilités laryngées unilatérales après chirurgie de la glande thyroïde. *Journal de Chirurgie*. 146, 6, 553–558.

[12] Van Lierde, K., D’haeseleer, E., Wuyts, F.L., Baudonck, N., Bernaert, L. and Vermeersch, H. 2010. Impact of thyroidectomy without laryngeal nerve injury on vocal quality characteristics: An objective multiparameter approach. *The Laryngoscope*. 120, 2, 338–345.

[13] Lo, C.-Y., Kwok, K.-F. and Yuen, P.-W. 2000. A Prospective Evaluation of Recurrent Laryngeal Nerve Paralysis During Thyroidectomy. *Arch Surg*. 135, 2, 204–207.

[14] Lombardi, C.P., Raffaelli, M., De Crea, C., D’Alatri, L., Maccora, D., Marchese, M.R., Paludetti, G. and Bellantone, R. 2009. Long-term outcome of functional post-thyroidectomy voice and swallowing symptoms. *Surgery*. 146, 6, 1174–1181.

[15] Lombardi, C.P., Raffaelli, M., D’Alatri, L., Marchese, M.R., Rigante, M., Paludetti, G. and Bellantone, R. 2006. Voice and swallowing changes after thyroidectomy in patients without inferior laryngeal nerve injuries. *Surgery*. 140, 6, 1026–1034.

[16] Schoentgen, J. 2001. Stochastic models of jitter. *The Journal of the Acoustical Society of America*. 109, 4, 1631–1650.

[17] Schoentgen, J. 2006. Vocal Cues of Disordered Voices: An Overview. *Acta Acustica united with Acustica*. 92, 5, 667–680.

[18] Stojadinovic, A., Shaha, A.R., Orlikoff, R.F., Nissan, A., Kornak, M.-F., Singh, B., Boyle, J.O., Shah, J.P., Brennan, M.F. and Kraus, D.H. 2002. Prospective functional voice assessment in patients undergoing thyroid surgery. *Annals of surgery*. 236, 6, 823–832.

[19] Timon, C.I., Hirani, S.P., Epstein, R. and Rafferty, M.A. 2010. Investigation of the impact of thyroid surgery on vocal tract steadiness. *Journal of voice: official journal of the Voice Foundation*. 24, 5, 610–613.

[20] Vorperian, H.K. and Kent, R.D. 2007. Vowel Acoustic Space Development in Children: A Synthesis of Acoustic and Anatomic Data. *Journal of speech, language, and hearing research : JSLHR*. 50, 6, 1510–1545.

[21] Yamanaka, H., Hayashi, Y., Watanabe, Y., Uematu, H. and Mashimo, T. 2009. Prolonged hoarseness and arytenoid cartilage dislocation after tracheal intubation. *British Journal of Anaesthesia*. 103, 3, 452 –455.

[22] Yumoto, E., Gould, W.J. and Baer, T. 1982. Harmonics-to-noise ratio as an index of the degree of hoarseness. *The Journal of the Acoustical Society of America*. 71, 6, 1544–1549.

[23] Yumoto, E., Sasaki, Y. and Okamura, H. 1984. Harmonics-to-Noise Ratio and Psychophysical Measurement of the Degree of Hoarseness. *J Speech Hear Res*. 27, 1, 2–6.