

A spatiotemporal prospective study of speech in patients with or without recurrent laryngeal nerve paralysis after thyroid surgery

Camille Fauth¹ Béatrice Vaxelaire¹, Jean-François Rodier², Pierre-philippe Volkmar², Fayssal Bouarourou¹, Fabrice Hirsch³, Rudolph Sock¹

¹Université de Strasbourg, Institut de Phonétique de Strasbourg – IPS & U.R. 1339 Linguistique, Langues et Parole – LilPa, E.R. Parole et Cognition

²Centre Paul Strauss – Strasbourg, Département de Chirurgie Oncologique, Centre Régional de Lutte Contre le Cancer

³Université Paul Valéry - Montpellier III, Praxiling UMR 5267, CNRS, et Institut de Phonétique de Strasbourg – IPS & U.R. 1339 Linguistique, Langues et Parole – LilPa, E.R. Parole et Cognition

camille.fauth@gmail.com

Abstract. *This work is an acoustic investigation on some spectral and temporal characteristics of the speech of patients with recurrent paralyses, and also of patients without diagnosed paralyses but with alteration of their voice after thyroid surgery. Speech is evaluated during two tasks: producing sustained vowels and also a sequence of nonsense words. Consequences of surgery on the voice of patients are evaluated in order to identify the different perturbations that such a surgery may provoke, and also to uncover probable compensatory or readjustment strategies which a patient might deploy, alone or with the help of speech therapy. Globally, results show that due to surgery, some aspects of the speech production system are perturbed, as highlighted by the spatial and temporal parameters measured. However, these parameters tend to become normal as the patient recovers, in time and with the help of speech therapy. This is a longitudinal study.*

1. Background

The current preliminary study is part of programmatic research work that is being conducted in our laboratory. This specific research is being carried out in collaboration with the *Department of Surgical Oncology* of the *Paul Strauss Centre* in Strasbourg. This centre offers a privileged advantage for our research as it receives, every year, between 650 and 750 patients suffering from some pathology of the thyroid gland. Consequently, it is possible to carry out systematic analyses of the diverse consequences of thyroid surgery on the voice and speech of the patients, and also to conduct longitudinal studies on a large number of speakers.

Let us recall that the thyroid gland, located at the base of the neck, is responsible for secretion of hormones indispensable for important functions of the organism such as cardiac rhythm, intestine activity, body temperature, transformation of fats into sugar, appetite, humour, sleep, body weight equilibrium and growth in the child.

During ablation of this gland, the surgeon takes great care to preserve the two recurrent nerves, responsible for mobility of the vocal folds. However, after surgery, at the laryngeal level, the voice and speech of the patient may be altered with the following consequences: a breathy or hoarse voice, irregular laryngeal vibrations or diplophonia, modification of voice periodicity, augmentation of jitter and shimmer, and of noise relative to harmonics, a reduction of intensity, effects of voice fatigue, alteration of airflow, etc. In some cases, when one recurrent nerve is inadvertently injured during surgery, unilateral paralysis may appear.

The *aim* of this preliminary study is to analyse some spectral and temporal characteristics of the speech of patients with *recurrent paralyses* and also of patients without diagnosed paralyses but with *severe alteration* of their voice and speech after thyroid surgery. This longitudinal study falls within the *perturbation and readjustment paradigm* preoccupied mainly with evaluating the *flexibility* of the speech production and perception system, determining the range of linguistically tolerated *deviations* from speech “targets”, and ultimately elaborating *viability constraints* in articulatory and acoustic terms.

2. Method

2.1. Patients

Today, 21 *speakers* (19 women and 3 men) have been recorded. This corresponds to 7 hours of audio recordings. All patients are native French speakers, between 52 and 60 years old. A short biography of each subject is kept in the medical file, with information pertaining also to pre- and post-surgery. For this specific work, 4 patients have been analysed, 2 patients with unilateral paralysis and 2 patients without unilateral paralysis, all the four patients have undergone thyroidectomy. When a patient had only been recorded in post-surgery phases, data are acquired from a control speaker, matched with age and gender, with no voice pathology and no surgery in the vocal tract area.

2.2. Corpus and recordings

The *corpus* consisted of *sustained vowels*, *VCV* nonsense words and a read *text*. The sustained vowels were the 3 extreme vowels /i, a, u/, which allow exploring the maximum vowel space of each speaker. As concerns the *VCV* sequences, when V1 was vowel /i/, V2 was vowel /a/, and when V1 was vowel /a/, V2 was vowel /i/. The consonant was one of the 6 plosives /p, t, k, b, d, g/. The combination of these sounds allowed creating 12 nonsense words that were inserted in the carrier sentence “Cet _____ ça”. Vowels /i/ and /a/ were specifically chosen in order to observe the acoustic consequences of larynx trajectory from a low position for /i/ to a high position for /a/, and *vice versa*. Consonants /b, d, g, p, t, k/ were retained in order to observe for effects of place of articulation on intra-oral pressure in relation to timing between burst-release and onset of vocal fold activity (VOT as defined by Lisker & Abramson, 1964), or timing between burst-release and onset of a clear formant structure (VOT as specified by Klatt, 1975). These six consonants were also chosen as they allow comparing entirely voiced sequences V1/b, d, g/V2 with partially voiced sequences V1/p, t, t/V2.

The latter sequences should impose more constraints on vocal fold activity since they require alternating closed and open configurations of the glottis. Reading a text allows evaluating the effects of vocal fatigue on the speaker's production.

During *recordings*, the vowel to be pronounced was presented to the subject on a cardboard in a graphic form /i, a, ou/ which the experimenter was sure would elicit the desired pronunciation. The three vowels were presented in random order, and the subject was required to pronounce each vowel, sustained for about 3 seconds, 10 times *per* vowel. The same method was also used to record the VCV nonsense words. The short text was read at a self-selected normal speaking rate. The data were obtained in 3 phases: 1) a pre-surgery phase, the day before surgery, which gives us the patient's unaltered *reference* voice; 2) a first post-surgery phase, the day after surgery, when the patient's voice is altered in varying degrees, due to paralysis or not of the vocal folds; 3) a second post-surgery phase, 15 days after surgery, which allows observing probable voice and speech recuperation. If unilateral paralysis is diagnosed, the patient will be recorded once a month during his or her entire speech therapy.

2.3. Measures

Measurements obtained were: VOTs (defined *supra*), VTT or Voice Termination Time, *i.e.* voicing decay after closure within the voiceless consonant hold in VCV sequences; vowel (V1 and V2) and consonantal durations; F0, F1 and F2 values in sustained vowels.

2.4. Hypotheses

It is *hypothesised* that in the first post-surgery phase, due to potential paresis or paralysis of the vocal folds: 1) VOT and VTT values, which respectively require precise offset and onset of laryngeal activity, relative to articulatory closure and release of the consonant, would be significantly longer; 2) Perturbation of the timing of VOT and VTT, as intra-segmental consonantal measures, would consequently have an impact on the duration of the consonantal duration, which should be longer, at least in the VOT transitional phase towards V2; 3) Difficulty in controlling voicing would also affect vowel durations; 4) Altered laryngeal activity would modify coupling between the larynx and the vocal tract, such that perturbation of voicing at the source would also have an impact on supraglottal resonances, *i.e.* on F1 and F2 values, and may affect the size and shape of vowel spaces; 5) Difficulty in controlling voicing would expectedly affect F0 values; 6) With speech therapy, the abovementioned parameters would be less modified, and may begin to resemble reference values observed for the patient in the pre-surgery phase, or those of the control speaker when no pre-surgery recordings are available.

3. Results

Preliminary *results*, based on mean values and standard deviations when patients had difficulties producing sufficient items, and based on ANOVAs when patients succeeded in carrying out the tasks required of them, show that: 1) VOT and VTT values had a *tendency* to be longer in post-surgery phases, and especially in the first post-surgery

one; *variability* as highlighted by standard deviations, was significantly more pronounced in post-surgery productions (typical e.g. of Klatt's VOT mean value for /iba/: 17ms for the control speaker (CS), std=4ms; 31ms in second post-surgery phase, std=15 ms; 24 ms in third post-surgery phase, std=7ms; 19ms in the fourth post-surgery phase, std 5ms).

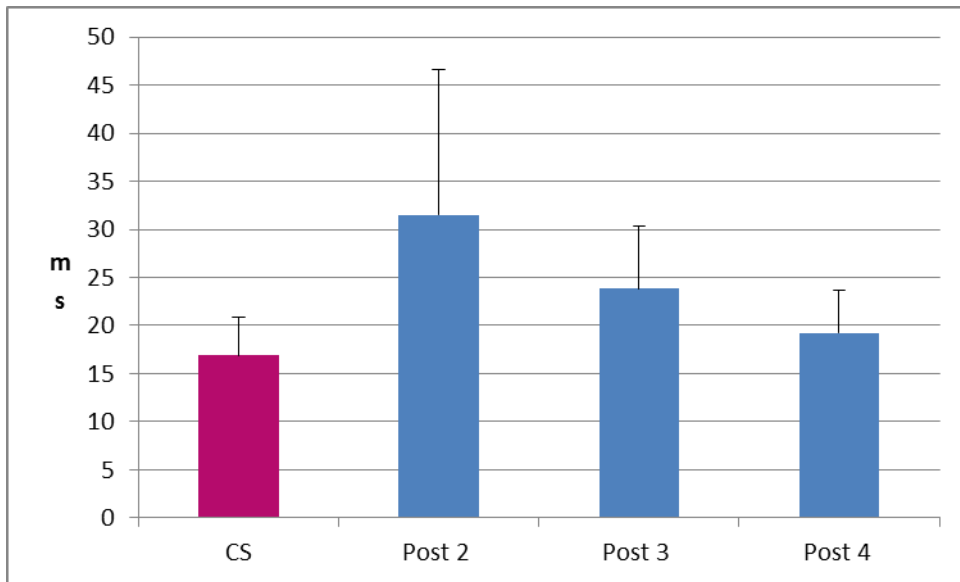


Figure 1. VOT values for [iba] for patient PLPH with unilateral paralysis

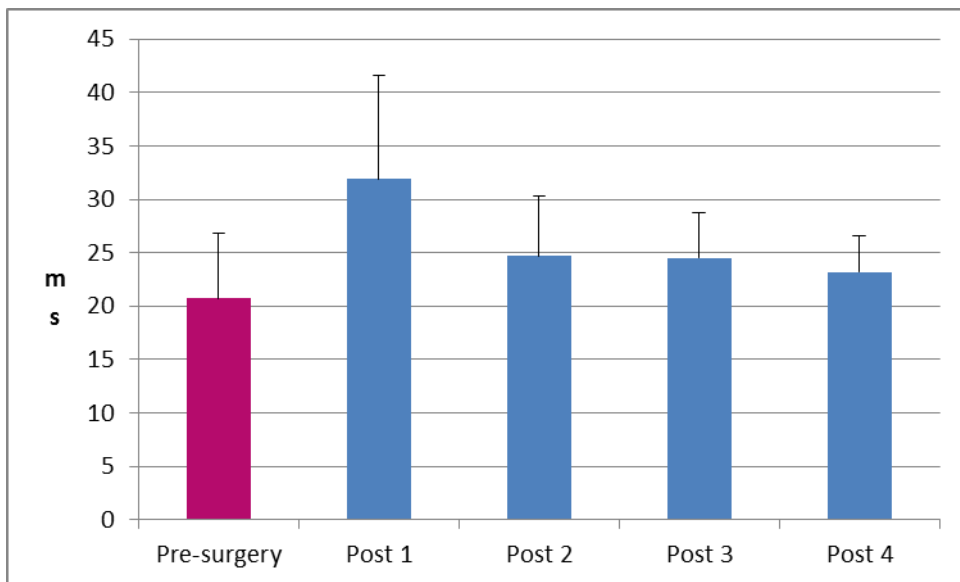


Figure 2. VTT values for [ika] for patient PLPH with unilateral paralysis

This finding is also true for V1 and V2 durations. See Figure 3 for a typical example of vowel durations for a patient (PLPH) with unilateral paralysis (example of vowel durations for /iba/: V1 /i/ 67ms for the control speaker (CS), std=15ms; 138ms in

second post-surgery phase, std=18ms; 122ms in third post-surgery phase, std=28ms; 92ms in the fourth post-surgery phase, std 17ms and V2 /a/ 68ms for the control speaker (CS), std=11ms; 80ms in second post-surgery phase, std=18ms; 70ms in third post-surgery phase, std=14ms; 73ms in the fourth post-surgery phase, std 8ms).

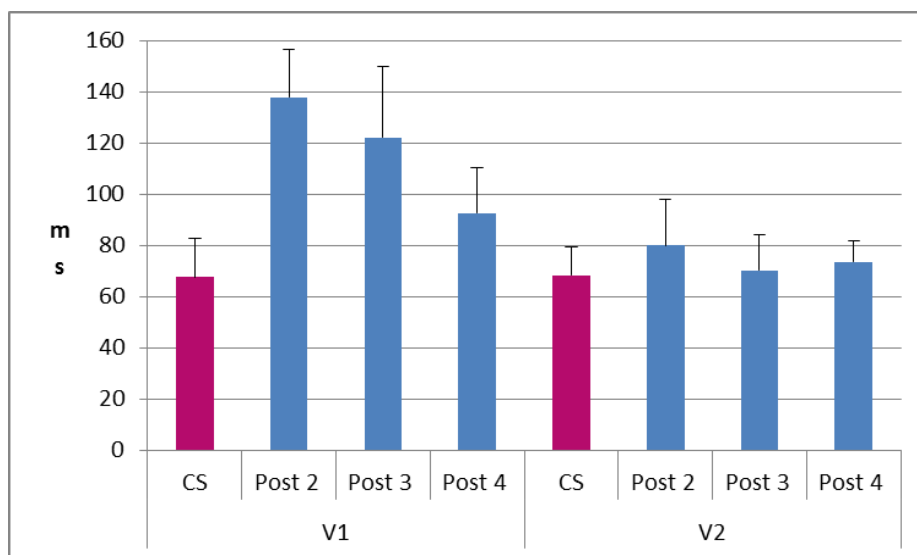


Figure 3. V1 and V2 durations for [iba] for patient PLPH with unilateral paralysis

Vowel spaces are significantly modified ($p < .01$) in the first post-surgery phase compared with the pre-surgery one; they look more conventional in the second post-surgery phase for patient (NPPG) without unilateral paralysis (see Figure 4). Figure 5 shows vowel space areas for subject CS (control speaker) and subject PLPH (patient with unilateral paralysis). Here also, differences are statistically significant ($p < .01$).

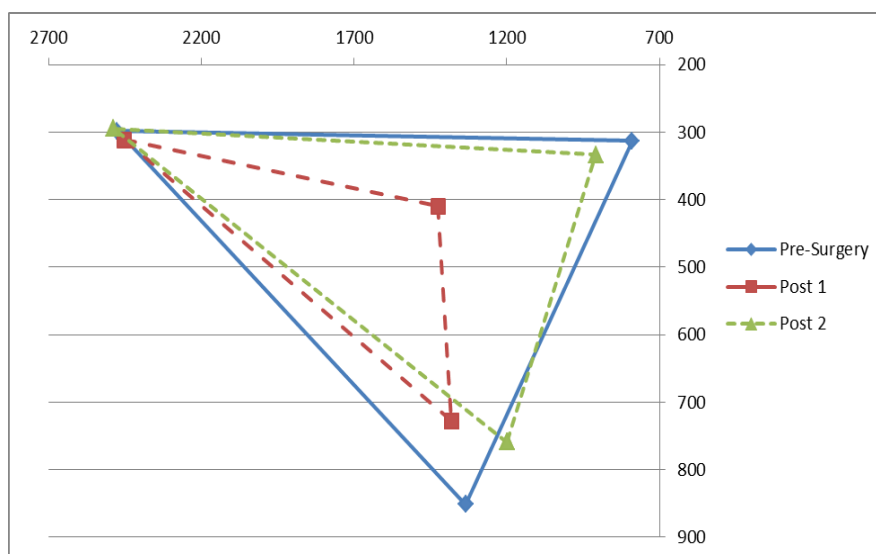


Figure 4. Vowel space areas for patient NPPG without unilateral paralysis

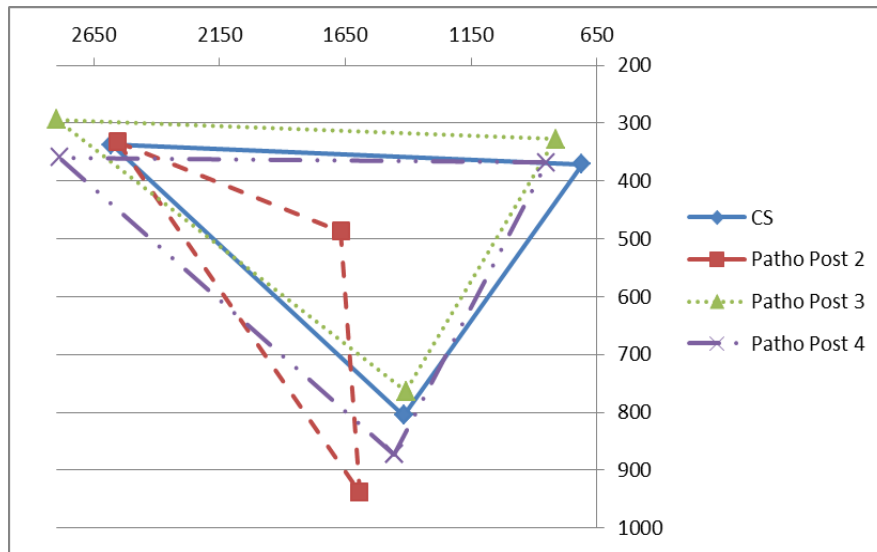


Figure 5. Vowel space areas for patient PLPH with unilateral paralysis

F0 values diminish significantly ($p < .01$) in the first post-surgery phase, compared with the pre-surgery phase (see Figure 5 for an illustration of this fact). Figure 6 shows that vowel /i/ varies from 219 Hz (std=7) in the pre-surgery phase to 187 Hz (std=15) in the first post-surgery one. Reduction of F0 values is also observed for /a/ and /u/, whose values go from 187 Hz (std=17) to 157 Hz (std=16) and from 219 Hz (std=13) to 157 Hz (std=11), respectively. In the second post-surgery phase, all F0 values increase; F0 is at 209 Hz for /i/ (std=4), 193 Hz for /a/ (std=8) and 212 Hz (std=12) for /u/.

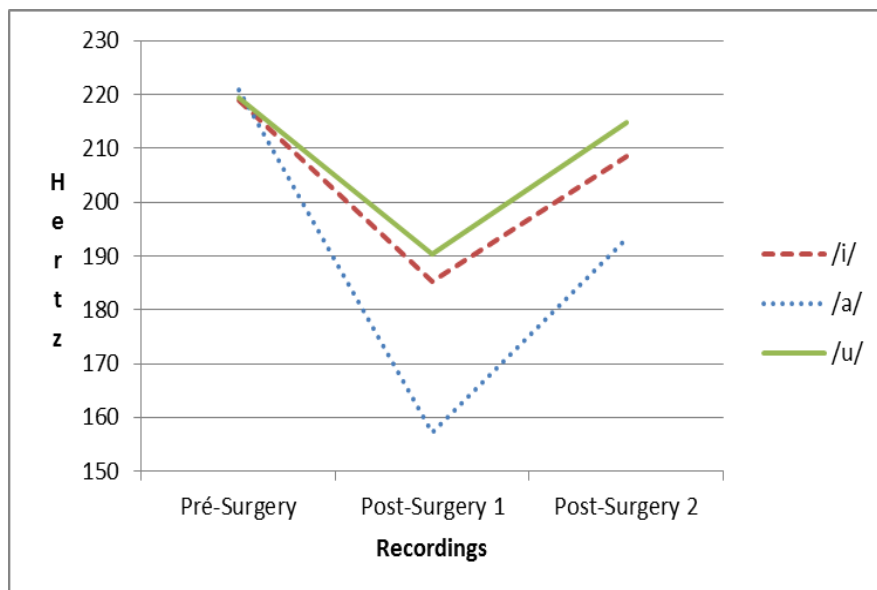


Figure 6. F0 for patient NPPH without unilateral paralysis

Reduction in F0 values is clear-cut for patients with unilateral paralysis in the post-surgery phase, compared with those of their respective control speakers (see Figure 7 for a typical example). It can be seen in the figure that F0 for CS is at 219 Hz (std=7) for /i/, 187 Hz (std=30) for /a/ and 219 Hz (std=31) for /u/. The patient with paralysis, in the second post-surgery phase, has lower F0 values, compared with those of the corresponding control speaker (CS): /i/ is at 172 Hz (std=12), /a/ at 155 Hz (std=17) and /u/ at 160 Hz (std=13). Results that are being analysed for the third post-surgery phase (130 Hz (std=50) for /i/, 130 Hz (std=40) for /a/ and 126 Hz (std=49) for /u/) and the fourth post-surgery phase (174 Hz (std=30) for /i/, 174 Hz (std=16) for /a/ and 177 Hz (std=33) for /u/) seem to indicate increase of F0 values, with time and speech therapy.

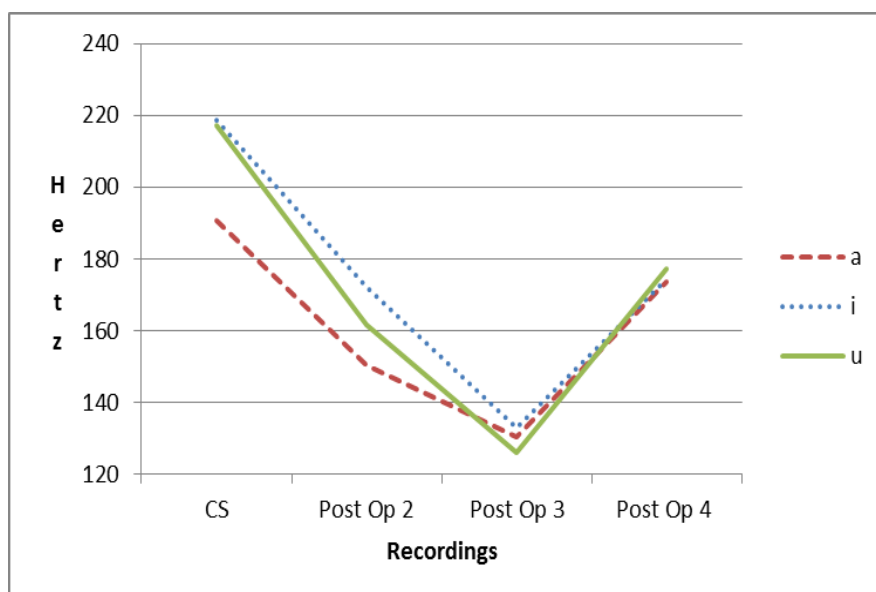


Figure 7. F0 values for patient PLPW with unilateral paralysis

4. Conclusions and discussion

As concerns hypotheses 1 and 2, VOT and VTT values *tend* to be longer in post-surgery phases. However, such results should be taken with caution as they rely mostly on nonsense words, and also reveal high inter-speaker variability. Systematic analyses of these parameters will be pursued in order to be able to evaluate their robustness in the current investigation.

Vowel duration has proven to be a relevant parameter in the current study (hypothesis 3). It seems that difficulty in controlling vocal fold vibrations has had an impact on the length of vocalic segments.

Vowel space area is reduced in post-surgery phases, in conformity with hypothesis 4. This is probably due more to alteration of the voice source than to potential perturbations in tongue body mobility. It is also plausible that the patient, in a post-surgery phase, is using gestural economy strategies. However, the latter conjectures must be verified with the help of articulatory data.

.

Hypothesis 5 is also confirmed, as F0 values diminish in post-surgery phases, with spontaneous recuperation in patients without paralysis, in the second post-surgery phase. Recuperation is slower in patients with unilateral paralysis, however speech therapy and time seem to have a positive effect on F0 values.

Hypothesis 6 was related to the benefits, in time, of speech therapy. Indeed, therapy has a positive effect on altered parameters in post-surgery phases. Patients without vocal fold paralysis recover, with therapy, as from the second post-surgery phase. Patients with unilateral paralysis, show slow recuperation as from the beginning of speech therapy; which is continually beneficial to the patient.

This investigation is being pursued, extending data collection to a larger set of subjects, with or without paralysis. Articulatory data will also be collected in order to confirm some of our hypotheses.

Acknowledgments

This work was supported by the program « *Perturbations et Réajustements : parole normale vs parole pathologique* » of the Maison Inter-Universitaire des Sciences de l'Homme Alsace (MISHA), 2008-2012, and also by an ANR Grant "DOCVACIM" to the Phonetics Institut of Strasbourg / U.R. LiLPa, E.R. Parole et Cognition.

References

- Crevier-Buchman L. & Ruiz C. Speech and voice rehabilitation following organ preservation surgery in Weinstein G., Laccourreye O., Brasnu D., Laccourreye H. *Organ Preservation Surgery of the Larynx*. Eds Singular Publishing Group, Inc. P171 175, 2000.
- Clément P, Hans S, Hartl DM, Maeda S, Vaissière J, Brasnu D. Vocal tract area function for vowels using three-dimensional magnetic resonance imaging. A preliminary study. *J Voice* ; e-pub, 2006.
- Hartl DM, Crevier-Buchman L, Vaissière J, Brasnu D. Phonetic effects of paralytic dysphonia. *Ann Otol Rhinol Laryngol*, 114:792-8, 2005.
- Hartl DM, Hans S, Vaissière J, Brasnu DF. Laryngeal aerodynamics after vocal fold augmentation with autologous fat versus thyroplasty in the same patient. *Arch Otolaryngol Head Neck Surg* ; 131:696-700, 2005.
- Scott AR, Chong PS, Hartnick CJ, Randolph GW, Spontaneous and evoked laryngeal electromyography of the thyroarytenoid muscles: a canine model for intraoperative recurrent laryngeal nerve monitoring. *An Otol Rhinol Laryngol*. 119(1):54-63, Jan 2010.
- Vaxelaire B., Sock R. Laryngeal movements and speech rate. An X-ray investigation. *Eurospeech'97. fifth European Conference on Speech Communication and Technology*, 22-25 septembre 1997, Rhodes, Grèce. Vol. 2, 1039-1042, 1997.