PERCEPTUAL AND ACOUSTIC CORRELATION IN CONSONANT IDENTIFICATION AFTER PARTIAL LARYNGECTOMY

L. CREVIER-BUCHMAN †&*, S. MAEDA ‡ , D. BRASNU * , J.VAISSIERE†

† Laboratoire de Phonétique et Phonologie (UMR 7018), CNRS/Sorbonne Nouvelle, Paris, France.

‡ Ecole Nationale Supérieure des Télécommunications, CNRS, Paris, France.

* Voice, Speech and Swallowing Laboratory, Hôpital Européen Georges Pompidou, AP-HP, Paris, France.

E-mail: lise.buchman@wanadoo.fr

ABSTRACT

The purpose of this study was two fold: i) to evaluate the voicing feature and ii) to determine the correlation between perceptual and acoustic characteristics of French voiced and voiceless stop consonants produced by 10 patients after supracricoïd partial laryngectomy (SCPL). In SCPL patients' speech, voice is produced by a neoglottis located at approximately 3 cm above the removed vocal folds, thus shortening the vocal-tract length. A perceptual evaluation was conducted for the 6 French stop voiced and voiceless consonants [p, t, k, b, d, g], in a syllabic context [CV]. The acoustic analysis of spectrograms served to determine what acoustic variables were used as cues for voicing distinction. The results are for i) perception test: voiced consonants were perceived as voiceless consonants (24%), ii) acoustic analysis: the absence of voicing murmur, the weakness of the burst and the short duration of the consonant seems to be responsible for perceived voiceless consonants for our patient's voiced consonants. Consonant articulation appears to impose certain constraints on voicing ability in SCPL patients. Presumably, this poor voicing ability is the direct consequence of the mechanical properties of the neoglottis.

1. INTRODUCTION

Supracricoid partial laryngectomy (SCPL) modifies the vibrator as it consists in a complete resection of the two vocal folds (VF) and the paraglottic space. The neoglottis is formed posteriorly by one or two arytenoid cartilages that move forward and inward, and anteriorly by the suprahyoid epiglottis and the base of the tongue that move backward to get in contact with the arytenoids. Reconstruction is achieved by suturing the cricoid cartilage to the hyoid bone, shortening the vocal tract about 3 cm. The only intrinsic muscles of the larynx preserved are the posterior crico-arytenoid (PCA), the lateral crico-arytenoïd (LCA) and the inter-arytenoïd (IA). The PCA abducts the neoglottis [1, 2] and participates in voicing distinction of consonants and is an antagonist of the LCA and IA muscles. The new vocal tract can nevertheless produce voice [3, 4].

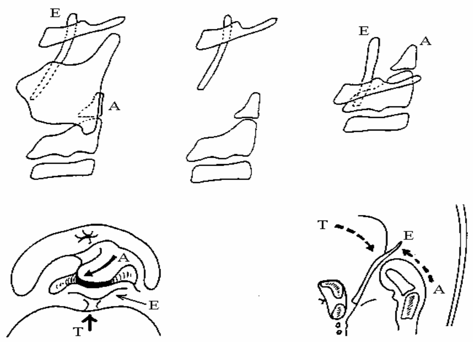

Figure 1 The top row schematizes, from left to right, the procedure of the supracricoid partial laryngectomy (SCPL) removing the vocal folds with the thyroid and the lower part of the epiglottis, followed by the pexy. The bottom row sketches the formation of a "neoglottis" for voicing, the view from above at the left and the sagittal view at right. The letter 'E', 'A', and 'T' denote respectively the epiglottis, arytenoid, and base of the tongue.

The larynx participates in distinguishing the voicing feature of consonants and vowels. In French, voicing is a distinctive feature. Confusion can occur between consonants with the greatest number of common features. Temporal perception cues play a predominant role in logatomal situations devoid of significance where cognition processes are not involved. [5].

For perceptual identification of stop consonant plus vowel syllables [CV], one commonly sees, in spectrograms, several acoustic variables that might be important in the auditory identification of the stop consonant phones. The main voicing feature depends on the existence of low frequency periodicity during the stop. A short burst of noise at the beginning of the syllable is presumed to be the acoustic counterpart of the articulatory explosion. The frequency position of the burst could serve as a cue for distinguishing among /p/, /t/, /k/. The intensity of the burst is less strong for voiced stops as the intraoral pressure is lower than for unvoiced stops [7]. Other cues might be found such as the VOT, and the length of the consonant, which is shorter for voiced stop consonant. Place of

ISBN 1-876346-48-5 © 2003 UAB

articulation can be distinguished by the general shape of the spectrogram and the slop of the formant frequencies transition [7, 8]. Such shifts in the frequencies of the vowel formants presumably reflects the articulatory movement.
The aim of this study was to understand how much potential do SCPL patients have for producing voicing distinction for consonants. Clinical experience has demonstrated however that these patients achieve a certain degree of independence for verbal communication in their social and occupational activities and proves to have a certain degree of effectiveness. Our study is the first prospective description and analysis of the perceptual and acoustic characteristics of voice and speech after SCPL. There were two objectives for this study conducted in ten SCPL patients. *i*) Identify stop consonants [p, t, k, b, d, g] in French and conduct perceptual tests for the voicing feature of these consonants in syllabic situations. *ii*) Better understand why confusions may occur by examining the perceptual and acoustic correlation.

2. METHOD

2.1 Speakers and corpus

Ten patients were recorded 18 months after SCPL. All were men, aged 51 - 73 years (mean, 63). The patients were asked to read a list of C-V syllables to explore the articulatory paradigm with C = [p, t, k, b, d, g] and V = /a/, recorded on a DAT recorder (DS 60, Sony, France) using a 22000 Hz sampling frequency. Non-sense syllables CV were preferred to words to avoid contextual information. We chose the vowel /a/ because of its relative frequency in the French language.

2.2 Listening test

Panels were composed of three expert listeners, two women and one man aged 25 to 35 years, who were trained speech therapists. The 6 syllables from each patient were presented in random order. The listeners transcribed their responses using an open-response paradigm.

2.3 Acoustic analysis

Identification of specific acoustic cues were looked for on spectrograms: presence of periodic vibration in low frequency for voiced consonant, the duration of the consonant, the burst and its intensity, the slope of the first 3 formant frequencies.

3. RESULTS

Listeners pooled responses were converted to confusion matrices and analysed for voicing features. Voicing feature was analysed, to distinguish voiceless / voiced consonants. Data in the matrices were analysed by column and by row. The columns corresponded to the target phonemes the patient wanted to pronounce, the rows corresponded to phonetic categories perceived by the listeners. The agreement between the three experienced listeners, measured with kappa statistics, ranged between 0.87 (se = 0.018) and 0.83 (se = 0.016).

3.1 Voicing distinction

The results showed that jury members had a real problem distinguishing between voiced and voiceless consonants pronounced by the patients. They confused 24% of the voiced consonants with a voiceless homorganic consonant, compared with 4% of the voiceless consonants heard as a voiced consonant. The identification error was greater for the consonant [g] with 31% confusion with [k]. The labial consonant [b] was confused with [p] in 31%. For the alveolar consonant [d], there was confusion with [t] in 29% of the cases.
The voicing feature of voiced consonants was confused with the homorganic voiceless consonants, predominantly for stops. The patients favoured consonant articulation, to the detriment of voicing.

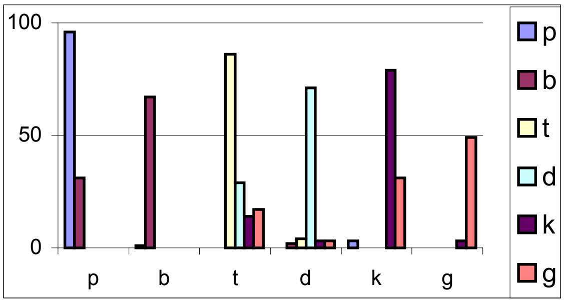

Figure 2 : histogram from the confusion matrices corresponding to the voicing feature of the stop consonants [p, t, k, b, d, g].

3.2 Acoustic cues

The transition slope, frequency and temporal parameters were measured. The absence of voicing murmur, the weakness of the burst, the rather flat formant frequency transition between consonant and vowel onset, and the short duration of the consonant seem to be responsible for perceived voiceless consonants for our patient's voiced consonants.

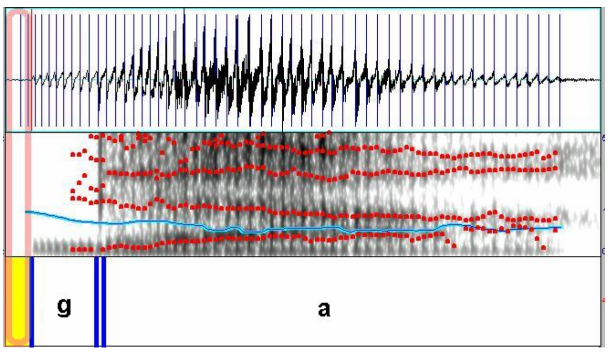

Figure 3 : spectrogram of the syllable [ga] produced by SCPL patient N°4. The low frequency voicing murmur, the burst and the formant transitions are visible although there is a lot of noise in those voices.

ISBN 1-876346-48-5 © 2003 UAB

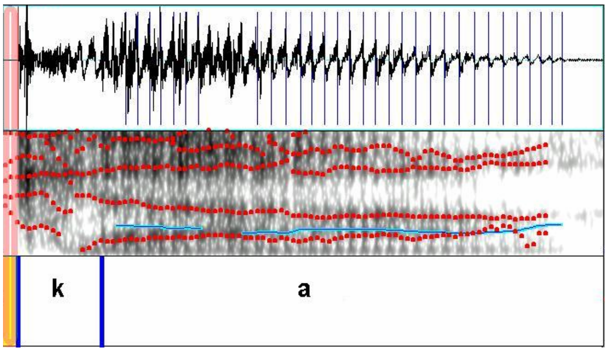

Figure 4 : spectrogram of the syllable [ka] produced by SCPL patient N°4. No voicing murmur, the burst and the formant transitions between the consonant and the onset of the vowel are visible although surrounded by a lot of noise.

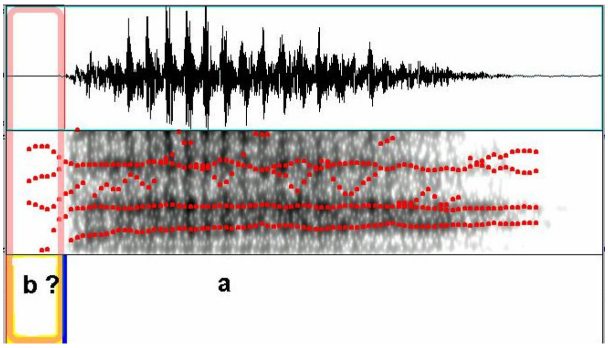

Figure 5 : spectrogram of the syllable [ba] produced by SCPL patient N° 2. No voicing was perceived and no voicing murmur was seen.

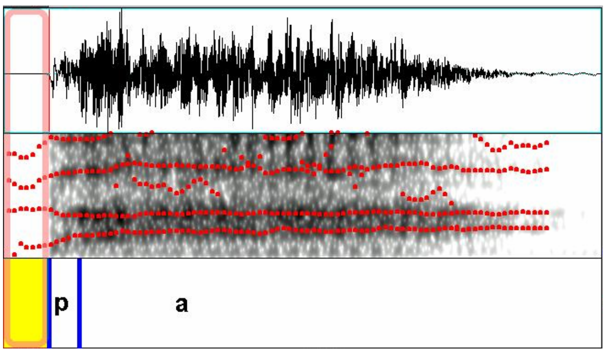

Figure 6 : spectrogram of the syllable [pa] produced by SCPL patient N° 2. No voicing was perceived and no voicing murmur was seen. It looks very much alike the [ba] spectrogram, but the acoustic signal of the vowel is more "noisy" and irregular after a voiceless stop consonant.

4. DISCUSSION

SCLP raises the larynx by 3 cm and modifies the volume of the vocal tract. These changes can produce more or less tension on the neolaryngeal vibrator. We have been unable to find any data in the literature related to the intelligibility of consonants pronounced by SCLP patients, making comparisons impossible.

4.1 Methodology

By considering 6 consonants, we were able to study the realization of distinctive voicing. The use of initial consonants in isolation could also have induced a certain degree of bias. First, the presence of a silence preceding the consonant could influence the results; when CV is embedded in a continuous stream, it may be easy to realize voicing in syllable-initial consonant: the adduction required to produce a voiced consonant in an intervocalic or postvocalic position is different from that required to initial phonation after à pause. Second, according to Bourciez and Bourciez [9], initial consonants are more easily identified than final consonants. These authors studied the evolution of words from Latin to French and observed that initial consonants have been preserved while final consonants have disappeared (for example, the Latin *campus* evolved into the French *camp* (camp): *campus* [?] > *campu* [?] > *camp* [?]). This observation represents an universal.

4.2 Voicing feature

As awaited, the main difficulty observed was in identifying the voicing feature of consonants. This can be explained by the fact that SCPL mainly modifies the anatomy of the vibrator. For SCPL voices, errors in recognising voicing feature mainly concerned voiced consonants heard as their unvoiced organic homologues with the same articulation position and the same constriction mode. The problem is thus limited to non-voicing. Indeed, to achieve phonation, the VF have to vibrate in a state of relative laxity [10]. The larynx has to be lowered to distend the VF giving them more flexibility. But after SCPL, the neolarynx is at least 3 cm above its normal position. Furthermore, to produce a voiced consonant, in addition to closing the VF, the pharyngeal cavities have to be widened to increase the volume and the length of the vocal tract between the site of constriction and the VF [11]. It is easier to achieve this widening for labial consonants /b/ where the volume of the cavity is easier to modulate, than for posterior consonants /g/ where the volume between the constriction and the source of the sound is reduced. After SCPL, the shorter vocal tract has only limited potential for size and volume adaptations; homorganic confusions of the voicing feature were more frequent for palatal consonants than for labial or alveolar consonants.

For Ladefoged [10], Stevens [12], and Ladegoged and Maddieson [13], voiceless stop consonants in normal phonation is associated with a state of muscle tension. This tension may be due to the ascension of the larynx which stretches the VC and narrows the pharyngeal cavities. But in SCPL patients, the pexia (sutured cricoid cartilage to hyoid bone with upraising of the larynx by 3 cm) has already created a state of muscle tension.

The voiceless stops /p/, /t/, /k/, result from the absence of VF vibration. For voiceless consonants, the posterior crico-arytenoid muscle, preserved after SCPL, is very active and opens the VF [14]. All these anatomic changes

ISBN 1-876346-48-5 © 2003 UAB

favour voiceless consonants, especially palatal consonants.

4.3 Acoustic cues

In spectrograms of stop consonant-vowel syllables, several acoustic variables might be important in the auditory identification of the stop consonant phones [15]. The short burst of noise, found near the beguinning of the syllable, and the frequency position of the burst could be the acoustic counterpart of the articulatory explosion and serve as a cue for distinguishing among [p, t, k]. A second possible cue to the perception of the stops lies in the formant frequency transition between the consonant and the vowel onset. It would reflect the articulatory mouvement. In our patients, the poor voicing ability is probably the direct consequence of the mechanical properties of the neoglottis.

5. CONCLUSIONS

After SCLP, changes in the neolaryngeal vibrator are the cause of voicing confusions, voiced consonants basically being perceived as their voiceless homologues. Ascension of the neolarynx and suture of the hyoid bone to the cricoid cartilage would be the cause of a conflict between articulation and voicing. The base of the tongue, which participates in occluding the neoglottis after SCPL, is crucial for voicing. However, correct articulation of consonants, respecting the site of constriction, requires a certain degree of freedom for tongue movement. The patients appeared to favour consonant articulation over voicing.

Analysing intelligibility and the voicing features as a phonetic dimension could be useful for managing the rehabilitation program for these patients.

REFERENCES

[1] KONRAD H.R., RATTENBORD C.C., KAIN M.L., BARTON M.D., LOGAN W.J., HOLADAY D.A. “Opening and closing mechanisms of the larynx”. *Otolaryngology Head and Neck Surgery*. **92** : 401-405, 1984.

[2] LUDLOW C.L., SEDORY S.E., FUJOTA M. “Neurophysiological control of vocal fold adduction and abduction for phonation onset and offset during speech”. In J. Gauffin & B. Hammarberg (Eds.), *Vocal Fold Physiology: Acoustic, perceptual and physiological aspects of voice mechanisms*. San Diego: Singular Publishing Group, 1991.

[3] CREVIER-BUCHMAN L., LACCOURREYE O., WEINSTEIN G., GARCIA D., JOUFFRE V., BRASNU D. “Evolution of speech and voice following supracricoid partial laryngectomy”. *The Journal of Laryngology and Otology*, **109** : 410-413, 1995.

[4] CREVIER-BUCHMAN L., LACCOURREYE O., WUYTS F.L., MONFRAIS-PFAUWADEL MC., PILLOT C., BRASNU D. “Comparison and evolution of perceptual and acoustic characteristics of voice after supracricoid partial laryngectomy with cricohyoidoepiglottopexy”. *Acta Otolaryngologica (Stockh.)*, **118** : 594-599, 1998

[5] SAERENS M, SERNICLAES W, BEECKMANS R. “Acoustic versus contextual factors in stop voicing perception in spontanious French”. *Language and Seech*, **32**, 291-314, 1989.

[6] SERNICLAES W. « Fenêtre de prélèvement temporel des indices d’occlusives”. *Actes des 13èmes JEP*, Bruxelles, pp. 69-78, 1984.

[7] STEVENS KN, KLATT DH. “Role of formant transitions in the voiced-voiceless distinction for stops”, *Journal of Acoustical Society of America*, **55**, 653-659, 1974.

[8] KEWLEY-PORT D. “Measurement of formant transitions in naturally produced stop consonant-vowel syllables”, *Journal of Acoustical Society of America*, **72 (2)** : 379-389, 1982.

[9] BOURCIEZ E., BOURCIEZ J. *Phonétique Française*. Klincksieck Ed, Paris, France, 1995.

[10] LADEFOGED P. “The linguistic use of different phonation types”. In D.M. Bless & J.H. Abbs (Eds.), *Vocal Fold Physiology: contemporary research and clinical issues.* San Diego: College Hill, pp. 351-360. 1983.

[11] OHALA JJ. “The relation between phonetics and phonology”, in *The Handbook of Phonetic Sciences,* Cambridge, Massachusetts, Ed. By WJ. Hardcastle, J. Laver, Blackwell Publishers Inc., pp. 674-693, 1997.

[12] STEVENS KN. “Vocal fold vibration for obstruent consonants”, In Gauffin J and Hammarberg B (Eds), *Vocal Fold Physiology: Acoustic, perceptual and physiological aspects of voice mechanisms*, San Diego, CA: Singular Publishing Group, pp. 22-36, 1991.

[13] LADEFOGED P. & MADDIESON I. *The Sound of the World’s Languages*, Cambridge, Massachustts, Blackwell Publishers Inc., 1996.

[14] HIROSE H. “Laryngeal adjustments in consonant production”, *Phonetica*, **34**, 140-164, 1977.

[15] LIBERMAN AM., DELATTRE PC., COOPER FS., GERSTMAN LT. “the role of consonant-vowel transitions in perception of stop and nasal consonants”. *Psychological Monograph*, **68, (8, Whole N° 379)**, 1-13, 1954.