VOICE INTELLIGIBILITY AFTER SUPRACRICOID PARTIAL LARYNGECTOMY

Lise Crevier-Buchman, MD°, Daniel Brasnu, MD°, Jacqueline Vaissière PhD#
°Voice and Speech Department, Laënnec Hospital, Paris, France;
 # Institut de Phonétique, CNRS/UPRESA 7018, Paris.

ABSTRACT

Supracricoid partial laryngectomy is the widest conservation technique for glottic carcinoma. Voice is produced without consonants and V = /a/) produced by 10 patients, vocal folds. 6 French listeners (3 naive and 3 experts) identified 2880 monosyllables [C-V] (where C represents the 16 French 18 months after surgery. Expert and naive listeners misidentified 23 % and 24 % of voiced consonants as voiceless, respectively . There was less than 2 % of manner confusions and there were 16 % and 28 % errors in place of articulation, mainly for palato-velar consonants, respectively. The sonorants /l, m, n, r/ were well identified. The jury's experience was somewhat determinant for consonant identification.

1.INTRODUCTION

Supracricoid partial laryngectomy [1] consists in the resection of the whole thyroid cartilage and paraglottic space. This surgery allows conserving the physiologic functions of the larynx (respiration, swallowing and phonation). The cricoid cartilage, the two mobile arytenoid cartilages, the hyoid bone and most of the epiglottis are conserved. Phonation results from the inward and forward movement of the arytenoid cartilage toward the remnant of the epiglottis[2]. The neoglottis has a T shape instead of the classical V shape. Voice is produced in the absence of vocal folds in the laryngeal tract but the rest of the vocal tract is not modified. Speech overall intelligibility is considered as satisfactory in every day life by the patients as long as it is produced in a quiet environment. Our goal here is to quantify the intelligibility of the consonants.

Voicing is phonologically distinctive in French for stops and fricatives. As the voice-source is modified, we can expect confusion in voicing, but not in the other dimensions (nasality, place of articulation, continuous/discontinuous, etc.). In our study, we considered the voicing distinction of the stops /p, t, k, b, d, g/ and of the fricatives /f, s, ʃ, v, z, j/ as well as the identification of the sonorants /l, m, n ,r/, for which voicing is not distinctive.

The second goal of this paper is to test the influence of the degree of experience of the listeners. We tested the relation between experienced and naive jury (i.e. not familiar with such kind of voices) to check if the training of the listening jury might influence the performance in identifying consonants distinction as it does for perceptual evaluation [3,4]. This is an important question, since for quality judgement, trained listeners are necessary.

2.METHODS

2.1. Population

A group of ten male patients, natives of French, who had undergone supracricoid partial laryngectomy 18 months earlier, was studied. Mean age was 64 years (range from 57 to 72).

The French listeners were set into two groups, A and N. The first group, A, consists of three trained listeners (speech-language therapists), who were familiar with this kind of voices. The second group, N, was composed by three students from a phonetic laboratory in Paris. Both groups were used to phonetic transcription.

2.2. Listening task and corpus

The ten patients were asked to read a list of monosyllables /C-V/ where C represents one of the 16 French consonants and V the /a/ vowel. The syllables were then digitized. The test was run on a Macintosh HyperCard stack. The syllables were randomly presented, three times each. The listeners could listen to each syllables at will. The six listeners were asked to write down on the computer, the consonant they heard. All together, there was 2880 stimuli representing the 16 consonants x 3 repetitions x 10 patients and x 6 listeners. The results were edited automatically as a text file and then exported to Excel for further statistical analysis.

3. RESULTS

Listeners' responses were registered in two confusion matrices for trained and naive listeners. Each of these two matrices included 1440 data (16 syllables x 10 patients x 3 repetitions x 3 listeners in each group). Consonant confusion was expressed in error percentage. Results are presented in two matrices, (see table 1 and table 2 at the end of this paper), one for the naive jury (N) and one for the trained jury (A).

3.1. Voicing distinction

The results confirm that the speakers have a real difficulty in producing the voiced/unvoiced distinction in obstruents. The present study revealed that voiced stops and fricatives in general were identified as voiceless in 23 % and 24 % for expert and naive jury, respectively.

The problem is the most severe for the consonant /g/. /g/ was recognized as a voiceless stop 48% of the times by the trained listeners, and 51% of the times by the untrained listeners. 31% of the /g/ was identified as /k/ for expert listeners and 40% for naive listeners. /g/ was identified as /d/ in 17% and 11% of the cases, by group A and N, respectively.

The labials presented the same type of confusion with /b/ being confused 31% and 23% for the two group of listeners into /p/.

29% of the alveolar consonants /d/ was also confused and heard /t/ .
A higher percentage of voicing confusions was expected for /g/ than for /b/, because of the reduced volume between the constriction and the place where phonation occurs. It is known that it is more difficult to maintain voicing in consonants with a more posterior place of constriction. Our results are compatible with such an explanation: voicing in /b/ is slightly better recognized than voicing in /d/ (73% versus 71%), and voicing in /g/ is the worst identified.
The problem arises also for the fricatives, but to a lesser degree. 14% and 20% respectively of the fricatives /j/ was identified as [ʃ]. Amplitudes of Voicing and of frication are known to be antagonistic, up to a certain level. There are less voiced fricatives than voiceless fricatives in the languages of the world. Closing the vocal folds for voicing is done to the expense of the volume of airflow going through the glottis, and there to the frication produced in the supraglottic cavity. But still, voicing in fricatives seems to be easier than voicing for stops, for the patients. This is probably due to the fact that the maintenance of voicing in stops is harder than in fricatives, because of a greater intra-oral pressure due to the complete obstruction of the air pathway.
Devoicing was predominant for the palato-velar stop consonants for the naive jury compared with expert jury for whom the confusion was equal for all groups of consonants.

3.2..Place of articulation confusion

/k/ and /t/ , and /g/ and /d/ were highly confused. Note that French /a/ is a rather front vowel, and /k/ before /a/ is palato-velar and not velar. The main confusion for the trained jury was 17% of the palato-velar /g/ heard as alveolar /t/ consonant. All together, there was 37% palato-velar identified as alveolar consonants. For the naive jury, 36% palato-velar identified as alveolar consonants, and 34% labial identified as alveolar consonants. The acoustic cues for the place of articulation of the stops consonants before /a/ are the formant transitions (F2 and F3 falling for bilabial, F2 and F3 rising for the dental, and F2 rising and F3 falling for the palato-velar) and the burst (diffuse and grave for labial, diffuse and acute for dental, and compact for the palato-velar). The confusion could be due either to the problem of voicing (part of the transition into the vowel is transformed into noise), or to a problem in placing correctly the tongue. We will argue later that the errors are due to surgery, but a detailed analysis of the formant transitions and of the burst shape should be done.
The place of articulation of the fricatives was generally well identified. This is due to the fact that frequencies enhanced in the frication noise are due to the cavity in the front of the constriction, and the perturbations at the level of the larynx do not have an effect (all the fricatives have a rather fronted place of articulation; different results may be explained with speakers from a different language, such as Arabis speakers).

3.3. Jury performance

Naive jury seemed to have more difficulty in identifying consonants pronounced by patients after supracricoid partial laryngectomy. It was the most relevant for place of articulation. The second group was trained for phonetic transcription, but it was not familiar with pathological voices. We did not test the improvement of the listeners with time, but it would be worth to do it.

4. DISCUSSION

4.1. Surgical considerations

To our knowledge, intelligibility distinction after supracricoid partial laryngectomy has not been studied. It has been studied after total laryngectomy [5], but the mechanism of voice production in that case is quite different and doesn't allow comparison. In our situation, voice is produced by a larynx but with no vocal folds. After supracricoid partial laryngectomy, the larynx is pulled upward, the vocal tract is shorter, and narrower. All these modification may contribute to a stiffer larynx.

4.2. Methodological considerations

The choice of monosyllables instead of words was to avoid a perceptual bias due to word intelligibility.
Note however that a large number of these monosyllabic words correspond to real words in French (the words "pas", "tas", "cas", "bas", "gars", "fa", "sa", "chat", "va", "la", "ma", "rat" do exist in French). This may explain in part while the non sense monosyllables "za" and "ja" were recognized as "sa" and "chat" ("his/her", and "cat" in French). French has a very large number of words CV, and this fact may cause a bias in most studies pretending to use non-sense CV words.
There are strong restrictions in our study, we are well aware of. Voicing depends on the position of the consonant in the word. First, the initial consonants are known to be more resistant to changes than the middle or final ones in French [10], as exemplified by sound changes from Latin to French [6]. The gesture for maintaining voicing in medial or final consonant are different than the gesture of starting voicing in initial position. The default value for voicing is position dependent and length dependent. Voicing is favored in intervocalic position, if the consonant is short, and long voiced stops tend to be devoiced in all positions. Devoicing is favored in initial and final position, whatever the duration of the consonants. Our results are restricted to initial position, which favors devoicing. Intervocalic position probably would modify the results. Second, the identification test is done on monosyllables, after a silence. Starting or stopping voicing and devoicing in continuous speech may require very fine and rapid adjustments in voicing that may be difficult to produce for such patients. All these aspects should be investigated before reaching definitive conclusions.
Third, only one vowel was selected. We chose the /a/ vowel as it is the most frequent in French. As well, we did not intend to analyze vowel intelligibility so we chose the /a/ vowel as it is the easiest to pronounce by our patients. The confusion matrix for the consonants followed by a larger variety of vowels should be different. As a matter of fact, the voice onset time (VOT) is the shortest in front of the open /a/ vowel. Close vowels lead to a higher oral pressure, and there is more difficulty starting voicing before a close vowel, such as /i/, /y/ and /u/. As well, the transitions also vary as a function of the vowel. For example, formant transitions contain all the necessary information for distinguishing the place of

articulation of the stops before /a/, but the burst is necessary for the distinction between /ti/ and /ki/, or /pu/ and /gu/.
The restrictions on the length of our corpus is due to the fact that it is difficult to ask the patients to read a too large number of syllables. Such problem is a constant problem when working with patients. Participating to a phonetic test is not part of their therapy.

4.3. Voicing feature

The main confusion was on the voicing feature; the voiced consonants were perceived as devoiced contrarily to what has been observed with alaryngeal voices [7]. As noted by Ladefoged et al. [8], devoicing can be the result of laryngeal muscular tensions either by high position of the larynx with stiff vocal folds or by shortening of the pharyngeal cavity. After supracricoid partial laryngectomy, we can explain this phenomenon by the fact that the vocal tract is shortened approximately by 3 cm, and the neoglottis has heavier mass and slower mobility. Initiating the voicing can be difficult for a voiced consonant in the initial position. As noticed before, the test should be carried on intervocalic consonants, to test the ability to maintain voicing.

The most affected consonants are the palato-velar and to a lesser degree the bilabials. We may explain this by physiological compensatory behavior: to produce voicing, the patients must use the base of the tongue to help sphincterisation of the neolarynx. Instead of closing the vocal folds, the patients have to move forward the arytenoid cartilages and to move backward the base of the tongue with the remnant of the epiglottis. The tongue has a backward movement for initiating phonation which modifies articulatory targets. The place of articulation of the back consonants are more affected than the place of articulation of the front consonants. It is worth to note that the consonant /r/ is perfectly recognized.

4.4. Place of articulation

Contrary to what was expected, not only voicing distinction was modified but also place of articulation. Consonants were perceived as dento-alveolar instead of palato-velar or bilabial. It is not entirely sure that our listeners perceived a coronal instead of a velar. Alveolar position is often considered as the unmarked position. It may be the fact that the not identified consonants are labeled as alveolar, but default.
Our patients may also compensate by using this position when not able to use the correct articulation place. Further investigation is needed.

4.4. Jury experience

For perceptual evaluation, we noticed in a previous study [9] that the experience of the jury influenced the results when evaluating voice quality after supracricoid partial laryngectomy. The same observation was done by Kreiman [3] and Hammarberg [4] for laryngeal dysphonic voices. In this study, contrary to regular perceptual evaluation, the jury was not asked to evaluate the quality of voice but just transcribe what they heard. Our results did show some difference mainly in identification of place of articulation. The naive jury did more confusions than the expert one. To understand why is it so, more studies are needed and we should analyze the jury's reliability and the way they react to severe dysphonic voices.

5. CONCLUSIONS

Despite the limitations of this study, mainly due to the fact that it is not possible to ask the patients to read a too large corpus, we have found that they have difficulties in realizing the voicing distinctions as expected. We found also confusions in the place of articulation, as unexpected or at least, less expected. Further work should be done to understand the mechanisms of articulation confusion, which may be explored by palatography and MNR.

Analysis of intelligibility and voicing distinction in terms of phonetic dimensions can be useful for speech-language pathologists in treatment programs by improving phoneme production problems. The work on tongue position for voicing and for articulation is a very important part of the rehabilitation. The kinds of confusions attested by this study should be taken into account in the training program of these patients.

REFERENCES

[1] Laccourreye, H., Laccourreye, O., Weinstein, G., Menard, M., Brasnu, D. 1990. Supracricoid laryngectomy with cricohyoidoepiglottopaxy: a partial laryngeal procedure for glottic carcinoma. *Annals of Otology Rhinology Laryngology*, 99, 421-426.

[2] Crevier-Buchman, L., Laccourreye, O., Weinstein, G., Garcia, D., Jouffre, V., Brasnu, D. 1995. Evolution of speech and voice following supracricoid partial laryngectomy. *The Journal of Laryngology and Otology*. 109, 410-413.

[3] Kreiman, J., Gerratt, B., Precoda, K., Berke, G.S. 1992. Individual differences in voice quality perception. *Journal of Speech and Hearing Research*. 35, 512-520.

[4] Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J., Wedin, L. 1980. Perceptual and acoustic correlates of abnormal voice qualities *Acta Otolaryngologica*. 90, 441-451.

[5] Hirose, H. 1996. Voicing distinction in esophageal speech. *Acta Otolaryngologica*. Suppl. 524, 56-63.

[6] Bourciez, E., & Bourciez, J. 1995. *Phonétique française*. Klincksieck Ed., Paris.

[7] Miralles, J.L., & Cervera, T. 1995. Voice intelligibility in patients who have undergone laryngectomies. *Journal of Speech and Hearing Research*. 38, 564-571.

[8] Ladefoged, P., Maddieson, I., Jackson, M. 1988. Investigating phonation types in different languages. In O. Fujimura (Ed.), *Vocal fold physiology: Voice production, mechanisms and functions*. New-York: Raven Press.

[9] Crevier-Buchman, L. et al. 1995. Evaluation subjective de la voix et de la parole après laryngectomie partielle supracricoïdienne. *Revue de Laryngologie Otologie Rhinologie*. 116, 273-276.

[10] Vaissière, J., 1995, From Latin to Modern French: on diachronic changes and synchronic variations", AIPUK, 8, 479-488.

A	p	t	k	b	d	g	f	s	ch	v	z	j	l	m	n	r	% E
p	**96**		**3**	**1**													4
t		**86**			**4**								**10**				14
k		**14**	**79**		**3**	**3**											21
b	**31**			**67**	**2**												33
d		**29**			**71**												29
g		**17**	**31**		**3**	**49**											51
f							**86**			**14**							14
s								**99**			**1**						1
ch									**100**								0
v										**92**				**8**			8
z								**20**			**80**						20
j			**2**						**14**			**80**					20
l													**100**				0
m														**93**	**7**		7
n															**100**		0
r																**100**	0
	127	**146**	**116**	68	84	52	86	**119**	**114**	**107**	81	80	**110**	**101**	**107**	100	

Table 1: Confusion matrix, trained listeners (speech therapists)

N	p	t	k	b	d	g	f	s	ch	v	z	j	l	m	n	r	% E
p	**97**	**2**		**1**													3
t		**78**			**12**								**10**				22
k		**19**	**76**		**1**	**4**											24
b	**23**	**1**		**72**	**3**												28
d		**29**			**71**												29
g		**11**	**40**		**4**	**44**											56
f							**92**	**4**		**3**							8
s								**97**			**3**						3
ch								**1**	**97**			**2**					3
v							**3**			**87**	**10**						13
z								**20**			**73**	**7**					27
j									**20**			**80**					20
l													**100**				0
m														**86**	**14**		14
n														**8**	**92**		8
r																**100**	0
	120	**140**	**116**	73	92	49	96	**122**	**117**	90	87	89	**110**	93	**107**	100	

Table 2: Confusion matrix, untrained listeners