Speech quality after major surgery of the oral cavity and oropharynx with microvascular soft tissue reconstruction

Irma M Verdonck-de Leeuw [1], *Louis ten Bosch*[2], *Li Ying Cao*[2], *Rico NPM Rinkel*[1], *Pepijn A Borggreven*[1], *Lou Boves*[2], *C René Leemans*[1]

[1] Department of Otolaryngology – Head & Neck Surgery, VU University Medical Center, Amsterdam, The Netherlands
[2] Department of Language and Speech, University of Nijmegen, Nijmegen, The Netherlands
im.verdonck@vumc.nl

10.21437/Interspeech.2007-383

Abstract

Speech quality of patients with oral or oropharyngeal carcinoma was assessed by perceptual and acoustic-phonetic analyses. Speech recordings of running speech of patients before and 6 and 12 months after treatment for oral or oropharyngeal cancer and of 18 control speakers were evaluated regarding intelligibility, nasality and articulation, which revealed deteriorated speech in 20% of the patients before treatment, and in 75% 6-12 months after treatment. Acoustic analyses comprised formant, duration, perturbation and noise measures of the vowels /i/, /a/, and /u/ and were performed on the speech samples 6 months after treatment and the controls. Patients appeared to have a smaller vowel space compared to controls, which was clearly related to speech intelligibility. Furthermore, voice perturbation appeared to be higher in patients. Although oropharyngeal treatment does not effect the function of the larynx itself, the acoustic coupling between source and filter may effect the smoothness of the voicing characteristics. The presented speech analyses may serve as part of an outcome measurement protocol for assessing efficacy of speech rehabilitation.

Index Terms: head and neck cancer, major surgery, reconstruction, voice and speech quality, intelligibility, formant analyses, perturbation, HNR

1. Introduction

Head and neck cancer and its treatment has a distinct impact on daily life. In addition to general complaints such as pain and fatigue, head and neck cancer patients are often confronted with changes in facial appearance, swallowing and voice and speech problems. Longitudinal data reveal that quality of life after treatment only gradually improves during the first year, but swallowing and speech difficulties continue to exist [1]. Patients with advanced oral and oropharyngeal cancer are especially prone to speech difficulties. Speech outcome is dependent on residual mobility of structures in the oral cavity and oropharynx. The past 20 years have shown an improvement in the technical possibilities of replacing ablated tissues in the oral cavity and oropharynx by regional or distant flaps. Free fasciocutaneous flaps, such as the radial forearm flap (RFFF), have become the preferred method of reconstruction for larger soft tissue defects in the oral cavity and oropharynx because of their reliability and improved functional characteristics of dynamic structures such as the tongue and pharynx. Speech outcome can be assessed by using indicators of speech production (oral function and articulation tests, aerodynamic and acoustical analyses), speech perception (intelligibility and acceptability), and self-reported speech adequacy in every-day-life situations (questionnaires). There are numerous methodological differences between studies on speech quality of patients treated for oral or oropharyngeal cancer [2-7]. Nevertheless, it can be concluded that speech difficulties are highly dependent on tumor size and site. Expectantly, patients undergoing resection of larger tumors have more speech difficulties. After resection of oral carcinomas, patients encounter articulation problems due to tissue loss, structure alteration, or tongue mobility impairment. Target sounds may be distorted, substituted, or omitted leading to decreased intelligibility. Speech production problems of patients with oropharyngeal defects include nasal resonance problems due to velopharyngeal inadequacy. In case of tissue loss or mobility impairment, air will escape through the nose, vowels sound nasal and insufficient pressure can be build up in the oral cavity to produce stops and fricatives. In case of continued velopharyngeal closure, the air stream cannot escape through the nose and the nasal consonants are denasalized.

The aim of this study is to investigate speech outcome in a well-defined head and neck cancer patient population surgically treated and reconstructed with up-to-date methods, in order to obtain insight in associated speech difficulties. Our long-term goal is the improvement of the efficacy of speech rehabilitation by applying specch analysis techniques in combination with perceptual ratings from experts.

2. Methods

2.1. Speakers

Seventy-nine patients who underwent composite resections for advanced oral or oropharyngeal squamous cell carcinoma with microvascular soft tissue transfer for the reconstruction of their surgical defects were included in the study after written informed consent. The study population was treated at the department of Otolaryngology-Head and Neck Surgery of the VU University Medical Center. Exclusion criteria were age more than 75 years or inability to participate in functional tests. Patients were operated by means of composite resections including excision of the primary tumor with en bloc ipsilateral or bilateral neck dissection. If the tumor encroached on the mandible, a marginal mandibulectomy was performed transorally or by using a cheek flap. In oropharyngeal carcinomas a paramedian mandibular swing approach was used. All free flaps were successful. Indications for postoperative radiotherapy included the most severe (T3-4) tumors, positive surgical margins, perineural tumor spread, multiple positive nodes or extranodal spread. Generally, the clinical target volume of the initial field included the entire surgical bed. The primary tumor area and neck nodes were irradiated using 2 Gy per fraction to a dose of 46 Gy. An additional boost was given at the primary site up to a total dose of 56 Gy (2 Gy per fraction, 5 times/week). In case of positive surgical margins or extranodal spread an additional boost was given to a total dose of 66 Gy (2 Gy per fraction, 5 times/week).

Collection of speech data for patients was performed at three points in time; before treatment, 6 months after treatment and 12 months after treatment. Identical speech tests were performed for an age and gender matched control group of 18 persons (controls).

2.2. Speech assessment

Speech recordings of a read aloud text were performed in a sound-treated room and digitized using Cool Edit PRO 1.2 (Adobe Systems Incorporated, San Jose, California, USA) with 22 kHz sample frequency and 16-bit resolution. Recording level was adjusted for each speaker to optimize signal-to-noise ratio. All recordings were made with a mouth to microphone distance of 30 cm. A computerized program was developed to perform blinded randomized speech evaluation. Overall intelligibility (judged by a panel of 2 trained speech therapists) was assessed on a 10-points scale (ranging from poor to excellent: the 10-point grading scale is commonly used in the Dutch educational system in which 5 or less is judged as insufficient and 6 or more as sufficient). Interrater reliability scores (Cronbach's alpha) on intelligibility was high: 0.86. Intrarater agreement (percentage within one scale value between the first and second, repeated speech fragment) were equally high, ranging from 40-90%. To obtain more insight in the cause of decreased intelligibility, evaluation of the quality of articulation and nasal resonance was performed by the same panel of speech therapists agreeing consensus on a 4-point scale ranging from normal to increasing deviant. Intrarater agreement was high with 100% equal scores between the ratings on the first and second, repeated speech fragments on articulation and nasality.

Speech analyses can be carried out on different types of stimuli, such as whole speech fragments or segmented vowel portions. Vowel portions are relatively simple to identify in the speech signal. Moreover, vowels are usually easier to analyse acoustically. On the other hand, vowels are not representative for continuous speech and for example voice onset problems may be underrepresented in this sort of material. Despite the disadvantages, often vowels are chosen due to the complexity of detailed analyses on continuous speech (Umapathy et al, 2005). In the present study, formant values (F1, F2, F3), durations, jitter, shimmer and harmonic-to-noise ratios (HNR) were measured on vowels. It is expected that formant values for patients with treatment in the oral cavity differ from those for control speakers. These effects are assumed to be maximum for the cardinal vowels /a/, /i/ en /u/. With respect to duration, it is known that certain types of pathological speech are less fluent and characterised by longer phoneme durations [8]. All vowels were segmented manually using the speech analysis program PRAAT (www.praat.org). Per speaker nine vowels were selected (3 /a/, 3 /i/ and 3 /u/). For a few speakers, only 2 /u/s were available in the speech material.

2.3. Statistical analyses

To test multiple group differences, groups in time (controls, patients tested before treatment and patients tested 6 and 12 months after treatment) and groups regarding tumor site (mobile tongue, floor-of-mouth, retromolare trigone, tonsil, base-of-tongue and soft palate) analyses of variance (ANOVA F-test) were carried out regarding intelligibility; in case of significant F-tests, posthoc tests were performed to test which groups differed from each other. F-tests were used to identify significant differences between pathological and control speakers regarding the acoustical-phonetic speech variables. Kruskal-Wallis H-tests were used to test group differences regarding articulation and nasality scores. Independent t-tests (intelligibility) and Mann-Whitney U-tests (articulation, nasality) were performed to determine the impact of tumor stage (T2 versus T3-4). To investigate relations between overall intelligibility and detailed speech outcome (articulation and nasal resonance scores, and acoustic-phonetic analyses), Spearman's correlation rho (r) coefficients were calculated. On patients who were tested at all points of time (i.e., before, and 6 and 12 months after treatment), comparable paired tests (ANOVA F-test with repeated measures, Friedman chi-square test) were performed.

3. Results

3.1. Patient characteristics

Patient characteristics are shown in table 1. For a total of 76 patients tested before treatment speech tests could be analyzed, while for patients tested 6 and 12 months after treatment this was possible for 51 and 42 patients respectively. For 38 patients speech tests could be analyzed on all points in time (i.e., patients tested before, and 6 and 12 months after treatment). Drop-out was caused by tumor recurrence, distant metastases, death, patients refusal, due to technical problems, or by patients lost to follow-up.

Table 1. Characteristics of 79 patients included in the study

	n	(%)
Gender		
Male	46	(58)
Female	33	(42)
Tumor site		
Oral cavity	37	(47)
Mobile tongue	18	(23)
Floor-of-mouth	15	(19)
Retromolare trigone	4	(5)
Oropharynx	42	(53)
Tonsil	24	(30)
Base-of-tongue	11	(14)
Soft palate	7	(9)
T		
2	35	(44)
3	41	(52)
4	3	(4)

3.2. Speech evaluation

Based on intelligibility scores, none of the control speakers was deviant, while 17% of the patients before treatment, and 71% of the patients after treatment had deviant scores. Analyses of variance revealed significant differences regarding intelligibility (F=46.57, p<0.01). Posthoc tests showed significant differences (p<0.05) between controls on the one hand and patients tested before, and 6 and 12 months after treatment on the other hand (figure 1). The differences between patients tested before treatment on the one hand and patients tested 6 and 12 months after treatment on the other hand were statistically significant. No significant differences were found between patients tested 6 and 12 months after treatment. Repeated analyses of variance on the 38 patients who underwent assessment at all points in time, revealed similar results: patients tested before, and 6 and 12 months after treatment were significantly less intelligible (F=55.79, p<0.01) compared to controls, patients tested before treatment were significantly better than patients tested 6 and 12 months after treatment, and no significant differences were found between patients tested 6 and 12 months after treatment.

Regarding tumor stage, patients were divided into patients having T2 tumors and patients having T3-4 tumors. Before treatment, 32 patients (42%) had T2 tumors and 44 patients (58%) T3-4 tumors. Six months after treatment this was 26 (51%) and 25 (49%) and 12 months after treatment this was 24 (57%) and 18 (43%) respectively. A significant difference between patients with T2 tumors and patients with T3-4 tumors was found, before treatment (t=2.68, p<0.01), 6 months after treatment (t=3.03, p<0.01) and 12 months after treatment (t=3.41, p<0.01).

Regarding tumor site, significant differences were found for intelligibility in patients tested before treatment (F=2.67, p<0.05). Posthoc tests revealed that, before treatment, intelligibility was significantly worse in patients with tumors of the mobile tongue (mean intelligibility score 5.72) compared to patients with tumors of the base-of-tongue (mean intelligibility score 7.22). For patients tested 6 and 12 months after treatment, no significant differences between tumor sites were found regarding intelligibility.

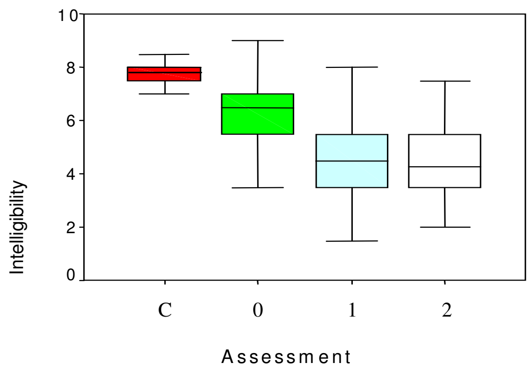

Figure 1. Box-plots of intelligibility scores for controls (C), patients before treatment (0) and 6 months (1) and 12 months (2) after treatment .

To obtain more insight in the causes of deteriorated intelligibility of patients, detailed evaluation was carried out on articulation and nasal resonance. Significant differences between controls and patients tested before, and 6 and 12 months after treatment were found for both articulation (H=62.63, p<0.01) and nasality (H=47.75, p<0.01). Regarding articulation, 94% of controls had normal scores, while for patients tested before treatment this was 63%, for those tested 6 months after treatment this was 14% and for patients tested 12 months after treatment this was 24%. Regarding nasality, no controls showed deviant scores, while 25% of patients tested before treatment, 67% of patients tested 6 months after treatment and also 67% tested 12 months after treatment showed deviant nasality scores. The same trend was found in the 38 patients who underwent assessment at all points in time regarding both articulation (Friedman chi-square ≥ 44.77, p<0.01) and nasal resonance (Friedman chi-square ≥ 25.26, p<0.01).

Articulation was significantly worse for patients with T3-4 tumors than for patients with T2 tumors tested before treatment (U=466, p<0.01) and tested 6 months after treatment (U=183, p<0.01), but not for those tested 12 months after treatment (U=144, p=0.06). No significant differences appeared between patients with T2 versus T3-4 cancer regarding nasal resonance. Regarding tumor site, no statistically significant results were found on articulation. Nasality scores revealed significant differences (H=14.36, p<0.01) for patients tested 6 months after treatment. Patients with floor-of-mouth tumors showed the best nasality scores, while patients with tonsil or soft palate tumors showed the worst overall scores.

Spearman correlations between intelligibility on the one hand, and articulation and nasal resonance on the other (all evaluated by the same panel of trained raters) revealed that intelligibility is more prominently correlated to articulation scores than to nasality scores for patients tested before treatment (r=0.64, r=0.44, respectively), tested 6 months after treatment (r=0.68, r=0.36, respectively) and tested 12 months after treatment (r=0.69, r=0.45, respectively).

3.3. Speech analyses

Between the patient and control group, a number of differences were significant. Both the male and the female patients had a lower F2 than the corresponding control speakers for /a/ and /i/, but the reverse was true for /u/. This means that for the patients the vowel triangle is shifted towards the closed-back position and has a smaller area (see figure 2) Moreover, for patients (P) vowels had a longer duration than for the control speakers (C). This was true for all vowels (/a/: P: 114 ms, sd = 29, C: 100 ms, sd = 32; /i/: P: 87 ms, sd = 29, C: 73 ms, sd = 22; /u/: P: 72 ms, sd = 21, C: 53 ms). In our study, patients did not use a significantly lower speech rate (P: 171.4 words/minute, sd = 32.6; C: 174.8 words/minute, sd = 22.5). Finally, patients did have a significantly higher shimmer than control speakers (Table 2).

Table 2. Significant differences between parameters measured on vowels from pathological and control speakers.

	Pathological vs. control speakers
Speech	
F1	-
F2	F = 9.9, p =.000
Duur	F = 29.5, p = .000
Voice	
Jitter	-
Shimmer	F = 8.1, p = .005
HNR	-

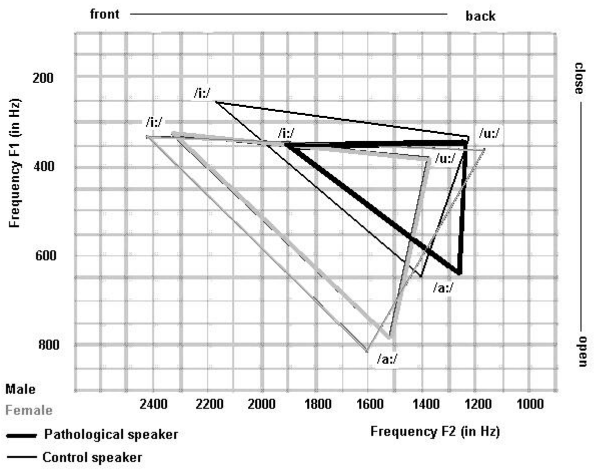

Figure 2. Position in the F1-F2 plane of the three vowel /i/, /a/, /u/ for male and female pathological and control speakers

4. Discussion

This study presents an inventory of speech performance before and after treatment in a well-defined head and neck cancer patient group. Overall intelligibility in patients before treatment was approximately 20% worse than in controls, while 6-12 months after treatment this was approximately 75%. No evident improvement was seen between 6 and 12 months after treatment. Worse postoperative overall speech quality was demonstrated for patients with larger tumors compared to patients with smaller tumors. These speech evaluations were clearly related to acoustic-phonetic analyses performed on the speech data 6 months after treatment. Although the speech data per patient were limited and recorded in realistic medical context, the analyses performed clearly suggest that pathological and control speakers differ in a number of acoustic characteristics. In this study, we focussed on vowels and looked at formants, jitter, shimmer and SNR in particular. It appears that both speaker groups can be differentiated on the basis of formants (F2). The vowel space of pathological speakers appeared smaller than the space spanned by control speakers, suggesting that the articulatory freedom in the vocal tract is reduced after oncological treatment. There was also a significant difference found in vowel duration. However, these duration measurements have not been normalised for speaking rate, and so may be an artefact of the chosen speaking style other than a direct consequence of mechanical articulation problems that are related to the oncological treatment. Our pathological speakers had a significantly higher shimmer level than the control speakers. Although oropharyngeal treatment does not effect the function of the larynx itself, the acoustic coupling between source (vocal folds) and filter (vocal plus nasal tract) may effect the smoothness of the voicing characteristics.

Speech deterioration after treatment is caused by the inevitable anatomical and functional alterations of surgery and radiotherapy. Currently, optimal functional results are observed with the application of free flaps and, in the oral cavity and oropharynx, the best outcome is obtained by the use of thin pliable fasciocutaneous flaps, such as the RFFF and anterolateral thigh flap [8]. The possible effect of speech rehabilitation is a matter of consideration in outcome analyses such as the present. Although speech rehabilitation can be beneficial in improving speech intelligibility, reference rates in clinical practice are low as was also found in the present study (only 2 patients received speech rehabilitation on their own request). Information on efficacy of speech rehabilitation is scarce and further evidence on the efficacy of speech rehabilitation is therefore clearly needed.

5. Conclusion

Speech quality of patients with oral or oropharyngeal carcinoma as assessed by perceptual and acoustic-phonetic analyses is deteriorated in 20% of the patients before treatment, and in 75% 6-12 months after treatment. The presented speech analyses may serve as part of an outcome measurement protocol for assessing efficacy of speech rehabilitation. Further development of a multidimensional speech outcome measurement protocol including consonant analyses and self-ratings of speech problems in daily life is ongoing [9].

6. Acknowledgement

Thanks to Toni Rietveld for advice on the statistical methods.

7. References

[1] Borggreven, P.A., Aaronson, N.K., Verdonck-de Leeuw, I.M., Muller, M.J., Heiligers, M.L.C.H., de Bree, R., Langendijk, J.A., Leemans, C.R. "Quality of life after surgical treatment for oral and oropharyngeal cancer: a prospective longitudinal assessment of patients reconstructed by a microvascular flap oral oncology", Oral Oncol., epub, 2007.

[2] Pauloski, B.R., Logemann, J.A., Rademaker, A.W., et al. "Speech and swallowing function after anterior tongue and floor of mouth resection with distal flap reconstruction", J. Speech Hear. Res., 36:267-276, 1993.

[3] Greven, A.J., Meijer, M.F., Tiwari, R.M., "Articulation after total glossectomy: a clinical study of speech in six patients", Eur. J. Disord. Commun., 29:85-93, 1994.

[4] Furia, C.L., Kowalski, L.P., Latorre, M.R., et al. "Speech intelligibility after glossectomy and speech rehabilitation", Arch. Otolaryngol. Head Neck Surg., 127:877-883, 2001.

[5] McKinstry, A., Perry, A., "Evaluation of speech in people with head and neck cancer: a pilot study", Int. J. Lang. Commun. Disord., 38:31-46, 2003.

[6] Borggreven, P.A., Verdonck-de Leeuw, I.M., Langendijk, J.A., Doornaert, P., Koster, M.N., de Bree, R., Leemans, C.R., "Speech outcome after surgical treatment for oral and oropharyngeal cancer: a longitudinal assessment of patients reconstructed by a microvascular flap", Head Neck, 27(9):785-93, 2005.

[7] Umapathy, K., Krishnan, S., Parsa V., and Jameison, D., "Discrimination of pathological voices using a time-frequency approach", IEEE Trans. Biomedical Eng., pp 421-430, Vol. 52(3), March 2005

[8] Su, W.F., Hsia, Y.J., Chang, Y.C., Chen, S.G., Sheng, H., "Functional comparison after reconstruction with a radial forearm free flap or a pectoralis major flap for cancer of the tongue", Otolaryngol. Head Neck Surg., 128:412-418, 2003

[9] Rinkel, R.N., Verdonck-de Leeuw, I.M., van Reij, E.J., Aaronson, N.K., Leemans, C.R., "The speech handicap index in oral and pharyngeal cancer patients: better understanding of patients' complaints", submitted.