SPEECH OUTCOME AFTER SURGICAL TREATMENT FOR ORAL AND OROPHARYNGEAL CANCER: A LONGITUDINAL ASSESSMENT OF PATIENTS RECONSTRUCTED BY A MICROVASCULAR FLAP

Pepijn A. Borggreven, MD,1 Irma Verdonck-de Leeuw, SLP, PhD,1 Johannes A. Langendijk, MD, PhD,2 Patricia Doornaert, MD,2 Marike N. Koster, MA, SLP,3* Remco de Bree, MD, PhD,1 C. René Leemans, MD, PhD1

1 Department of Otolaryngology–Head and Neck Surgery, VU University Medical Center, P.O. Box 7057, 1007 MB, Amsterdam, The Netherlands. E-mail: pa.borggreven@vumc.nl
2 Department of Radiation Oncology, VU University Medical Center, Amsterdam, The Netherlands
3 Utrecht Institute of Linguistics OTS, University of Utrecht, Utrecht, The Netherlands

Accepted 11 March 2005
Published online 10 June 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hed.20236

Abstract: Background. The aim of the study was to analyze speech outcome for patients with advanced oral/oropharyngeal cancer treated with reconstructive surgery and adjuvant radiotherapy.

Methods. Speech tests (communicative suitability, intelligibility, articulation, nasality, and consonant errors) were performed in a control group and in patients before treatment (n = 76), and 6 months (n = 51) and 12 months (n = 42) after treatment.

Results. Speech tests were significantly worse for patients before and after treatment compared with the controls. Speech did not improve between 6 and 12 months. After treatment, patients with T3–4 tumors showed a significantly worse score for communicative suitability, intelligibility, and articulation than patients with T2 tumors. No significant differences were found for subsites after treatment, although patients with mobile tongue tumors showed the best results.

Conclusion. Speech difficulties are significant, and with the knowledge of this study better counseling and vigilance as to speech difficulties may be possible in patients undergoing treatment for oral/oropharyngeal cancer. © 2005 Wiley Periodicals, Inc. Head Neck 27: 785–793, 2005

Keywords: head and neck cancer; oral cavity; oropharynx; microvascular reconstruction; speech

Head and neck cancer and its treatment may negatively affect the patient’s functional status. Shortly after treatment, a significant number of patients encounter speech deterioration, swallowing limitations, facial appearance changes, and psychological problems.1–5 Longitudinal data reveal that quality of life after treatment only gradually improves during the first year,6–8 but swallowing and speech difficulties continue to exist.9–11 Patients with oral and oropharyngeal cancer are especially prone to speech difficulties.12
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 free 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.13–15

Speech outcome is assessed by the use of indicators of speech production (oral function and articulation tests, aerodynamic and acoustical analyses), speech perception (intelligibility and acceptability), and self-reported speech adequacy in everyday life situations (questionnaires). Numerous methodologic differences exist between studies on speech quality of patients treated for oral or oropharyngeal cancer.16–20 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 because of 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 because of velopharyngeal inadequacy. In the case of tissue loss or mobility impairment, air will escape through the nose, vowels sound nasal, and insufficient pressure can be built up in the oral cavity to produce stops and fricatives. In the case of continued velopharyngeal closure, the air stream cannot escape through the nose, and the nasal consonants are denasalized.

The aim of this study was to investigate speech outcome by means of a multidimensional speech assessment protocol in a well-defined patient population with head and neck cancer reconstructed with up-to-date methods to obtain insight into speech difficulties in relation to tumor classification and site.

MATERIALS AND METHODS

Speakers. Seventy-nine patients who underwent composite resections for advanced oral or oropharyngeal squamous cell carcinoma with microvascular soft tissue transfer (ie, RFFF) for the reconstruction of their surgical defects, operated on between January 1998 and December 2001, 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 older than 75 years or inability to participate in functional tests. Patients were operated on 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 use of a cheek flap. In patients with oropharyngeal carcinomas, a para-median mandibular swing approach was used. All free flaps were successful. Indications for postoperative radiotherapy included 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 with 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 the 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).

Patient data (age and sex), tumor data (TNM stage21 and tumor site), and treatment data (postoperative radiotherapy and rehabilitation) were registered. 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-matched and sex-matched control group of 18 persons (controls).

Speech Assessment. Speech analyses were performed according to a standardized speech assessment protocol. Speech recordings of a standardized read-aloud text and a subset of Dutch consonants were performed in a sound-treated room and digitized using Cool Edit PRO 1.2 (Adobe Systems Incorporated, San Jose, CA) 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.
Communicative suitability is defined as the speaking-dependent adequacy of speech as judged by naive listeners. Communicative suitability (judged by a panel of 13 naive listeners) and overall intelligibility (judged by a panel of two trained speech therapists) was assessed on a 10-point 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 communicative suitability and intelligibility were high, 0.98 and 0.86, respectively. Intrarater agreement (percentage within one scale value between the first and second repeated speech fragment) were equally high, ranging from 50% to 100% for the naive raters on communicative suitability and ranging from 40% to 90% for the experts on intelligibility.

To obtain more insight into 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. Furthermore, consonant error rate was assessed. Patients were asked to repeat five times the consonant-vowel (CV) syllables ta, da, na, sa, xa, ka, and la. In a computerized listening experiment, all Dutch consonants were presented on a computer screen. The CV syllables with the seven target consonants were presented in random order to a panel of five naive raters, who were asked to click with the mouse on the consonant they recognized in each CV syllable. Consonant error rate was assessed by the percentage incorrect target consonant.

### 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 on intelligibility and communicative suitability; in case of significant F tests, post hoc tests were performed to test which groups differed from each other. Kruskal–Wallis H tests were used to test group differences regarding articulation and nasality scores. Independent t tests (intelligibility and communicative suitability) and Mann–Whitney U tests (articulation, nasality) were performed to determine the impact of tumor classification (T2 vs T3–4). The influence of tumor classification and tumor site on consonant error rate was tested by chi-square tests. To investigate relations between overall intelligibility and detailed speech outcome (articulation and nasal resonance scores), Spearman’s correlation rho (r) coefficients were calculated. On patients who were tested at all points of time (ie, before and 6 and 12 months after treatment), comparable paired tests (ANOVA F test with repeated measures, Friedman chi-square test) were performed.

### RESULTS

#### Patient Characteristics

Patient characteristics are shown in Table 1. The ages of the 79 patients included in this study ranged from 23 to 74 years (mean, 56 years). In most patients (73%), the tumor did not extend over the median line. One patient was operated on for a recurrent tumor after prior transoral excision (1 year before) in which the rTNM stage was used, and one patient received previous radiotherapy (3 years before) for a neck node with an unknown primary tumor. Four patients had a synchronous second primary cancer.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>46 (58)</td>
</tr>
<tr>
<td>Female</td>
<td>33 (42)</td>
</tr>
<tr>
<td>Tumor site</td>
<td></td>
</tr>
<tr>
<td>Oral cavity</td>
<td>37 (47)</td>
</tr>
<tr>
<td>Mobile tongue</td>
<td>18 (23)</td>
</tr>
<tr>
<td>Floor of mouth</td>
<td>15 (19)</td>
</tr>
<tr>
<td>Retromolare trigone</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Oropharynx</td>
<td>42 (53)</td>
</tr>
<tr>
<td>Tonsil</td>
<td>24 (30)</td>
</tr>
<tr>
<td>Base of tongue</td>
<td>11 (14)</td>
</tr>
<tr>
<td>Soft palate</td>
<td>7 (9)</td>
</tr>
<tr>
<td>T classification</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>35 (44)</td>
</tr>
<tr>
<td>3</td>
<td>41 (52)</td>
</tr>
<tr>
<td>4</td>
<td>3 (4)</td>
</tr>
<tr>
<td>N classification</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>24 (30)</td>
</tr>
<tr>
<td>1</td>
<td>15 (19)</td>
</tr>
<tr>
<td>2a</td>
<td>2 (3)</td>
</tr>
<tr>
<td>2b</td>
<td>30 (38)</td>
</tr>
<tr>
<td>2c</td>
<td>6 (7)</td>
</tr>
<tr>
<td>3</td>
<td>2 (3)</td>
</tr>
</tbody>
</table>
tumor, and in these, the T classification of the largest tumor was used. In total, 73 patients (92%) received postoperative radiotherapy. Only two patients were rehabilitated by speech therapy, which precluded statistical analyses.

Speech tests could be analyzed for a total of 76 patients tested before treatment and for 51 and 42 patients tested 6 months and 12 months after treatment, respectively. For 38 patients, speech tests could be analyzed on all points in time (ie, patients tested before and 6 and 12 months after treatment). After treatment, there were no patients with retromolare trigone tumors. Dropout was caused by tumor recurrence, distant metastases, death, patient refusal, technical problems, or patients lost to follow-up.

**Overall Speech Outcome.** On the basis of intelligibility scores, none of the control speakers were deviant, whereas 17% of the patients before treatment and 71% of the patients after treatment had deviant scores. ANOVA results revealed significant differences regarding communicative suitability ($F = 51.26$, $p < .01$) and intelligibility ($F = 46.57$, $p < .01$). Post hoc tests showed significant differences ($p < .05$) between controls on the one hand and patients tested before and 6 and 12 months after treatment on the other hand, both regarding communicative suitability and intelligibility (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 for both communicative suitability and intelligibility. No significant differences were found between patients tested 6 and 12 months after treatment. Repeated ANOVA 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 communicative suitable ($F = 36.23$, $p < .01$) and less intelligible ($F = 55.79$, $p < .01$) than 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 classification, 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%) had 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 on communicative suitability and intelligibility (Figure 2), before treatment ($t = 2.03$, $p < .05$ and $t = 2.68$, $p < .01$, respectively), 6 months after treatment ($t = 3.84$, $p < .01$ and $t = 3.03$, $p < .01$, respectively), and 12 months after treatment ($t = 3.22$, $p < .01$ and $t = 3.41$, $p < .01$, respectively).

Regarding tumor site, there were no significant differences between patients in communicative suitability tested before and 6 and 12 months after treatment, but statistically significant differences were found for intelligibility in patients tested before treatment ($F = 2.67$, $p < .05$; Figure 3). Post hoc tests revealed that, before treatment, intelligibility was significantly worse...
in patients with tumors of the mobile tongue (mean intelligibility score, 5.72) than in 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, but patients with mobile tongue tumors showed the best scores on both communicative suitability and intelligibility.

**Detailed Speech Outcome.** To obtain more insight into the causes of deviant communicative suitability and intelligibility of patients, a 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 (\( H = 62.63, \ p < .01 \)) and nasality (\( H = 47.75, \ p < .01 \)) (Figure 4). Regarding articulation, 94% of controls had normal scores, whereas for patients tested before treatment, this was 63% and for those tested 6 months after treatment, this was 14%; for patients tested 12 months after treatment, this was 24%. Regarding nasality, no controls showed deviant scores, whereas 25% of patients tested before treatment, 67% of patients tested 6 months after treatment, and 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 \( \geq 44.77, \ p < .01 \)) and nasal resonance (Friedman chi-square \( \geq 25.26, \ p < .01 \)).

Articulation was significantly worse for patients with T3–4 tumors than for patients with T2 tumors tested before treatment (\( U = 466, \ p < .01 \)) and tested 6 months after treatment (\( U = 183, \ p < .01 \)) but not for those tested 12 months after treatment (\( U = 144, \ p = .06 \)). 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 (\( H = 14.36, \ p < .01 \)) for patients tested 6 months after treatment. Patients with floor-of-mouth tumors showed the best nasality scores, whereas 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 articu-
lation 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).

To obtain more insight in the rate and type of articulation problems, consonant errors were assessed. Comparisons were made between controls and patients after treatment. For the patients, the latest assessment was chosen, and, in total, 52 measurements of patients were analyzed. Consonant error rate seemed to be low for controls (2% error rate) and significantly higher for patients (17% error rate) (chi-square = 91.9, $p < .01$). Patients with T3–4 tumors showed significantly more consonant errors (25%) than patients with T2 tumors (9%) (chi-square = 78.67, $p < .01$). Tumor site seemed to affect consonant error rate (chi-square = 9.79, $p < .05$). Patients with base-of-tongue tumors made the fewest consonant errors (11%), whereas patients with tonsil tumors made the most consonant errors (19%) compared with patients with other tumor sites (17%). The type of consonant errors seemed to be diffuse, but a few clear observations can be extracted. The target consonants $k$ (velair), and $s, d,$ and $t$ (alveolair) were the most difficult consonants to produce correctly. The velair $k$ was often confused with the velair $g$, which was observed in all patient groups, except for patients treated for base-of-tongue tumors. The alveolar consonants were often nasalized ($d$ confused with $n$) in the case of patients treated for oropharyngeal tumors (tonsil or soft palate) or retracted ($s$ confused with $sj$ as in sheep, $t$ confused with $tj$ as in chill) in patients treated for mobile tongue tumors.

**DISCUSSION**

The outcome of speech in patients with head and neck cancer undergoing extensive ablative and reconstructive surgery was investigated in a fairly homogeneous group of 79 patients. Speech
was examined on the level of communicative suitability and intelligibility, which were thought to be influenced by nasality, articulation, and consonant errors. Speech before treatment and speech outcome were assessed relative to tumor classification and site. As expected, all aspects of speech both before and after treatment turned out to be significantly worse for patients than controls. Furthermore, patients scored significantly more consonant errors than controls. McKinstry and Perry\textsuperscript{20} also found impaired speech in 20 patients with various head and neck cancers before treatment compared with a control group. This study also shows that speech was significantly worse 6 and 12 months after treatment than before treatment. This effect after treatment is caused by the inevitable anatomic 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.\textsuperscript{23,24} Su et al\textsuperscript{25} reported better speech function after reconstruction with RFFF than with the more bulky pectoralis major flap. Intelligibility was also found to be improved for patients who underwent reconstruction by RFFF than those who underwent a somewhat thicker lateral upper arm flap as described by Hara et al.\textsuperscript{13} Seikaly et al\textsuperscript{15} reported that the use of the RFFF is presently the best reconstructive option, especially when the created defect takes up multiple anatomic sites.

Tumor classification and size play an important role in speech. In this study, patients with advanced primary tumors did significantly worse than patients with smaller tumors in the assessments before and after treatment concerning communicative suitability, intelligibility, articulation, and consonant errors. Speech was minimally impaired in patients with T2 tumors before treatment, a finding also shown by Schonweiler et al.\textsuperscript{26} These objective speech data are in concordance with patient’s self-assessment in which patients with smaller tumors assessed their speech before treatment significantly better than did those with larger tumors.\textsuperscript{27} Although Colangelo et al\textsuperscript{28} found no stage effect in the phase before treatment, a clear influence was observed after treatment, as shown in this study. However, irrespective of tumor classification, speech deteriorates after treatment and has no real tendency to improve after 12 months.

A statistical difference in communicative suitability as assessed by naive listeners between patients with different tumor subsites could not be established. Assessment before treatment by expert listeners, however, regarding intelligibility and articulation, revealed a clear difference between tumor subsites. On the two extremes of the spectrum were patients with mobile tongue and base-of-tongue tumors, the latter doing significantly better than the former. Speech in all its aspects deteriorates markedly after treatment for all subsites. Although it seemed that communicative suitability, intelligibility, and articulation were somewhat better after treatment for an oral cavity cancer than for an oropharyngeal cancer and site analyses pointed to distinct problems for different subsites, there were no statistical differences. Communicative suitability and intelligibility were most markedly preserved in patients who were treated for mobile tongue lesions, whereas articulation problems and nasality (significantly) were especially present in patients with tonsil and soft palate cancers. In contrast, Haughey et al.\textsuperscript{29} who investigated speech for 43 patients with tongue or floor-of-mouth carcinomas who also underwent fasciocutaneous flap reconstruction, found improved intelligibility for patients with base-of-tongue defects compared with those with tumors of the mobile tongue. This study, however, included 51% of patients who had undergone prior treatment (partial glossectomy, radiation therapy, or chemotherapy), which may have influenced their results. Colangelo et al\textsuperscript{28} reported significantly worse speech in T3-staged patients with oral cavity tumors than with those with cancer of the oropharynx, whereas this difference was absent in the same dichotomized site groups with T1–2 or T4 tumors. In this study, average results for the assessments before and (3 months) after treatment were analyzed.

Whereas, as expected, particularly patients with larger tumors made significantly more consonant errors than the control group or patients with smaller tumors, patients with base-of-tongue tumors showed the least consonant errors after treatment. Similar findings were reported by Haughey et al.\textsuperscript{29}.

When articulation scores after treatment were compared with consonant errors, no clear pattern could be found. The reason for this was unclear, but personal articulation strategies and tumor classification might influence both articulation and consonant errors. Analyses of type of
consonant errors showed that for all tumor sites, pronunciation of the \( k \) was the most abnormal. In a study by Pauloski et al,\textsuperscript{30} this was also reported, but most markedly for patients with base-of-tongue tumors.

It is obvious that in any patient group treated for cancer, longitudinal follow-up is often inadequate. Factors such as marked comorbidity, alcohol abuse, depression, and a relatively high recurrence rate are typical of patients with head and neck cancer, and a certain bias because of selective dropout seems, therefore, inevitable.\textsuperscript{31–34} In our patient cohort, the initial fraction of patients with T3–4 tumors was 58\%, whereas this dropped to 49\% and 43\% at 6 and 12 months, respectively. Likewise, patients with mobile tongue and floor-of-mouth tumors, as well as those who scored worse at baseline, dropped out more frequently. The possible effect of speech rehabilitation is another 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.\textsuperscript{17,18,20} This also counts for this study; only two patients received speech rehabilitation at their own request. Regarding the objective speech deterioration, it might be recommended to analyze the objective and subjective effect of speech rehabilitation in a future study. In the literature, though, information on efficacy of speech rehabilitation is scarce. Only one study was found demonstrating a positive effect of speech rehabilitation on 27 patients after glossectomy.\textsuperscript{19}

This study presents an inventory of speech performance before and after treatment in a well-defined patient group with head and neck cancer. Overall, speech quality in patients before treatment was approximately 20\% worse than in controls, whereas 6 to 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. Patients with mobile tongue tumors showed a worse overall speech quality than those with tumors at other subsites, albeit without significant difference. Speech rehabilitation might be beneficial for these patients, but further research on its efficacy is clearly needed. The results of this study could assist in improved patient counseling and vigilance as to speech difficulties in patients who are treated for oral cavity and oropharynx cancer.

Acknowledgment. The authors thank the control group patients and the panel of naive listeners for their contributions to this study.

REFERENCES


