Use of real-time MRI in assessment of consonant articulation before and after tongue surgery and tongue reconstruction

K. MÁDY[1], R. SADER[1], A. ZIMMERMANN[1], Ph. HOOLE[2], A. BEER[3], H. F. ZEILHOFER[1], Ch. HANNIG[3]
[1] *Department of Oral and Maxillofacial Surgery, University of Technology, Munich*
email: mady@mkg.med.tu-muenchen.de
[2] *Department of Phonetics and Linguistic Communication, Ludwig Maximilian University, Munich*
[3] *Department of Radiology, University of Technology, Munich*

Abstract.
The present paper describes first results on impaired articulation after partial glossectomy and oral reconstruction. The tongue movements are recorded by dynamic MRI (8 images per second) while the patients utter test words involving all lingual consonants of German. Along 13 lines of a semicircular grid, the tongue–palate distance is measured for all consonants, and pre- and postoperative movements are compared. To supplement the MR recordings, the speech is recorded acoustically. Eventually, the results will be compared with results from the video fluoroscopy technique, which has already proved to be useful in the evaluation of tongue and velum movements.

1. Introduction

The assessment of articulatory impairments following surgery to structures of the oral cavity in the course of cancer treatment is of considerable clinical relevance. While the surgeon has no options regarding location and amount of tissue to remove, there is some room for manoeuvre in the technique chosen for reconstructing: one possibility is to close a defect by fixing the tongue to the floor of the mouth, the second is to use a local (platysma) or a free (pectoral or radial) flap, or to leave the defect open to heal by itself. Much research has been done on evaluating the clinical advantages and disadvantages of the reconstruction methods, but the potential articulatory consequences of surgery are still not well-understood (Bressmann et al., 2000). In addition to immediate clinical relevance, the compensatory articulatory strategies employed by the patients to cope with gross structural changes of the oral structures and with impaired mobility (for example, when the tongue is attached to the floor of the mouth to cover a defect) are of great theoretical interest in the field of speech motor control.

For investigations in this area the new technique of real-time MRI appears promising for several reasons: (1) It covers the whole vocal tract from lips to larynx, which is important when it is difficult to determine a priori what compensatory behaviours may occur; (2) for patients with painful conditions in the oral region it is advantageous that the technique does not require sensors to be attached to the tongue; (3) it does not use ionizing radiation, so the amount of speech material that can be collected is not very severely restricted.

2. Methods and materials

The patients involved in our study suffer from cancer of the anterior and lateral oral cavity, the tongue and/or the floor of mouth being concerned. They had neither speech or language disorders before their current disease, nor any tumours in the head and neck area previously. The first recordings are made a few days before the operation, the second set is run approximately four weeks postoperatively, before the optional commencement of the irradiation therapy.

Acoustic recordings are being made prior to every session, firstly because MR imaging is accompanied by a strong noise that could only be eliminated by a special recording technique. (Different solutions are being tested currently, but a satisfying quality has not been achieved yet.) Secondly, the circumstances in the MR tomograph are hard to bear for some patients, especially after the operation, therefore we try to keep the recording time at a minimum. The acoustic signals are recorded and analysed by MultiSpeech 3700 (KAY Elemetrics). As a comparison, a video fluoroscopy recording follows each MR session, as this technique has already been used successfully for the evaluation of tongue movements (Sader, 1999). Because of the X-ray radiation in this method, again, the amount of data recorded on the video tape is as small as possible.

At the present state of our study, the possible maximum recording speed is 8 images per second, using a T1 fast gradient echo sequence on Philips ACS NT Gyroscan (TR = 4.0 ms, TE = 1.1 ms, fa = 6, slice thickness = 10 mm) with the new “sensitivity encoding” system (SENSE). With this feature combination the entire head and neck area can be viewed, while resolution and slice thickness are sufficient for gaining information about the

movement of all articulators from the lips to the larynx, and also, the contrast between tissue and air is sharp enough. In addition, images that are computed in the SENSE system lack the artefacts (reflection of the back part of the head to the front) that occurred in the previous gradient echo recordings.

The test material includes nine existing German words with simple CV-structures, in which all vowels and lingual consonants of Standard Modern German are represented (though obviously not every possible coarticulation effect could be taken into account). Real words proved to be the only possible test material, since the communication environment during the MRI recordings would make understanding of nonwords very difficult. In addition to the test words, fluent speech is being recorded acoustically, and the visual recordings (MRI and video fluoroscopy) also include the performance of extreme tongue movements, like protruding the tongue, moving it up and down to the maximum extension, or pulling it backwards towards the velum. These tasks should provide additional information about tongue mobility and motility. As oral cancer often causes a swallowing disability, a swallow test follows the speech recordings in both MRI and video fluoroscopy, what is also part of the diagnostic routine.

The resection and reconstruction of the defect is documented in a surgical mapping protocol (Mackenzie Beck et al., 1998) where resection and reconstruction are separately mapped, providing a good distinction between tumour location and reconstruction method. This documentation should allow for conclusions regarding the connection between muscle function and articulatory impairment. As a supplement, patients are requested to complete a set of personality tests, which should shed light on personal motivation and personality type and on a possible correlation with the postoperative speech ability. The pre-operative personality test will be compared with a second questionnaire six months after the operation.

3. Preliminary results

Our study is designed to rely on results from about 100 patients in all, of which we would now like to present preliminary results. Previous research has been done on movements of the tongue including lips and larynx during vowel articulation using MRI (Demolin et al., 2000; Kröger et al., 2000), but the present speed of the SENSE sequence of 8 pictures per second allows first steps towards imaging of dynamic consonant articulation. (After further technical improvements, we hope to achieve a maximum of 15 MR images per second soon, with unchanged resolution and slice thickness.)

As mentioned above, the aim of our study is not only to get a deeper insight into articulatory movements as such, but also to define categories which contain sufficient information for the evaluation of the degree of the postoperative articulatory impairment in the glossectomee. For this reason, it was necessary to create a general analysis method that can be used not only for the same patient pre- and postoperatively, but also for different patients, resulting in comparable data. Thus, two points had do be defined which are neither influenced by the articulation, nor by the tumour resection. The first such point is the inferior apex of the alveolar ridge, the second is the anterior inferior edge of the second cervical vertebra (axis), corresponding approximately to the vertical height of the hyoid bone (see Figure 1). From the middle point of this line, a grid of 13 vectors is constructed (where adjacent lines are at an angle of 15° to each other). The resulting semicircle covers the entire oral area where consonant articulation takes place. Along each line, two points are measured for every tongue position, beginning from the origin of the grid: (a) the outer edge of the tongue and (b) the closest point of the palate, thus the origin–palate distance is identical with the maximal possible tongue extension (100%) along the line. In compensatory articulation, the lips are likely to participate in some movements where they are normally not involved. Therefore the lip distance is measured separately, bilabial distance and lip protrusion being the two main factors.

Figure 1: Reference points for the semicircular grid. Fast gradient echo image with an artefact on the left side of the picture.

Two kinds of analysis can be run on this data: (1) The movement of 13 points of the tongue can be illustrated in a dynamic way (see Demolin et al., 2000), and (2) the overall extension of the tongue in the sagittal plane can be calculated within the "palatal semicircle", completed with the (3) reconstruction method and (4) the acoustic impression. These four factors play an equal, though different role in the evaluation of the postoperative speech ability, thus their correlation is extremely important. This can be exemplified by the following case.

Figure 2 and 3 show the articulation of /l/ and /k/ by the same patient, before and after the operation. The place of articulation for /l/] is not grossly impaired; however, the characteristic combination of apical articulation and rather constricted pharynx is replaced by a laminal-looking articulation and wide pharynx. According to the acoustic recording, there is no clear closure in the /l/ articulation postoperatively, the sound could be identified as /l/ or /j/.

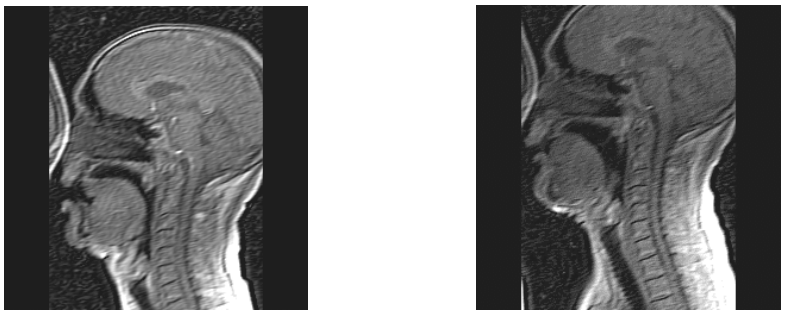

Figure 2 a & b: articulation of /l/ in the word *lang* by the same patient, (a) pre- and (b) postoperatively.

In the case of /k/ there is a change from a postpalatal closure preoperatively (normal for the nominally velar consonants in this front vowel context) to an incomplete constriction in the pre-palatal region postoperatively. The acoustic impression is not intelligible as /k/.

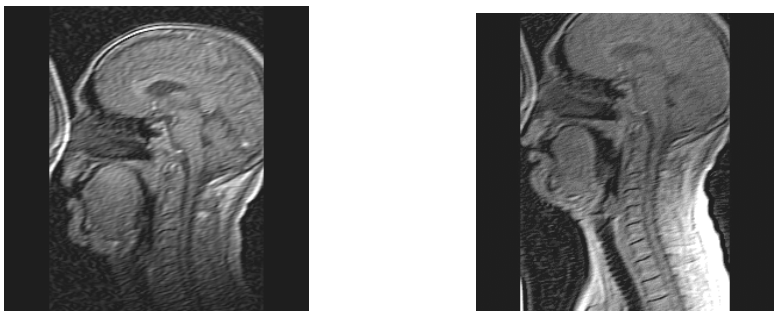

Figure 3 a & b: articulation of /k/ in the word *Theke* by the same patient, (a) pre- and (b) postoperatively.

As can be seen, the tongue movements are restricted mostly to the anterior area of the palate which makes a velar closure impossible. This fact is also well documented on the MR image of postoperative /x/ articulation (see Figure 4b), where the dorsum elevation is clearly not sufficient in the posterior area. Despite the changed articulation, the acoustic features of /x/ are not definitely impaired. This can be explained by the fact that in case of non-strident fricatives, precise control of constriction size is not crucial.


Figure 4 a & b: articulation of /x/ in the word *Tuch* by the same patient, (a) pre- and (b) postoperatively.

It is important to bear in mind that the reason for the reduced elevation of the tongue dorsum can be complex: the articulation defect can be caused by impaired muscle activity as a consequence of the operation, by the reconstruction method, or the combination of both. In the present case, the entire anterior part of the tongue on the right side had to be removed, including the apex and the median frontal region of the left tongue. The superior anterior part of the remaining tongue was folded down and fixed to the inferior side, while the lateral defect on the right side was fixed at the floor of the mouth. As a consequence, the elevation of the right tongue part is only possible to some extent, as shown in Figure 5b. As mentioned above, this also seems to influence the articulation of dentialveolar consonants like /l/.


Figure 5 a & b: backward movement of the tongue – maximum extent (a) before and (b) after the operation.

A possible shortcoming of our method is the fact that no other parts of the oral cavity are imaged than the midsagittal plane. But this does not seem to lead to inconsistent results, because consonant articulation elsewhere than in the midsagittal area does in fact mean an incorrect articulation, which is also supported by the acoustic impression.

4. Conclusion

Studying articulatory and compensatory mechanisms in a changed oral environment is not only of theoretical relevance, but it is also useful in providing clinical guidelines regarding oral reconstruction methods. The present speed of the MR imaging technique reveals processes of articulation in a novel way, so that the evaluation of speech can be completed with the careful control of consonant articulation, and more can be learned about compensatory articulation. In conjunction with the personality questionnaire, this method may also help to pinpoint psychosocial factors that have an influence on speech impairment following glossectomy.

References

Bressmann, Tim et al. (2000). Quantitative aspects of glossectomee speech production. In *Proc. ICPLA 2000* (in press).

Demolin, Didier, Metens, Thierry & Soquet, Alain. (2000). Real time MRI and articulatory coordinations in vowels. In *Proc. 5th Speech Production Seminar* (pp. 86–93).

Kröger, Bernd J. et al. (2000). Estimation of vocal tract area function from magnetic resonance imaging: preliminary results. In *Proc. 5th Speech Production Seminar* (pp. 333–336).

Mackenzie Beck, Janet et al. (1998). Surgical mapping and phonetic analysis in intra-oral cancer. In W. Ziegler & K. Deger (Eds.), *Clinical Phonetics and Linguistics* (pp. 481–492). London: Whurr.

Sader, Robert (1999). Untersuchungen zur velopharyngealen Insuffizienz bei Patienten mit Lippen-Kiefer-Gaumenspalten und Dysgnathien. Habilitationsschrift, Fakultät für Medizin der TUM.