Articulatory parameters in consonant production after tumour surgery: a real-time MRI investigation

Katalin Mády[1] & Ambros Beer[2]

[1]Institute of German Studies, Pázmány Péter Catholic University
2087 Piliscsaba, Egyetem u. 1., Hungary[2]Department of Radiology, University of Technology
Ismaninger Straße 22, 81675 München, Germany

mady@btk.ppke.hu, beer@roe.med.tu-muenchen.de

ABSTRACT

This study introduces a technique used in clinical research of speech production for the first time: real-time magnetic resonance imaging (MRI). Our goal was to show the clinical relevance of the use of real-time MRI in speech production. Based on the MRI recordings, four parameters of consonant production were tested: (1) localisation of minimal distance, (2) length of constriction, (3) minimum distance and (4) difference of minimum and maximum distance in front of the uvular region. In the preoperative analysis, parameters (1), (2), and (3) are ranked for the sounds /l/, /s/, /S/, and /x/ from minimum to maximum and parameter (4) inversely. The postoperative results show that this hierarchy is only maintained for parameter (3): minimum distance. The postoperative outcome regarding the other parameters becomes unsystematic depending on the severity of the impairment.

1 Introduction

Articulatory phonetics has long been a discipline with mainly qualitative analysis methods. In the last decades, several visualisation techniques, such as electropalatography (EPG), electromagnetic midsagittal articulography (EMMA) or ultrasound have been developed, which also allow for a quantification of articulatory parameters.

The visualisation of speech production can be even more crucial in the evaluation of distorted speech, as the researcher cannot rely on articulatory processes known from normal speech. However, not all articulatory research methods that are commonly used with normal subjects can be adapted for clinical purposes. This is especially true when the functional outcome of surgery is to be investigated in a relatively short period after the operation, as in the case of oral tumour surgery. Although several visualisation methods have been used in this field (for a summary, see [1]), they all have disadvantages: either they are invasive/inconvenient for the subjects (e.g. X-ray, EMMA or EPG) or their output is difficult to quantify (e.g. ultrasound).

In our study, articulatory analysis was based on a relatively new method that has not been used in clinical research so far: real-time magnetic resonance imaging (MRI). This technique has several advantages: it is non-invasive (i.e. not harmful for the subjects), objective (i.e. independent from the researcher), and the output signal is digital, so there is no information loss during data transfer. One disadvantage of real-time MRI is that the recordings are accompanied by a loud noise; another shortcoming is the relatively poor time resolution, although there is at present a rapid development in this field.

As the analysis methods to be presented here have not been used so far, the evaluation of impaired consonant production had to be preceded by an attempt to find appropriate parameters that reflect the adequate features of consonant production. Previous studies show that not only place and manner of articulation (PoA and MoA), but also tongue shape and some characteristics of the constriction play an important role in consonant distinction (see [2], [3], and for an overview, [4]). For example, Hoole et al. [2] investigated alveolar /s/ and postalveolar /S/. According to the traditional distinction, the two sibilants are differentiated by their PoA which is more fronted for /s/ than for /S/. In fact, it seems that the crucial difference is not the exact localisation of the PoA but rather the shape of the tongue surface immediately behind the constriction: due to a groove that is characteristic for the production of /s/ (see Section 2), the midsagittal tongue shape is flat, while it is domed for /S/, where the deep median groove is missing. Thus, PoA is one, but not necessarily the crucial distinctive articulatory feature in the production of these sibilants.

In the following, an analysis method for real-time MRI recordings will be presented and tested for four consonants. The results for normal (i.e. preoperative) speech will be then compared with the impaired sound production in postoperative recordings. Finally, we will point out some general tendencies which deserve more attention in forthcoming research.

2 Articulatory characteristics of /s/, /S/, /l/, and /x/

Before describing the analysis, some important characteristics of the unimpaired production of the sounds investigated here will be pointed out. The MoA of fricatives can be described as the emission of an air jet through a constriction at some part of the oral cavity, in which the tongue or the lips are moved towards another anatomic structure, such as the alveolar ridge, palate, teeth, etc. This is characteristic for /x/ where the constriction involves the tongue back and the soft palate. However, the sibilants /s/ and /S/ require a second obstacle: the teeth or the incisors ([4], [5]).

The production of the alveolar /s/ and postalveolar /S/ can be differentiated by three aspects: (1) the constriction in /s/ is formed somewhat more frontal than in /S/, (2) /s/ is produced with a deep and narrow longitudinal groove in the midsagittal part of the tongue that is less distinct in /S/, and (3) /s/ is produced with a flat tongue shape (concave form) whereas in /S/, the tongue back is domed (convex form). This coincides with a longer constriction of /S/ compared to /s/.

The difference between the tongue shapes of the alveolar lateral /l/ and the alveolar fricative /s/ can be described as follows: there is a midsagittal closure and lateral opening in /l/ whereas /s/ requires a midsagittal opening and lateral closure. /l/ has a slightly different PoA from /s/: /l/ is always produced by the tongue tip whereas the production of German /s/ mostly involves the tongue blade. Thus, the constriction is supposed to be somewhat more frontal in /l/ than in /s/. The tongue contour of /l/ is even more flat than in /s/, without showing the characteristics of a groove.

The German fricative /x/ can have a velar or uvular PoA. According to Kohler [6], velar realisations occur after the high vowels /o/ and /u/, uvular allophones after low /a/ and /O/. Thus, in the target word *Tuch* ('scarf'), the velar realisation of /x/ is more probable than the uvular one.

The fricative /x/ differs from the other three consonants not only by the PoA, but also by the constriction size, i. e. the distance between soft palate and dorsum. It has been shown that German /x/ is associated with a relatively large articulatory variability as opposed to alveolar sounds, as the number of back consonant phonemes in German is smaller than that of front consonants and thus, the perceptual categorisation of back consonants is easier even in case of a less careful production. Due to the smaller mobility of the tongue back, the constriction length in /x/ is greater than in apical and laminal sounds.

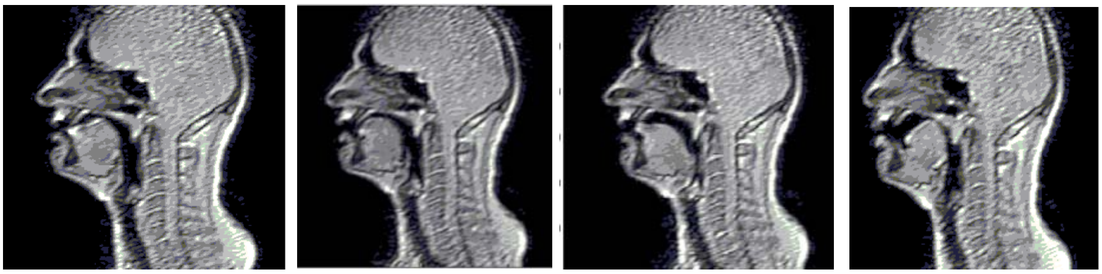

Figure 1a–d. *Preoperative MRI scan of the consonants (a) /s/, (b) /S/, (c) /l/, and (d) /x/.*

3 Material and methods

3.1 Subjects

Eight male subjects (age mean: 55.8 years, SD: 12.0 years) suffering from a squamous cell carcinoma of the oral cavity were involved in this study. All of them were regular patients at the Department of Oral and Maxillofacial Surgery of the Technical University, Munich, between November 2000 and October 2001. Before the operation, no impairment of the speech or hearing ability was found in the subjects. The tumours were sized between 0 and 4 cm and located at the tongue and/or floor of the mouth. The overall size of the resection (amount of removed tissue) was larger in order to maintain a security distance to the surrounding healthy tissue. Thus, the largest resection involved nearly 50% of the tongue body (but not the tongue root), while the smallest defect was only a few mm large and affected only the lateral tongue surface.

The reconstruction techniques regarding defect closure were also different among subjects: the tongue tissue was either covered by a platysma flap or fixed to the floor of mouth. (For details, see [1].)

3.2 Recording technique

The analysis was performed on the sounds /s/, /S/, /l/ and /x/ embedded in real German words. All subjects were recorded twice: a few days before the operation and approximately 4 weeks after surgery, before the starting of an optional radiotherapy.

The articulatory analysis was based on real-time MRI recordings (Philips ACS NT Gyroscan, magnetic field: Bo = 1.5 Tesla, gradient echo, SENSE technique). The imaging was based on the midsagittal plane with a slice thickness of 10 mm. The technical set-up allowed a time resolution of 8 images/sec. Each target word was repeated over 10 seconds and thus occurred 5–11 times per sequence, depending on the speed of speech and on word length.

The MRI recording sessions were accompanied by additional acoustic recordings that allowed an acoustic and auditory evaluation of the sounds in question.

3.3 Methods of articulatory analysis

The articulatory analysis was based on individual images that had been segmented from the target words. For an intra- and intersubject comparability, two anatomically defined reference points were chosen (Figure 2a). On these, a semi-polar coordinate system was imposed, in which the tongue and the palate contour were plotted (Figure 2b). Along each axis of the coordinate system (0–180°, 180 values), the distance of the tongue and the palate was measured in a Carthesian coordinate system (Figure 2c). Articulatory characteristics like the place and length of constriction or tongue shape were thus visualised by tongue-palate distance (TPD) trajectories which allow a direct comparison between pre- and postoperative realisations of the sounds by different subjects.

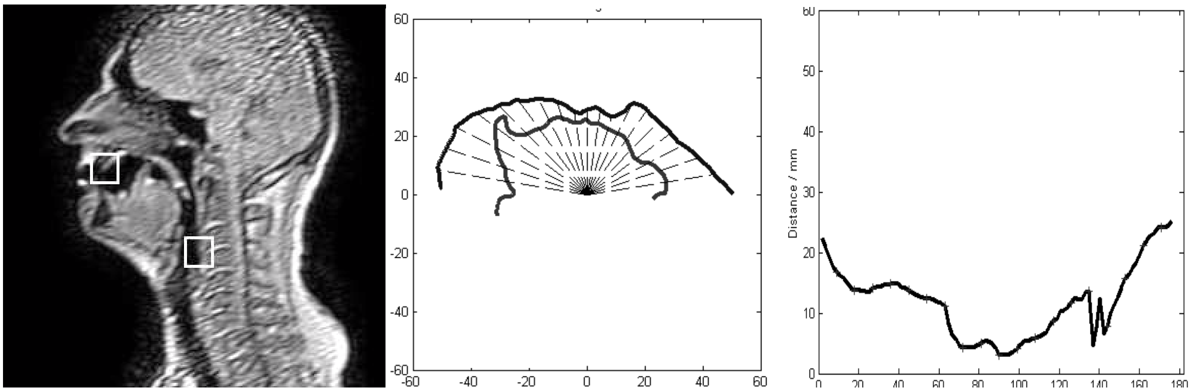

Figure 2a–c. *Backward movement of tongue tip along the hard palate. (a) Front and back fixed points for the semi-polar coordinate system, (b) tongue and palate contour in the semi-polar coordinate system, (c) tongue-palate distance trajectory.*

The TPD trajectory enabled us the define place and length of the constriction. In order to get a quantitative description of the tongue shape, the first derivate (FD) of the TPD trajectory was computed. In other words, each distance value was subtracted from the distance value connected to the next grid point, thus, the first value of the FD trajectory equalled the difference between the second and first TPD distance, the second value the difference of the third and second TPD distance etc. The output curve makes rises and falls of the TPD trajectory more obvious: FD values above zero indicate a rise, values below zero a fall, and values around zero a stagnation of the TPD trajectory (Figure 3c).

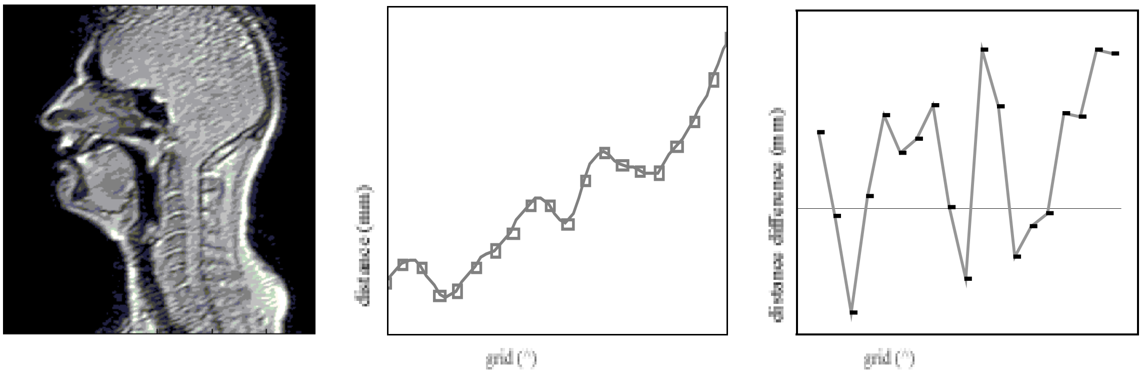

Figure 3a–c. */s/, preoperatively, (a) MRI scan, (b) TPD trajectory and (c) first derivate for the anterior part of the oral cavity.*

Based on the TPD trajectory and on FD, following parameters were used to investigate consonant production: (1) localisation of minimal distance, (2) length of constriction (number of FD values between –0.1 and 0.1), (3) minimum distance and (4) difference of maximum and minimum distance in front of the uvular region. Parameter (1) can be roughly interpreted as the PoA, while parameters (2) and (3) point to the characteristics of the constriction. Parameter (4) shows whether the midsagittal tongue shape is flat behind the constriction like in /s/ and /l/, i.e. the difference between minimum and maximum is large, or whether the tongue shape is domed like in /S/, the difference thus being smaller.

Parameters (1) and (2) are given in degree, while parameters (3) and (4) are distances in mm. The distance between each degree is smaller than 1 mm (the length of the vocal tract being 170 mm, from which not all parts were being considered in our analysis). However, as no cross-parameter comparison was performed, the different characters of the measures were ignored in the analysis.

4 Results

In the following section, the consonants /s/, /S/, /l/, and /x/ will be described in terms of the articulatory parameters presented in Section 2. First, the main parameters of the preoperative realisations will be elaborated. The second step will be to compare these with the postoperative realisations, which were classified according to the severity of the impairment for each consonant, evaluated by the first author. The presentation of the postoperative results will be based on this classification.

4.1 Preoperative analysis

The analysis was based on the mean values of the preoperative realisations among all subjects. The values of parameter (1) correspond to the previous expectation (Figure 4): while it is evident that the distance minimum in /x/ is connected to a high value (mean: 58.4°), it was not quite as obvious that /s/ and /S/ are distinct regarding their PoA – as is indeed not always the case within a single subject. The apical character of /l/ as opposed to the laminal /s/ (both being alveolar sounds) is reflected by the fact that the lateral is produced somewhat more frontal than the sibilant. The same difference is indicated by the constriction length for these sounds: a contact produced with the apex is naturally shorter than a constriction involving the tongue blade. The longer constriction in /S/ as opposed to /s/ is due to the domed tongue shape in /S/, while the long constriction in the velar fricative /x/ is due to its dorsal production – the dorsum being a less flexible organ compared to the tip of the tongue.

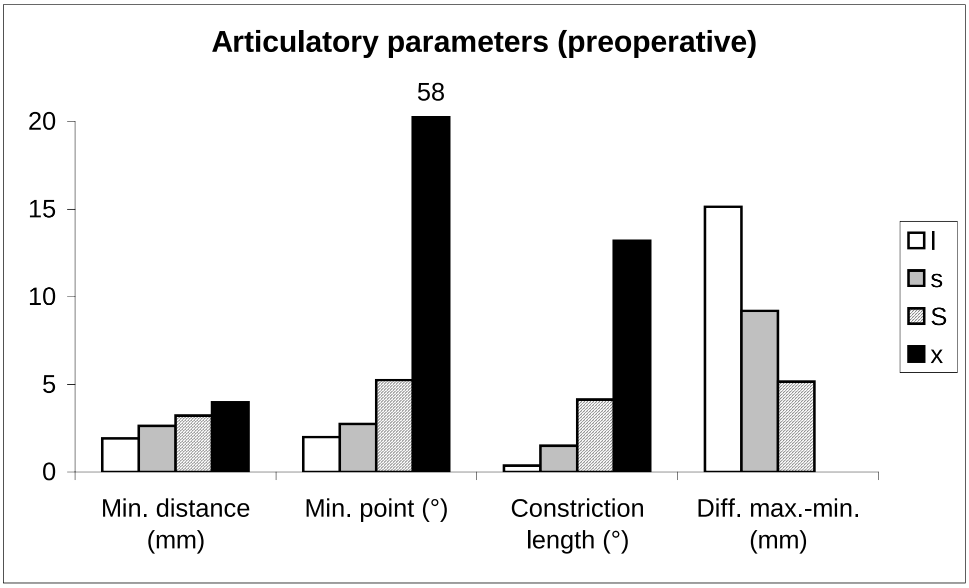

Figure 4. *Articulatory parameters for the consonants /l/, /s/, /S/, and /x/, preoperatively.*

While /s/, /S/, and /x/ are fricatives and thus produced with a constriction instead of a complete closure of the articulatory organs as opposed to /l/, it is surprising that the mean minimal distance for /l/ is definitely larger than 0 mm (1.92 mm). Two explanations are possible: (1) some of the MRI scans used for the analysis did not exactly correspond to the moment of the closure, thus, the tongue tip was not yet or no more producing a total contact with the alveolar ridge, or (2) the measured distance is due to an artefact that is sometimes present in MRI segments at contact surfaces. Still, the fact that the smallest distance was measured for /l/ meets our expectations regarding the different closure type between the lateral sound and the three fricatives.

The flat tongue shape in /l/ and /s/ is expressed by the large maximum-minimum differences in these sounds. Although /S/ is produced with a domed tongue, there is still a distance maximum in the region of the hard palate. This is due to the fact that the palate itself is also domed (see Figure 1). The same is not true for the velar sound, for which the anterior tongue shape does not play a role.

The somewhat larger distance minimum in /x/ is due to the involvement of the dorsum instead of the tongue tip, as was argued above.

4.2 Postoperative analysis

The postoperative analysis relied on auditory evaluation (subjective impressions of the first author) and on acoustic analysis described in [1] and in [7]. While /x/ was normally not connected with any postoperative altering, an impairment of /S/ was characterised by a lateralisation and that of /s/ by a palatalisation. The latter led to a complete neutralisation of the two sounds showing roughly the characteristics of preoperative /S/. The impairment of the /l/ production involved a laminal instead of apical production, together with a palatalisation. In more severe cases, the lateral character of the sound was missing, and no closure could be found.

The analysis of the postoperative results is based on three groups for each sound: Group 1 includes realisations that remained unimpaired after the operation. Group 2 involved realisations that correspond to the tendencies described above (moderate impairment). Realisations that were not identifiable as the target were counted as Group 3 (severe impairment) (Table 1).

Table 1. *Consonant realisations based on postoperative impairment.*

	Group 1	**Group 2**	**Group 3**
/l/	1	4	3
/s/	3	4	1
/S/	3	4	1
/x/	7	1	-

The parameters for Group 1 (Figure 5a) resemble the preoperative tendencies only partly: while the constriction length and the tongue shape (i.e. the difference between maximum and minimum distance) show the same rank order as in normal articulation, this is not quite true for the minimal distance which is larger for /l/ than for /s/ or /S/. However, these results do not allow for general conclusions, as in Group 1, /l/ is only represented by the realisations of a single subject. There seems to be a tendency not to distinguish the PoA for /s/ and /S/, as the PoA of /s/ is located slightly further back in the oral cavity than for /S/. It is interesting that this does not seem to influence the perceptual image (i.e. the auditory judgement) of these sounds.

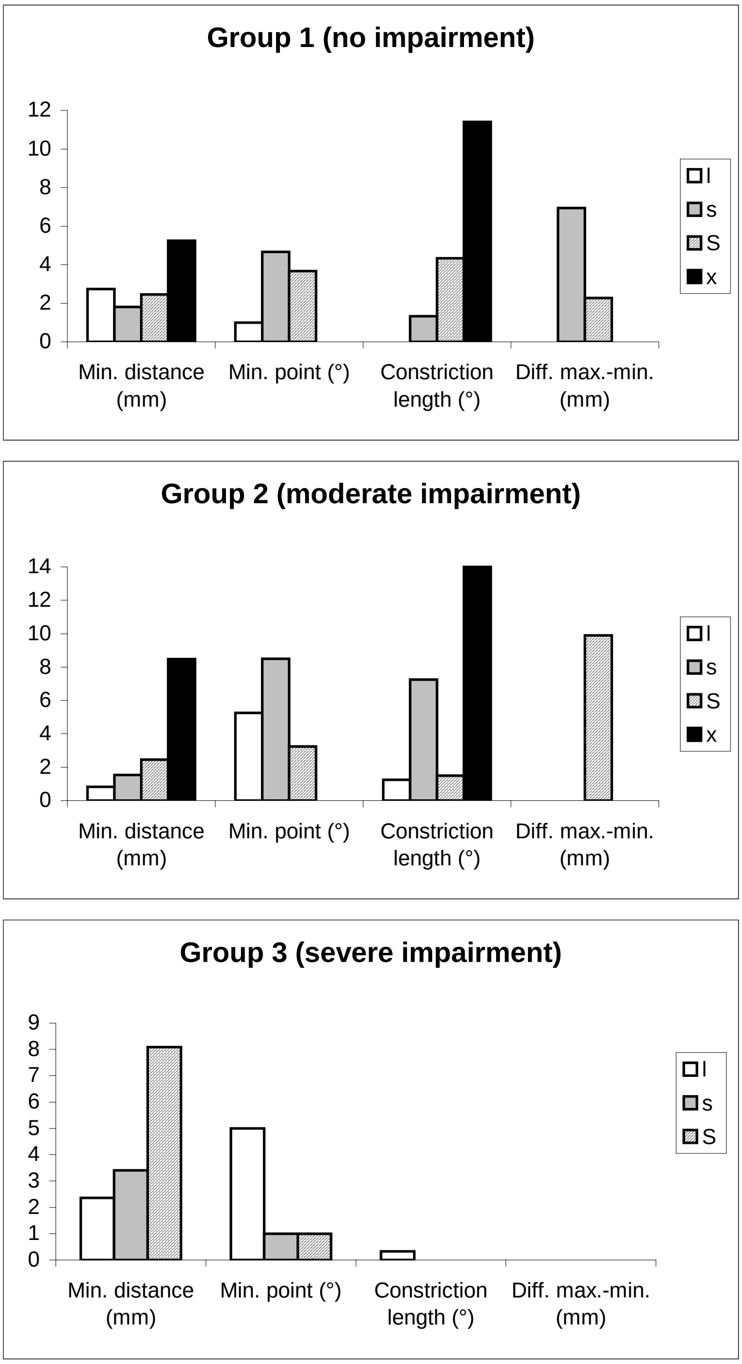

Figure 5a–c. *Postoperative consonant realisations with (a) no impairment. (b) moderate impairment, and (c) severe impairment.*

In Group 2, the overall tendencies of preoperative sound production are almost completely missing: difference values could be calculated only for few sounds, and neither the PoA nor the constriction length reflected the rank order found in normal articulation. However, the minimal distance of the constriction is maintained for this group.

The same can be said about Group 3 (NB. no /x/ realisation belonged to this group!): while most parameters were not organised along a hierarchy or were not even measurable, the minimal distances reflected exactly the same rank order as in the preoperative articulation.

5 Discussion and conclusions

In our study, articulatory analysis based on real-time MRI recordings was performed for a group of subjects suffering from a tumour of the oral cavity. Although tumour size and localisation were different for the subjects, the surgery and the reconstruction had led to an involvement of the tongue in all of them. The analysis method presented here allowed careful generalisations of the pre- and postoperative consonant production. On this basis, a rank order of four parameters regarding the size and localisation of the constriction, on one hand, and of tongue shape, on the other hand, were investigated for normal and impaired speech.

The normal (i.e. preoperative) production of /l/, /s/, /S/, and /x/ was characterised by following tendencies: (1) the apical and alveolar lateral /l/ was produced with the furthest PoA, followed by the laminal and alveolar fricative /s/, by the laminal and postalveolar fricative /S/, while the dorsal velar fricative /x/ had naturally the most posterior PoA. (2) The more posterior the PoA was in the oral cavity, the longer was the constriction. (3) fricatives were produced with a larger distance between tongue and alveolar ridge/palate than the lateral /l/, for which a complete closure between apex and alveolar ridge is characteristic.

These tendencies were only partly present in postoperative realisations that did not reveal detectable impairments via auditory analysis (Group 1). While the rank order for the minimal distances was only partly maintained, the localisation of the minimum indicating PoA lacked any recognisable tendency. It is not fully surprising, as for the distinction between /s/ and /S/, the tongue shape in the frontal region seems to be more crucial than the PoA itself (see [2], [7]). Thus, the maintenance of a flat tongue shape is probably a better indicator for the correctness of /s/ production than the fact that the constriction for /s/ is located more frontal than for /S/.

In the moderately and severely impaired production of the investigated consonants, only one of four tendencies was maintained: the minimal distance between tongue and alveolar ridge/soft palate. This finding did not correspond to the degree of impairment, i.e. it was present in moderately as well as severely distorted sound realisations. In other words, even if the motor skills are not sufficient to fulfil more crucial requirements that belong to the production of a given consonant, a less important distinction is maintained instinctively, thus becoming the only indicator of the intended production type instead of being a side effect, as in normal production.

The relatively poor mobility of the tongue tip is possibly due to the reconstruction technique that was used for this group of subjects: after surgery, only one patient was able to use his apex without restriction. This fact is also indicated by his being the only subject with an unimpaired postoperative /l/ production. It seems that the reduced mobility of the tongue tip does not necessarily lead to an impairment of coronal sounds in general, as long as their normal production is laminal instead of apical.

It is important to note that the results described above represent only a small group of subject and only a few reconstruction techniques that are used in oral tumour surgery. Thus, the findings presented here have to be interpreted carefully. This paper was rather intended to show whether the analysis method that was developed in our study was appropriate for consonant analysis.

Based on the described tendencies that reveal a connection between preoperative and postoperative consonant production, on one hand, and between the degree of speech impairment, on the other hand, it is to conclude that real-time MRI is a useful method in clinical speech evaluation and can contribute to future research. More exact and reliable results could be gained from recordings, where more subjects and more consonants are involved.

References

[1] Mády, K. & Beer, A., 2004, A real-time MRI evaluation of consonant production after oral tumour surgery. *Grazer Linguistische Studien* **62**, 77–94.

[2] Hoole, P., Ziegler, W., Hartmann, E., & Hardcastle, W., 1989, Parallel electropalatographic and acoustic measures of fricatives. *Clinical Linguistics & Phonetics* **3**, 59–69.

[3] Hoole, P., Nguyen-Trong, N., & Hardcastle, W., 1993, A comparative investigation of coarticulation in fricatives: electropalatographic, electromagnetic, and acoustic data. *Language and Speech* **36**, 235–260.

[4] Ladefoged, P. & Maddieson, I., 1996, The sounds of the world’s languages. Oxford: Blackwell.

[5] Shadle, C. H., 1990, Articulatory-acoustic relationships in fricative consonants. In: Speech production and speech modelling, W. J. Hardcastle & A. Marchal (Eds.). Dordrecht: Kluwer, 187–209.

[6] Kohler, K. J., 1995, Einführung in die Phonetik des Deutschen, 2., reviewed edition.. Berlin: Erich Schmidt.

[7] Mády, K., 2004, Akustische, artikulatorische und perzeptive Parameter in der Konsonantenproduktion nach Zungenteilresektion. Diss, Universität München. To appear in *Forschungsberichte Institut für Phonetik und Sprachliche Kommunikation der Universität München.*
URL: http://edoc.ub.uni-muenchen.de/archive/00003060/01/Mady_Katalin.pdf