Original Paper

Folia Phoniatrica et Logopaedica

Folia Phoniatr Logop 2008;60:151–156
DOI: 10.1159/000121004

Published online: March 14, 2008

# Automatic Quantification of Speech Intelligibility of Adults with Oral Squamous Cell Carcinoma

M. Windrich A. Maier R. Kohler E. Nöth E. Nkenke U. Eysholdt M. Schuster

Department of Phoniatrics and Pediatric Audiology, Universitätsklinikum Erlangen, Erlangen, Germany

**Key Words**

Oral carcinoma · Perceptual assessment · Automatic speech recognition · Intelligibility

**Abstract**

***Objective:*** It has been the aim of the present study to introduce a novel automatic technique for the objective and quantitative assessment of speech intelligibility to the evaluation of postoperative outcome. ***Patients and Methods:*** Forty-six patients with oral carcinomas, mean age 59.8 ± 10.1 years, and an age-matched control group of 40 subjects without oral diseases. Recordings of a standard text read by the patients and the control group were analyzed by an automatic speech recognition system. ***Results:*** For the patients, automatic speech recognition yielded word recognition rates between 8 and 82% (mean 49 ± 19%), for the control group between 60 and 91% (76 ± 7%). Automatic evaluation closely correlated with the experts' perceptual evaluation of intelligibility ($r = -0.93$; $p < 0.01$). The multi-rater kappa of the experts alone (0.55) differed only slightly from the multi-rater kappa of the experts and the speech recognition system (0.58). ***Conclusion:*** For adults with speech disorders, automatic speech recognition may serve as a valuable tool to assess global speech outcome after treatment of oral squamous cell carcinoma objectively and quantitatively for clinical and research purposes.



## Introduction

Oral squamous cell carcinoma is one of the ten most common malignant diseases [1]. Its treatment often negatively affects expressive speech skills and reduces speech intelligibility. As a consequence, patients' postoperative quality of life can be significantly impaired [2]. There is no standardized assessment of speech disorders in adults or children at national or international level [3–5]. Moreover, a method that allows for assessing speech disorders objectively and independently is missing. Therefore, it is difficult to compare the outcome of different therapeutic options in a reliable way.

Semistandardized instruments for the analysis of speech disorders in adults are well known [6–10]. But the assessment of speech disorders or intelligibility is usually performed subjectively and therefore lacks reliability due to differences in individuals' experience and variable test conditions [11]. Therefore, a panel of several listeners is often used for scientific evaluation of speech. This method is still the most widely used technique for assessing speech intelligibility [12–20], phonematic disorders and temporal structure of speech [21–26]. For more reliable results, transcription tasks and multiple-choice tasks for several listeners have been found to be appropriate [7, 9, 15, 17, 19, 21, 25]. Unfortunately, the use of several listeners is rather time-consuming and is thus only used for

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com


1021–7762/08/0603–0151$24.50/0

Accessible online at:
www.karger.com/fpl

Maria Schuster
Department of Phoniatrics and Pediatric Audiology, Universitätsklinikum
Bohlenplatz 21, DE–91054 Erlangen (Germany)
Tel. +49 9131 853 3146, Fax +49 9131 853 9272
E-Mail maria.schuster@phoni.imed.uni-erlangen.de

**Table 1.** Age and gender of 46 patients with oral squamous cell carcinoma and the control group of 40 subjects

| Age | Patients | | | Control group | | |
|---|---|---|---|---|---|---|
| | women (n = 13) | men (n = 33) | total (n = 46) | women (n = 10) | men (n = 30) | total (n = 40) |
| Range | 34.4 – 80.3 | 45.4 – 80.1 | 34.4 – 80.3 | 27.8 – 78.8 | 30.2 – 82.1 | 27.8 – 82.1 |
| Mean ± SD | 62.3 ± 13.4 | 58.8 ± 8.5 | 59.8 ± 10.1 | 50.5 ± 12.8 | 60.8 ± 12.6 | 58.1 ± 13.3 |

research projects. For clinical purposes, usually only one expert evaluates the patient's speech.

Objective and independent diagnostic tools for the assessment of speech intelligibility after treatment have only been performed for the quantification of nasalance [27] and the spectral characteristics and intensity of the voice signal [28]. However, these methods do not allow for assessing speech intelligibility in a comprehensive and reliable way. A new technique for objective evaluation of speech intelligibility has been tested as a diagnostic tool in adult patients who suffered from neurological diseases [29], who stuttered [30, 31], in laryngectomees with tracheo-esophageal speech [32] and in children with cleft lip and palate [33, 34]. This method is based on a statistical analysis of speech with established methods of automatic speech recognition. It was the aim of the present study to test this method for the follow-up of patients treated for oral squamous cell carcinoma and to compare the results of automatic evaluation of speech intelligibility with a perceptual rating of intelligibility by a panel of expert listeners.

## Material and Methods

*Patients*

The study group comprised 46 patients with different speech disorders. Each patient was recorded during regular outpatient care (table 1). All of them had suffered from squamous cell carcinoma of the mouth (table 2). They were treated by surgery, microvascular reconstruction of the defect for tumors with at least $T_2$ staging, underwent modified radical neck dissection on the ipsilateral side and suprahyoid neck dissection on the contralateral side and had adjuvant radiation therapy. All patients were native German speakers, using a local dialect. Forty subjects without oral diseases speaking the same local dialect served as the control group (table 1).

The patients read a German text ('Der Nordwind und die Sonne', 'The North Wind and the Sun'), a fable by Aesop. It is known as the reference text for the International Phonetic Alphabet by the International Phonetic Association. The German text consists of 108 words. It is phonetically balanced and includes all possible phonemes of the German language. The text was divided into 10 sequences (10.8 ± 2.4 words) according to syntactic boundaries and shown on a computer screen. The speech samples were recorded with a close-talking microphone (Call4U Comfort-Headset, DNT GmbH, Dietzenbach, Germany, sampling frequency 16 kHz, 16 bit).

**Table 2.** Clinical TNM classification of patient group

| | $T_1$ | $T_2$ | $T_4$ |
|---|---|---|---|
| Patients | 16 | 8 | 22 |

*Automatic Speech Recognition System*

Intelligibility was measured by means of a state-of-the-art automatic speech recognition system [34]. It segments speech data into temporal units of 16 ms. Then temporal and spectral characteristics are analyzed and compared with word models. A speech recognition system evaluates two different information channels: acoustic signals modeled by stochastic word models (hidden Markov model) and word frequency in the text represented by a stochastic language model. A speech recognition system typically has a so-called bi- or trigram-language model. Hence, the probability of a word depends on the acoustic signal (obtained by the hidden Markov models) and the probability that this word follows the last spoken (bigram) or the two last spoken words (trigram). For our purposes we used a unigram-language model, which assumes that the current word is independent of previously spoken words. The only linguistic parameter included is the frequency of each word in the recognition vocabulary. Thus, recognition mainly depends on the acoustic signal of each single word. Our recognition system is polyphone-based at the acoustic level, i.e. the acoustic properties of a phoneme are computed with respect to the coarticulatory modulation caused by its phonetic context. It is well known [35] that the realization of a phoneme is influenced by its phonetic context. This is why most speech recognition systems use triphone models as elementary recognition units, i.e. a different model is trained for each phonetic context. For instance, a /ʃ/ /t/ /r/ model for /t/ with /ʃ/ on the left and /r/ on the right is trained with exactly this context. Word models then concatenate the respective elementary models. In our system, we use a pho-

**Table 3.** Pearson's correlation between expert listeners judging the intelligibility of 46 speech recordings

| Listener | Listener | | | | Mean of all other listeners |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| 1 | 1 | 0.82 | 0.87 | 0.90 | 0.91 |
| 2 | | 1 | 0.74 | 0.80 | 0.81 |
| 3 | | | 1 | 0.96 | 0.91 |
| 4 | | | | 1 | 0.95 |

'All' (right column) means the average of the remaining 3 listeners that is compared to each single listener.

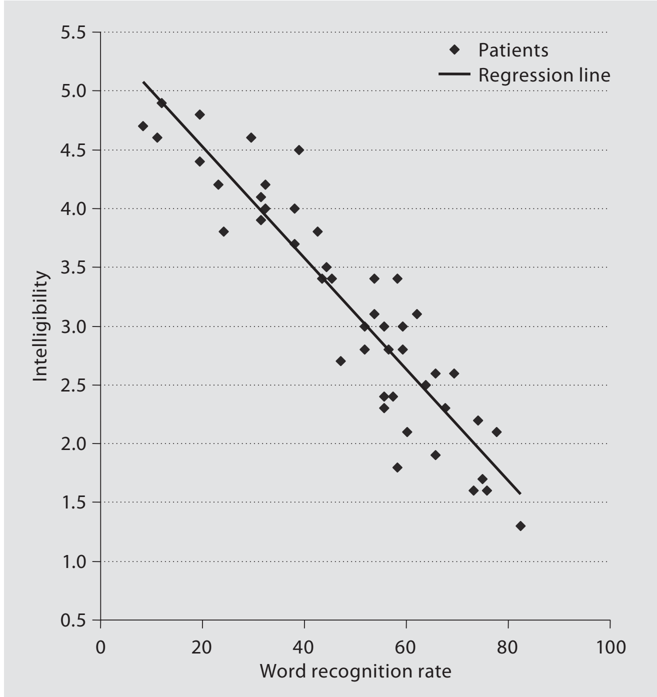


**Fig. 1.** Comparison of the listeners' ratings of intelligibility and WR. Perceptual rating is performed using categories from 1 (very high intelligibility) to 5 (very low intelligibility). WR is the percentage of correctly recognized words of the standard text. There is a high correlation between perceptually judged intelligibility scores and the automatic recognition system (r = –0.93).

netic context of arbitrary length (polyphones), provided it is observed often enough in the training database (more than 50 times). The acoustic model has been trained with speech recordings of 578 adult speakers (304 male, 274 female) without any speech disorder.

We computed the so-called word recognition rate (WR) of the speech data using the PEAKS evaluation software [36]. The computation time is less than half of real time. The WR describes the percentage of correctly recognized words in the whole text. It is calculated as follows:

$$WR\ [\%] = C/R \times 100\%$$

where C is the number of correctly recognized words and R is the number of words of the reference text.

*Perceptual Evaluation of Speech Intelligibility*

Three experienced clinicians and a speech recognition expert, who was well trained also in perceptual evaluation, estimated speech intelligibility independently by listening to a playback of the recordings presented via headphones. None of the expert listeners was familiar with the patients. The recordings were presented segment by segment and judged after each segment. The listeners used a five-point Likert scale (1 = very high, 2 = rather high, 3 = medium, 4 = rather low, 5 = low) to rate the intelligibility of all segments of the recordings. The average of these estimations formed the final score on a quasi-continuous scale for every record.

*Statistics*

The statistical analysis was performed using Microsoft Excel and our own evaluation software implemented in Java, which computed Pearson's product-moment correlation coefficient (r), Spearman's rank correlation coefficient (ρ), Student's t test, and multirater kappa [37]. The latter poses a special problem: the averages of listener ratings and word accuracy are continuous measures with completely different ranges. Thus, correlation techniques should be applied. If one wants to use kappa, mapping onto discrete values has to be performed. Therefore, we rounded the listeners' average intelligibility scores to the next integer and set thresholds on the WR results, so that the difference between the listeners' scores and the scores derived from the WR values was minimal. Although thresholds should not be set on the test data, the small size of our database does not allow us to define a validation set. Therefore, the kappa values for the WR results should be interpreted with caution.

## Results

*Subjective Rating*

Listener ratings showed good agreement (table 3). The multirater kappa for all raters was 0.55. (A kappa value of 0.4 is considered as moderate agreement and a value of 0.75 as strong agreement.) Pearson's correlation for all raters ($r = 0.90 \pm 0.06$) was found to be very close to Spearman's correlation ($\rho = 0.89 \pm 0.06$).

*Comparison between Subjective and Objective Assessment of Intelligibility*

There was a strong correlation between listeners' ratings and automatic assessment (word recognition rate) ($r = -0.93$; $\rho = -0.90$; $p < 0.01$, multirater kappa 0.58). The

**Table 4.** Results of automatic quantification of intelligibility in 46 patients with oral cancer and a control group of 40 patients without speech disorders

| | WR, % | |
|---|---|---|
| | range | mean ± SD |
| Patients' speech (n = 46) | 8–82 | 49 ± 19 |
| Speech of control group (n = 40) | 60–91 | 76 ± 7 |

distance between objective evaluation and subjective rating was less than one point on the Likert scale (fig. 1).

*Comparison between Patients and Control Group*

Automatically computed data of the control group showed significantly better WR values than for the patients (table 4). The WR of the control group was ranging from 60 to 91%. There was no correlation between age and WR in the patient group ($r = -0.06$; $p > 0.3$), while a weak correlation in the control group ($r = -0.3$; $p < 0.05$) was found.

## Discussion

In the present study, a new method for the automatic evaluation of speech intelligibility is introduced. This technique is based on the WR of spoken language as a means of representing intelligibility by speech recognition technique. The study revealed a significant correlation between the results of the automatic speech evaluation system and listeners' evaluation, despite the fact that evaluation of speech intelligibility by listeners is always biased due to considerable intraindividual variability.

The limitations of speech evaluation by listeners are evident in the results given in table 3. Although the listeners' evaluations show good correlation, they vary between listeners.

Several methods exist for perceptual assessment. We decided to use a five-point Likert scale because it produced similar results as visual analogue scales in preliminary experiments. Furthermore, quasi-linear scales are well-tested tools for voice evaluations, e.g. GRBAS or RBH scales. Nevertheless, Whitehill [4] and Schiavetti [38] questioned the validity of equidistant scales. Schiavetti [38] pointed out that 'listeners cannot effect a linear partition of the dimension with the result that the intervals used to scale speech intelligibility are not equal'. Nevertheless, he adds that practical limitations may force one to use a suboptimal method. So we decided for a descriptive five-point scale. To analyze the informative value of the Likert scale for this application, we compared Pearson's correlation coefficient with Spearman's rank correlation, since the latter is independent of the underlying scale used in an experiment. Spearman's correlation considers only the rank of the ratings instead of the actual value on the underlying scale. As described in the 'Results' section, both results were almost identical. For our experiments we used a standard test for all patients for better comparability, e.g. when comparing different patients or one patient's intelligibility over time. This, however, precludes the use of transcription or multiple-choice tasks as applied by other authors [39, 40]. Furthermore, transcriptions by multiple naive listeners have comparable inter-rater correlation as perceptual evaluations [41].

In general, speech recognition depends on five factors: the speaker, speech (read speech, spontaneous speech), vocabulary, grammatical complexity or perplexity (average probability of words possibly following a sequence of others), and the input medium [42]. For diagnostic purposes, the influence of most of these factors was minimized by using a standard text and a stable setting. Thus, the speaker remains the main factor of influence as required for the evaluation of intelligibility [38].

Automatic speech recognition techniques have been used successfully in the evaluation of communication disorders, such as severe voice disorders in laryngectomees, stuttering and speech disorders in children. In these cases the method showed a close correlation of automatically evaluated intelligibility with perceptual ratings by a panel of expert listeners. The present study shows that this method can also be used in speech disorders of adults. The correlation between 4 expert listeners and the automatic evaluation of intelligibility is very high ($r = -0.93$). To demonstrate the reliability of the new method, patients with a wide range of speech disorders formed the cohort. The study group included patients with small tumors, whose speech was only minimally disturbed, but also patients with large tumors and severe speech disorders. When speech intelligibility was perceived as high or very high, WR was within the range of the control group. Comparison between the data of the control group and those of the study group showed that speech intelligibility was significantly reduced ($p < 0.01$) in the study group (table 4). Moreover, the results of the control group demonstrated the variability of 'normal' speech that also depends on age. The finding that some of the speakers from the control group had poor results is not surprising. The recognizer was trained with mostly young people. In comparison, elderly people and children deviate from the 'standard' speaker and within the speaker group [43], showing lower WR. Also, the speech recognition system mostly uses acoustic information ('unigram language model'), which leads to restricted word recognition results. One has to keep in mind that absolute WR is not crucial, it is dependent on the training population and adaptations of the system, i.e. a speech recognition system that achieves a better recognition rate on the same corpus does not necessarily reflect intelligibility more accurately. The recognizer should rather cover the whole spectrum of intelligibility levels. Such is the case with the applied system, where the results vary from 8 to 91% and the range of the patient group differs from the range of the control group: $(\mu \pm \sigma)_{\text{patient group}} < (\mu - \sigma)_{\text{control group}}$.

In the future, a larger control group will allow for correcting patient data for age and gender effects on the WR in order to quantify deviations of a patient's WR from normal speech. It can be expected that the new method will be valuable and appropriate for clinical and scientific use. Automatic speech evaluation considers every single word and is independent of prosodic, pragmatic and contextual information that influence perceptual ratings. Therefore, it describes the acoustic properties of speech precisely and facilitates comparison between different speech samples independent of time and place of recording.

Further adaptations should enable the recognition of different phonematic disorders. In the future this will enhance the scientific evaluation of different surgical strategies concerning speech outcome and the identification of the best and least impairing treatment modality for oral cancer. Nevertheless, as with every diagnostic tool it does not replace the expert's interpretation and detailed examination.

## Conclusion

The automatic speech recognition system is a valuable technique for investigating speech intelligibility of adult patients treated for oral squamous cell carcinoma. It objectifies and quantifies global speech outcome for clinical and scientific purposes. The method is easily available and applicable for the global assessment of speech outcome.

## Acknowledgments

This work was supported by the ELAN-Fonds of Friedrich-Alexander University of Erlangen-Nuremberg. We thank Dr. Ralf Frisch for his valuable comments on this article.

## References

1 World Health Organization: The World Oral Health Report 2003. Geneva, World Health Organization, 2003, pp 6–7.
2 Gellrich NC, Schimming R, Schramm A, Schmalohr D, Bremerich A, Kugler J: Pain, function, and psychologic outcome before, during, and after intraoral tumor resection. J Oral Maxillofac Surg 2002;60:772–777.
3 Michi K: Functional evaluation of cancer surgery in oral and maxillofacial region: speech function. Int J Clin Oncol 2003;8:1–17.
4 Whitehill TL: Assessing intelligibility in speakers with cleft palate: a critical review of the literature. Cleft Palate Craniofac J 2002; 39:50–58.
5 Lohmander A, Olsson M: Methodology for perceptual assessment of speech in patients with cleft palate: a critical review of the literature. Cleft Palate Craniofac J 2004;41:64–70.
6 Schönweiler R, Altenbernd C, Schmelzeisen R, Ptok M: Articulatory capacity and intelligibility of speech of patients with carcinomas of the mouth cavity: a comparison of pre- and postoperative results of various reconstruction techniques. HNO 1996;44: 634–639.
7 Panchal J, Potterton AJ, Scanlon E, McLean NR: An objective assessment of speech and swallowing following free flap reconstruction for oral cavity cancers. Br J Plast Surg 1996;49:363–369.
8 Pauloski BR, Logemann JA, Colangelo LA, Rademaker AW, McConnel FM, Heiser MA, Cardinale S, Shedd D, Stein D, Beery Q, Myers E, Lewin J, Haxer M, Esclamado R: Surgical variables affecting speech in treated patients with oral and oropharyngeal cancer. Laryngoscope 1998;108:908–916.
9 Mády K, Sader R, Hoole PH, Zimmermann A, Horch HH: Speech evaluation and swallowing ability after intra-oral cancer. Clin Linguist Phon 2003;17:411–420.
10 Enderby PM: Frenchay-Dysarthrie-Untersuchung. Idstein, Schulz-Kirchner-Verlag, 2004.
11 Keuning K, Wieneke G, Dejonckere P: The intrajudge reliability of the perceptual rating of cleft palate speech before and after pharyngeal flap surgery: the effect of judges and speech samples. Cleft Palate Craniofac J 1999;36:328–333.
12 Robbins KT, Bowman JB, Jacob RF: Postglossectomy deglutitory and articulatory rehabilitation with palatal augmentation prostheses. Arch Otolaryngol Head Neck Surg 1987;113:1214–1218.
13 Wachter R, Diz Dios P: Effect of adaptation and compensation mechanisms on postoperative function of patients with tumors of the oral cavity. Laryngorhinootologie 1993; 72:333–337.

14 Bodin IK, Lind MG, Arnander C: Free radial forearm flap reconstruction in surgery of the oral cavity and pharynx: surgical complications, impairment of speech and swallowing. Clin Otolaryngol Allied Sci 1994;19:28–34.
15 Schmelzeisen R, Ptok M, Schönweiler R, Hacki T, Neukam FW: Reconstruction of speech and chewing function after extensive tumor resection in the area of the jaw and face. Laryngorhinootologie 1996;75:231–238.
16 Brown JS, Zuydam AC, Jones DC, Rogers SN, Vaughan ED: Functional outcome in soft palate reconstruction using a radial forearm free flap in conjunction with a superiorly based pharyngeal flap. Head Neck 1997;19:524–534.
17 Knuuttila H, Pukander J, Maatta T, Pakarinen L, Vilkman E: Speech articulation after subtotal glossectomy and reconstruction with a myocutaneous flap. Acta Otolaryngol 1999;119:621–626.
18 Haughey BH, Taylor SM, Fuller D: Fasciocutaneous flap reconstruction of the tongue and floor of mouth: outcomes and techniques. Arch Otolaryngol Head Neck Surg 2002;128:1388–1395.
19 Seikaly H, Rieger J, Wolfaardt J, Moysa G, Harris J, Jha N: Functional outcomes after primary oropharyngeal cancer resection and reconstruction with the radial forearm free flap. Laryngoscope 2003;113:897–904.
20 Markkanen-Leppanen M, Isotalo E, Makitie AA, Asko-Seljavaara S, Pessi T, Suominen E, Haapanen ML: Changes in articulatory proficiency following microvascular reconstruction in oral or oropharyngeal cancer. Oral Oncol 2006;42:646–652.
21 Mahanna GK, Beukelman DR, Marshall JA, Gaebler CA, Sullivan M: Obturator prostheses after cancer surgery: an approach to speech outcome assessment. J Prosthet Dent 1998;79:310–316.
22 Pauloski BR, Rademaker AW, Logemann JA, Colangelo LA: Speech and swallowing in irradiated and nonirradiated postsurgical oral cancer patients. Otolaryngol Head Neck Surg 1998;118:616–624.
23 Furia CL, Kowalski LP, Latorre MR, Angelis EC, Martins NM, Barros AP, Ribeiro KC: Speech intelligibility after glossectomy and speech rehabilitation. Arch Otolaryngol Head Neck Surg 2001;127:877–883.
24 Su WF, Hsia YJ, Chang YC, Chen SG, 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 2003;128:412–418.
25 Bressmann T, Sader R, Whitehill TL, Samman N: Consonant intelligibility and tongue motility in patients with partial glossectomy. J Oral Maxillofac Surg 2004;62:298–303.
26 Terai H, Shimahara M: Evaluation of speech intelligibility after a secondary dehiscence operation using an artificial graft in patients with speech disorders after partial glossectomy. Br J Oral Maxillofac Surg 2004;42:190–194.
27 Kuttner C, Schönweiler R, Seeberger B, Dempf R, Lisson J, Ptok M: Normal nasalance for the German language: nasometric values for clinical use in patients with cleft lip and palate. HNO 2003;51:151–156.
28 Zecevic A: Ein sprachgestütztes Trainingssystem zur Evaluierung der Nasalität; Dissertation Universität Mannheim, Mannheim, 2002.
29 Sy BK, Horowitz DM: A statistical causal model for the assessment of dysarthric speech and the utility of computer-based speech recognition. IEEE Trans Biomed Eng 1993;40:1282–1298.
30 Noeth E, Niemann H, Haderlein T, Decher M, Eysholdt U, Rosanowski F, Wittenberg T: Automatic stuttering recognition using hidden Markov models; in Yuan B, Huang T, Tang X (eds): Proc Int Conf on Spoken Language Processing, ICSLP, Beijing, 2000, vol 4.
31 Haderlein T, Steidl S, Noeth E, Rosanowski F, Schuster M: Automatic recognition and evaluation of tracheoesophageal speech; in Sojka P, Kopecek I, Pala K (eds): Proc 7th Int Conf on Text, Speech and Dialogue, Brno, 2004. Lecture Notes in Artificial Intelligence, vol 3206. Berlin, Springer, 2004, pp 331–338.
32 Schuster M, Haderlein T, Noeth E, Lohscheller J, Eysholdt U, Rosanowski F: Intelligibility of laryngectomees' substitute speech: automatic speech recognition and subjective rating. Eur Arch Otorhinolaryngol 2006;263:188–193.
33 Schuster M, Maier A, Haderlein T, Nkenke E, Wohlleben U, Rosanowski F, Eysholdt U, Noeth E: Evaluation of speech intelligibility for children with cleft lip and palate by means of automatic speech Recognition. Int J Pediatr Otorhinolaryngol 2006;70:1741–1747.
34 Stemmer G: Modelling Variability in Speech Recognition. Berlin, Logos-Verlag, 2005.
35 Menzerath P, De Lacerda A: Koartikulation, Steuerung und Lautabgrenzung. Berlin, Dümmlers Verlagsbuchhandlung, 1933.
36 Maier A, Noeth E, Batliner A, Nkenke E, Schuster M: Fully automatic assessment of speech of children with cleft lip and palate. Informatica 2006;30:477–482.
37 Davies M, Fleiss JL: Measuring agreement for multinomial data. Biometrics 1982;38:1047–1051.
38 Schiavetti N: Scaling procedures for the measurement of speech intelligibility; in Kent RD (ed): Intelligibility in Speech Disorders: Theory, Measurement and Management. Amsterdam, Benjamins, 1992, pp 11–34.
39 Zenner H: The postlaryngectomy telephone intelligibility test (PLTT); in Herrmann I (ed): Speech Restoration via Voice Prosthesis. Berlin, Springer, 1986, pp 148–152.
40 Pauloski BR, Logemann JA, Rademaker AW, McConnel FM, Heiser MA, Cardinale S, Shedd D, Lewin J, Baker SR, Graner D, Cook B, Milianti F, Collins S, Baker T: Speech and swallowing function after anterior tongue and floor of mouth resection with distal flap reconstruction. J Speech Hear Res 1993;36:267–276.
41 Haderlein T, Riedhammer K, Maier A, Nöth E, Toy H, Rosanowski F: An automatic version of the post-laryngectomy telephone test; in Matousek V, Mautner P (eds): Proc 10th Int Conf on Text, Speech and Dialogue, Pilsen, 2007. Berlin, Springer, 2007, vol 1, pp 238–245.
42 Gallwitz F, Niemann H, Nöth E: Speech recognition: state of the art, applications, and future prospects. Wirtschaftsinformatik 1999;41:538–547.
43 Wilpon JG, Jacobsen CN: A study of speech recognition for children and the elderly. Proc ICASSP 1996, pp 349–352.