International Journal of Pediatric Otorhinolaryngology (2006) **70**, 1741—1747

ELSEVIER

International Journal of Pediatric Otorhinolaryngology

www.elsevier.com/locate/ijporl

Evaluation of speech intelligibility for children with cleft lip and palate by means of automatic speech recognition

Maria Schuster *, **Andreas Maier, Tino Haderlein, Emeka Nkenke, Ulrike Wohlleben, Frank Rosanowski, Ulrich Eysholdt, Elmar Nöth**

Universitaetsklinikum der Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Phoniatrie und Paedaudiologie, Bohlenplatz 21, 91054 Erlangen, Franken, Germany

Received 23 February 2006; received in revised form 19 May 2006; accepted 23 May 2006

KEYWORDS
Cleft lip and palate;
Speech disorders;
Automatic speech recognition;
Intelligibility

Summary

Objective: Cleft lip and palate (CLP) may cause functional limitations even after adequate surgical and non-surgical treatment, speech disorders being one of them. Interindividually, they vary a lot, showing typical articulation specifics such as nasal emission and shift of articulation and therefore a diminished intelligibility. Until now, an objective means to determine and quantify the intelligibility does not exist.
Method: An automatic speech recognition system, a new method, was applied on recordings of a standard test to evaluate articulation disorders (psycholinguistic analysis of speech disorders of children PLAKSS) of 31 children at the age of 10.1 ± 3.8 years. Two had an isolated cleft lip, 20 a unilateral cleft lip and palate, 4 a bilateral cleft lip and palate, and 5 an isolated cleft palate. The speech recognition system was trained with adults and children without speech disorders and adapted to the speech of children with CLP. In this study, the automatic speech evaluation focussed on the word accuracy which represents the percentage of correctly recognized words. Results were confronted to a perceptive evaluation of intelligibility that was performed by a panel of three experts.
Results: The automatic speech recognition yielded word accuracies between 1.2 and 75.8% (mean 48.0 ± 19.6%). The word accuracy was lowest for children with isolated cleft palate (36.9 ± 23.3) and highest for children with isolated cleft lip (72.8 ± 2.9). For children with unilateral cleft lip and palate it was 48.0 ± 18.6 and for children with bilateral cleft lip and palate 49.3 ± 9.4. The automatic evaluation complied with the experts' subjective evaluation of intelligibility ($p < 0.01$). The multi-rater kappa of the experts alone differed only slightly from the multi-rater kappa of experts and recognizer.

* Corresponding author. Tel.: +49 913132612.
E-mail address: maria.schuster@phoni.imed.uni-erlangen.de (M. Schuster).

doi:10.1016/j.ijporl.2006.05.016

Conclusion: Automatic speech recognition may serve as a good means to objectify and quantify global speech outcome of children with cleft lip and palate.

1. Introduction

Cleft lip and palate (CLP) is the most common craniofacial malformation with an incidence of 1.0–2.21/1000 [1]. It can result in morphological and functional disorders, whereby one has to differentiate primary from secondary disorders. Primary disorders include problems of nutrition, breathing and mimic disorders. Speech and voice disorders [2–6] as well as conductive hearing loss that might also affect speech development [4,6], belong to secondary disorders. Speech disorders can persist even after adequate reconstructive surgical therapy. The characteristics of speech disorders are mainly a combination of different articulatory features, e.g. enhanced nasal air emissions, a shift in localization of articulation and a modified articulatory tension [2,7]. Voice disorders may occur as well, especially with boys [8]. Speech disorders affect not only the intelligibility but social competence and emotional development of an effected child [9–11]. One major aim of therapy is to enhance or normalize communication skills.

In clinical practice, articulation disorders are mainly evaluated by subjective tools. The simplest method is the auditive perception. For speech disorders standardized methods in different languages exist (e.g. PLAKSS in Germany). For Europe, a minimal standard for cleft care was published in 2000 [12]. Speech evaluation is performed subjectively and is therefore influenced by setting, experience, and the patient's compliance. However, procedures for speech assessment continue to vary considerably and the validity of results can still be questioned [13]. Objective means only exist for quantitative measurements of nasal emissions [14–16] and for the detection of secondary voice disorders [8,15,17]. But other specific or non-specific articulation disorders in CLP as well as a global outcome parameter of speech quality cannot yet be sufficiently quantified.

In this paper, we present a new technical procedure for the measurement and evaluation of speech disorders and compare the results obtained with subjective ratings of a panel of expert listeners.

2. Materials and methods

2.1. Patients

Acoustic files were recorded from 31 children, 4–16 years old with CLP (mean 10.1 ± 3.8 years, median 9.45), 9 girls aged 10.0 ± 4.67 years (median 9.96 years) and 22 boys aged 10.12 ± 3.11 years (median 9.41 years). Two had an isolated cleft lip, five an isolated cleft palate, 20 a unilateral cleft lip and palate and four a bilateral cleft lip and palate.

No indication of mental disability was documented in the patients' records. The examination was included into the regular out-patient clinic proceedings for all children with CLP. Informed consent had been obtained by all parents of the children prior to the examination. All children were native German speakers, some using a local dialect.

2.2. Speech samples

The children were asked to name pictures that were shown according to the PLAKSS test [18]. The German test consists of 99 words. It includes most possible phonemes of the German language in different positions (beginning, centre and end of a word). Furthermore, the children were asked to sustain all vowels and nasals and repeat six sentences from the "Heidelberger Rhinophonie Inventar" [19]. They consist of five sentences without nasal consonants and one only with nasal consonants ("Nenne meine Mama Mimi"). The speech samples were recorded with a close-talk microphone (dnt Call 4U Comfort-Headset) at a sampling frequency of 16 kHz and quantized with 16 bit.

2.3. Automatic speech recognition system

For objective measurement of functional limitation due to speech disorders, an automatic speech recognition system was applied, a state of the art word recognition system developed at the Chair for Pattern Recognition at the University of Erlangen-Nuremberg. In this study, the latest version as described in detail by Stemmer [20] was used. The evaluation was restricted on the PLAKSS-data only.

In general, a recognizer nowadays can handle spontaneous speech with mid-sized vocabularies of up to 10,000 words. A recognizer "knows" all common phonemes of the language. In a first acoustic analysis, a speech recognizer converts the speech signal into a set of features which represent the spoken utterance. Then, it compares spectral and temporal characteristics of the speech signal with an internal "dictionary" which was trained by acoustic speech samples. Each word in the "dictionary" is represented by a sequence of elementary Hidden Markov Models (HMM). These describe the likelihood of an acoustic signal to be identical with a certain phoneme. So the probability for each word can be obtained. In the end the recognized word chain is calculated as the most likely sequence of words that match a spoken text.

The current recognizer works polyphone-based which means that the acoustic characteristics of one phoneme are represented by HMMs with co-articulatory modulations of the phoneme. This is an extension of the triphone approach [21] that takes only the phoneme's predecessor and successor into account.

We used a so-called unigram language model to weight the outcome of each word model [22]. It was trained with the transliteration of the spoken test. Thus, the frequency of occurrence for each word in the used text was known to the recognizer. This helps to enhance recognition results by including linguistic information. The test set perplexity of the unigram language model is 94 which is high enough to represent low intelligibility as low word accuracies (WA, see below) and vice versa.

The speech recognition system had been trained with acoustic information from 23 male and 30 female children from a local school who were between 10 and 14 years old (mean age 11.88 ± 0.81 years, median age 12 years; 6.9 h of speech). To make the recognizer more robust, we added data from 85 male and 47 female adult speakers from all over Germany (2.3 h of spontaneous speech from the VERBMOBIL project [23]). 90% of the adult speakers had been younger than 40 years. The adults' data were adapted with "vocal tract length normalization" (VTLN, [24]) to better correspond to the acoustic properties of children's speech. Further improvement of speech recognition was enabled by MAP adaptation [25], i.e. we adapted the recognizer to each of the 31 children with CLP using all the data from the respective child.

We calculated the so called "word accuracy" (WA) of the children's speech data from the PLAKSS-test with the recognizer. WA shows how much a recognized sequence of words deviates from the spoken utterance, i.e. the percentage of correctly recognized words.

It is calculated with the following formula:

$$\mathrm{WA}\,[\%] = 100 \times \frac{(\mathrm{NC} - \mathrm{NW})}{N} \qquad (1)$$

NC, number of correctly recognized words; NW, number of wrongly inserted words by the recognizer; N, number of all spoken words.

Thus, if the child said "*Ball, Gabel und eine Blume*" and the output of the recognizer is "*Ball, eine Gabel oder eine Blume*" then the WA is 60%: The sentence consists of five words ($N = 5$), four words were correctly recognized (NC = 4), while one word was recognized wrongly ("*oder*") and one was as wrongly inserted ("*eine*", NW = 1). Using formula (1) WA is calculated as:

$$\mathrm{WA}\,[\%] = 100 \times \frac{(4 - 1)}{5} = 60$$

Note that the WA for children with normal speech is not expected to be 100% since the speech recognizer is trained on data from different speakers and different acoustic environments.

2.4. Subjective evaluation of substitute voice

A panel of three communication medicine professionals subjectively estimated the intelligibility of the children's speech while listening to a playback of the PLAKSS recordings. A five-point Likert scale (1 = very high, 2 = rather high, 3 = medium, 4 = rather low, 5 = very low) was applied to rate the intelligibility of all individual samples. The samples were divided into short passages from one up to four words. The intelligibility rating was calculated as the average of all ratings for each child and each rater.

2.5. Analysis and evaluation of the data

Statistical analysis was performed using scripts written in the Perl programming language. For the agreement computations between different raters on the one hand and raters/recognizer on the other hand, correlations and the weighted multi-rater kappa by Davies and Fleiss [26] were used. The multi-rater kappa allows comparing an arbitrary number of raters and weights the difference between the values of intelligibility or WA, respectively.

Several problems occur when comparing the ratings of the human experts and a speech recognition system. First of all, the human ratings were made on a quasi-continuous Likert scale while the word accuracy is a continuous measure within a completely different range. A mapping of the WA to the Likert scale had to be defined, since the kappa value can only be computed on discrete data. We rounded the experts' average intelligibility scores to the next integer and set thresholds defined as intervals on the WA scale on the recognizer's results manually, so that the difference between the experts' scores and the scores derived from the WA values was minimal.

3. Results

The total duration of the children's audio files was 120 min, consisting of 5330 words. The vocabulary of the word recognizer contained all 831 words occurring in the test data (99 unique words of the test, 34 words of the Heidelberger Rhinophonie Inventar, 266 additional adjectives and nouns which were used by the children to explain the pictures, and 432 additional representing reading errors). Due to the simple setup of the PLAKSS test the average turn length of fluently spoken following words is very short with 2.4 words. The recordings show a wide range in intelligibility. We only used the data from the PLAKSS test for the calculation of the WA. The highest WA is 75.8%, the worst is 1.2% (mean 48.0% $\pm$ 19.6%). Mean WA for each cleft type is given in Table 1. Age effects on WA could not be observed (correlation coefficient 0.14; $p > 0.5$).

Table 1 Word accuracies for children with different cleft types vary and show high standard deviation especially for those cleft types that include the palate

Cleft type	WA (mean and standard deviation)
Isolated cleft lip (n = 2)	72.9 $\pm$ 2.9
Isolated cleft palate (n = 5)	36.9 $\pm$ 23.3
Unilateral cleft lip and palate (n = 20)	48.0 $\pm$ 18.6
Bilateral cleft lip and palate (n = 4)	49.3 $\pm$ 9.4

The two children with isolated cleft lip show the highest WA.

Table 2 Inter-rater correlations between three experts (K, S, L) judging the intelligibility of 31 recordings of children with CLP

Rater	K	S	L
Other raters	0.92	0.93	0.93

For instance, the ratings of K correlate with 0.92 with the average of the ratings from S and L.

Table 3 Correlation coefficient between the objective speech evaluations by the polyphone-based recognizer concerning word accuracy (WA) and the perceptive evaluations by three experts concerning "intelligibility"

Rater	K	S	L	Mean
Automatic speech recognizer	−0.84	−0.93	−0.85	−0.90

The automatic speech evaluations mostly agree with evaluation by experts. Negative correlation coefficients result from the opposite scales for intelligibility (1, high intelligibility to 5, low intelligibility) and WA in percent.

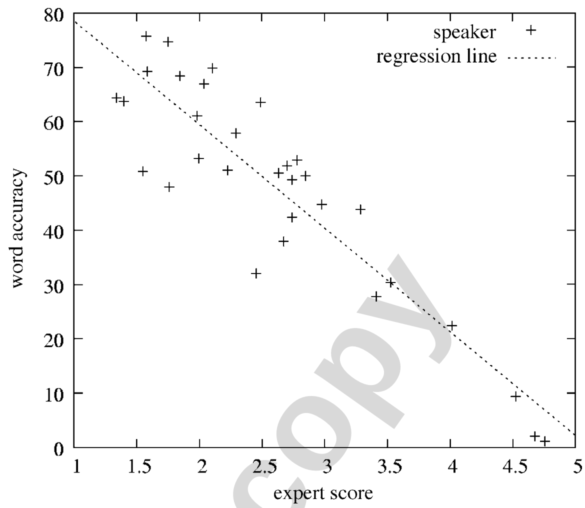

Fig. 1 Word accuracy (WA) vs. the average of the three experts' estimation of intelligibility of 31 children with cleft lip and palate with corresponding regression line. The children are ordered w.r.t. decreasing WA. It is clearly visible that there is a strong correlation (−0.90) between the human and the automatic evaluation results.

Subjective speech evaluation showed high consistency (see Table 2). The results for the correlations of the WA and the subjective speech evaluation are shown in Table 3. Considering the average of the raters, the WA for the recognizer has a correlation of −0.90 ($p < 0.01$) as shown in Fig. 1. The coefficient is negative because high recognition rates came from "good" speech with a low experts' score number and vice versa. The multi-rater kappa for the group of the three raters is 0.48. The kappa for the rater group versus the polyphone based recognizer is 0.52 (note that a result greater than 0.4 is said to represent fair agreement beyond chance [27]).

4. Discussion

Until now, no objective method of determining global speech outcome for individuals with CLP exists. Here, we present a new automatic objective measurement of speech quality: the recognition of spoken words, i.e. the word accuracy, by an automatic speech recognizer. The automatic speech evaluation was compared to the results of perceptive evaluations and shows high consistency between the perceptive ratings of intelligibility and the computed word accuracy. Until now, recordings of only 31 children with CLP have been evaluated yet, but there is every indication that the method can yield precise information in evaluating intelligibility on an expert's level.

Today, automatic speech recognition is used in many domains: for professional and private use as dictating machines, in call centres when a restricted vocabulary and "normal" voice quality and speech without background noise can be expected, and in the support of handicapped persons. Nevertheless, the technique often does not qualify for higher requirements such as low speech quality. Further research is in progress to enhance the possibilities of automatic speech recognition as it can also be of special interest as application in a medical field, i.e. in diagnostics, e.g. of speech disturbances. So far, it has been tested on severe voice disorders and has also shown promising results [28].

Speech recognition depends on five factors: on the speaker, on the speech (read speech, spontaneous speech), on the vocabulary, on the grammatical complexity or perplexity (average probability of words possibly following a sequence of others) and on the input medium [22]. The influence of most of these factors can be minimized when using a standard test such as the PLAKSS and stable recording setting as practiced in this study for diagnostic application. Thus, the speaker remains the most influencing factor.

We examined speech samples of children with cleft lip and palate after primary reconstructive surgical therapy with different speech outcome: most of them showing rather good intelligibility, and some showing very low intelligibility, thus including every occurrence of global speech outcome (see Fig. 1). The experts' evaluations of the intelligibility show high consistency and therefore can be a measure for the comparison with automatic evaluation. For this study, we applied an adapted speech recognizer for automatic speech evaluation that has previously proved to be adequate for "normal" speech samples, but covering also children without speech disorder. We used a polyphone-based recognizer as it considers also co-articulatory effects and even longer phoneme-sequences. This might be of special interest for this medical purpose since articulatory variances depend on previous and following phonemes.

According to previous studies about speech outcome of children with CLP, a high variability of intelligibility had to be assumed. In fact, the 31 children show considerable variation of the results of both the perceptive evaluation and the WA according to different speech outcome. Of course, this might also be due to different types of cleft [29]. As expected, the two children with cleft lip only show a higher mean WA than the others. There are rather slight differences between the other groups with high dispersion of values within these groups. Children with isolated cleft palate had lowest WA. No differences can be revealed between unilateral and bilateral CLP which is according to the literature [15]. However, as the number of patients examined is small for some groups, the data obtained in this pilot study with a methodical background would not allow drawing general conclusions for the relationship of cleft type and WA, yet. Furthermore, age effects on WA would have been expected [30] but could not be observed in our study. We suspect this is due to the small size of the study group. Further investigation will need to prove for age effects on a higher number of children per cleft type.

Of course, word accuracy is actually not the same as intelligibility. Both are influenced by voice quality, phonematic and morpho-syntactic structure, background noise, amplitudes and speaking velocity. But intelligibility includes also the "human factor". Even if one does not understand every word or syllable of a spoken sequence, the meaning can be understood by using contextual, pragmatic and prosodic elements of speech. We avoided this distortion of perceptive evaluation comparing to WA by using a test that mostly asks for single words. Though word accuracy does not include all aspects of speech quality, it is a strict criterion and obviously represents global speech outcome after CLP therapy in this evaluation setting.

The comparison of automatic evaluation with subjective evaluation by means of Likert scales is practical, but limited. The experts' estimation is demonstrated on a linear ordinate in Fig. 1, although the relation of distances between the values 1 and 5 could not be determined, i.e. the distance between "rather low" and "medium" is commonly not half the distance of "very low" to "medium". We would expect a bigger distance between the outer values (1 or 5) to their adjacent value than between the inner values (2, 3, and 4). Another characteristic of Likert scales is the reduced use of extreme scales, here 1 or 5. The discrepancy between the automatic speech recognizer and the experts' estimation in some recordings especially in the scales 2 and 3 (rather high, medium intelligibility) might arise from a reduced discriminatory power of Likert scales [31,32]. An automatic evaluation lacks this impreciseness and will lead to quick results without time-consuming perceptive evaluation.

The results of this pilot study demonstrate the automatic speech evaluation for German language. This ought to be alike in other languages, especially when testing only single words. For clinical application, further adaptation of the test is needed: synonyms of the test words have to be included into the speech recognizer's vocabulary as well as added words such as "This is a ... " Recordings of children without speech disorders of different age will serve as further training population for the speech recognizer and for the analysis of standard values. The determination of specific acoustic features that represent different articulation disorders such as hypernasality will lead to the automatic detection of these articulation disorders and will also quantify them. This will require a huge amount of recordings of both effected children and children without speech disorders. In future, automatic speech recognition systems might simplify speech evaluation for clinical and scientific purposes.

5. Conclusion

In this pilot study, automatic speech evaluation by a speech recognizer proved to be a precise means for research and clinical purpose in order to determine the global speech outcome of children with CLP. As a perspective, we assume that adaptation of the here presented technique will lead to further application to differentiate and quantify articulation disorders.

Acknowledgments

This work was supported by the Johannes and Frieda Marohn foundation at the Friedrich-Alexander University of Erlangen-Nuremberg. The authors are responsible for the content of this article. We thank PD Dr. med. Peter Kummer and Dr.-Ing. Jörg Lohscheller for their valuable comments on this article.

References

[1] A. Derijcke, A. Eerens, C. Carels, The incidence of oral cleft: a review, Br. J. Oral Maxillofac. Surg. 34 (6) (1996) 488—494.
[2] A. Harding, P. Grunwell, Characteristics of cleft palate speech, Eur. J. Disord. Commun. 31 (4) (1996) 331—357.
[3] M.C. Pamplona, A. Ysunza, M. Gonzalez, E. Ramirez, C. Patino, Linguistic development in cleft palate patients with and without compensatory articulation disorder, Int. J. Pediatr. Otorhinolaryngol. 54 (2—3) (2000) 81—91.
[4] F. Rosanowski, U. Eysholdt, Phoniatric aspects in cleft lip patients, Facial Plast. Surg. 18 (2002) 197—203.
[5] B. Schönweiler, R. Schönweiler, R. Schmelzeisen, Untersuchungen zum Spracherwerb bei Kindern mit Spaltbildungen, Folia Phoniatr. Logop. 48 (1996) 92—97.
[6] R. Schönweiler, B. Schönweiler, R. Schmelzeisen, Hörvermögen und Sprachleistungen bei 417 Kindern mit Spaltfehlbildungen, HNO 42 (1994) 691—696.
[7] U. Wohlleben, Die Verständlichkeitsentwicklung von Kindern mit Lippen-Kiefer-Gaumen-Segel-Spalten, Schulz-Kirchner Verlag, Idstein, Germany, 2004.
[8] K.M. Van Lierde, S. Claeys, M. De Bodt, P. Van Cauwenberge, Vocal quality characteristics in children with cleft palate: a multiparameter approach, J. Voice 18 (3) (2004) 354—362.
[9] K.L. Chapman, K. Tecco Graham, J. Gooch, C. Visconti, Conversational skills of preschool and school-age children with cleft lip and palate, Cleft Palate Craniofac. J. 35 (6) (1998) 503—516.
[10] L.C. Richman, M. Eliason, Psychological characteristics of children with cleft lip and palate: intellectual, achievement, behavioral and personality variables, Cleft Palate J. 19 (4) (1982) 249—257.
[11] T. Millard, L.C. Richman, Different cleft conditions, facial appearance, and speech: relationship to psychological variables, Cleft Palate Craniofac. J. 38 (1) (2001) 68—75.
[12] W.C. Shaw, G. Semb, P. Nelson, V. Brattström, K. Mölsted, B. Prahl-Andersen, The Eurocleft Project 1996—2000. Standards of care for cleft lip and palate in Europe, in: European Commission Biochemical and Health Research, IOS Press, Amsterdam, 2000.
[13] A. Lohmander, M. Olsson, Methodology for perceptual assessment of speech in patients with cleft palate: a critical review of the literature, Cleft Palate Craniofac. J. 41 (1) (2004) 64—70.
[14] C. Küttner, R. Schönweiler, B. Seeberger, R. Dempf, J. Lisson, M. Ptok, Objektive Messung der Nasalanz in der deutschen Hochlautung, HNO 51 (2003) 151—156.
[15] K.M. Van Lierde, M. De Bodt, J. Van Borsel, F.L. Wuyts, P. Van Cauwenberge, Effect of cleft type on overall speech intelligibility and resonance, Folia Phoniatr. Logop. 54 (3) (2002) 158—168.
[16] T.T. Hogen Esch, P.H. Dejonckere, Objectivating nasality in healthy and velopharyngeal insufficient children with the nasalance acquisition system (Nasal View). Defining minimal required speech tasks assessing normative values for Dutch language, Int. J. Pediatr. Otorhinolaryngol. 68 (8) (2004) 1039—1046.
[17] T. Bressmann, R. Sader, M. Merk, W. Ziegler, R. Busch, H.F. Zeilhofer, H.H. Horch, Perzeptive und apparative Untersuchung der Stimmqualität bei Patienten mit Lippen-Kiefer-Gaumenspalten, Laryngorhinootologie 77 (12) (1998) 700—708.
[18] A.V. Fox, PLAKSS—Psycholinguistische Analyse kindlicher Sprechstörungen, Swets & Zeitlinger, Frankfurt a.M., 2002. now available from Harcourt Test Services GmbH, Frankfurt a.M.
[19] W. Heppt, M. Westrich, B. Strate, L. Möhring, Nasalanz—ein neuer Begriff der objektiven Nasalitätsanalyse, Laryngo-Rhino-Otol 70 (1991) 169—228.
[20] G. Stemmer, Modelling Variability in Speech Recognition, Logos-Verlag, Berlin, 2005.
[21] E.G. Schukat-Talamazzini, H. Niemann, Das ISADORA-System—ein akustisch-phonetisches Netzwerk zur automatischen Spracherkennung. In: Radig, B., (Ed.), Mustererkennung 1991, vol. 290 Informatik Fachberichte Springer, Berlin, 1991. pp. 251—258.
[22] F. Gallwitz, H. Niemann, E. Nöth, Speech recognition—state of the art, applications, and future prospects, Wirtschaftsinformatik 41 (6) (1999) 538—547.
[23] W. Wahlster (Ed.), Verbmobil: Foundations of Speech-to-Speech Translation, Springer, Berlin, 2000.
[24] G. Stemmer, C. Hacker, S. Steidl, E. Nöth, Acoustic normalization of children's speech, in: Proceedings of European Conference on Speech Communication and Technology, vol. 2, Geneva, Switzerland, 2003, pp. 1313—1316.

[25] J. Gauvain, C. Lee, Maximum a-posteriori estimation for Gaussian mixture observations of Markov chains, IEEE Trans. Speech Audio Process. 2 (1994) 291—298.
[26] M. Davies, J.L. Fleiss, Measuring agreement for multinomial data, Biometrics 38 (1982) 1047—1051.
[27] J.L. Fleiss, Statistical Methods for Rates and Proportions, second ed., John Wiley & Sons, New York, 1981.
[28] M. Schuster, T. Haderlein, E. Nöth, J. Lohscheller, U. Eysholdt, F. Rosanowski, Intelligibility of laryngectomees' substitute speech: automatic speech recognition and subjective rating, Eur. Arch. Otorhinolaryngol. 263 (2) (2006) 188—193.
[29] M.-L. Haapanen, Cleft type and speech proficiency, Folia Phoniatr. Logop. 46 (1994) 57—63.
[30] J.G. Wilpon, C.N. Jacobsen, A study of speech recognition for children and the elderly, Proc. ICASSP (1996) 349—352.
[31] R. Likert, A technique for the measurement of attitudes, Arch. Psychol. 140 (1932) 1—55.
[32] P.E. Spector, Choosing response categories for summated rating scales, J. Appl. Psychol. 61 (1976) 374—375.