International Journal of Pediatric Otorhinolaryngology 76 (2012) 362–369

Contents lists available at SciVerse ScienceDirect

International Journal of Pediatric Otorhinolaryngology

journal homepage: www.elsevier.com/locate/ijporl

Automatically evaluated degree of intelligibility of children with different cleft type from preschool and elementary school measured by automatic speech recognition

Maria Schuster [a,*], Andreas Maier [b], Tobias Bocklet [c], Emeka Nkenke [d], Alexandra Holst [e], Ulrich Eysholdt [c], Florian Stelzle [d]

[a] Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital, Ludwig-Maximilians-University Munich, Marchioninstrasse 15, D-81377, Munich, Germany
[b] Pattern Recognition Lab, Technical Faculty, Friedrich-Alexander-University Erlangen-Nuremberg, Martensstrasse 3, D-91058 Erlangen, Germany
[c] Department of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Bohlenplatz 21, D-91054 Erlangen, Germany
[d] Department of Oral and Maxillofacial Surgery, University Hospital Erlangen, Glücksstrasse 11, D-91054 Erlangen, Germany
[e] Clinic for Orthodontics, University Hospital Erlangen, Glückstrasse 11, D-91054 Erlangen, Germany

ARTICLE INFO

Article history:
Received 25 September 2011
Received in revised form 9 December 2011
Accepted 10 December 2011
Available online 10 January 2012

Keywords:
Cleft lip and palate
Cleft palate
Cleft lip
Automatic speech recognition
Speech intelligibility
Objective evaluation

1. Background

Cleft lip and palate is the most common orofacial deformity with a prevalence of about 1.0–2.2 per 1000 newborns [1,2]. The morphological alterations might lead to functional limitations in breathing, hearing and speech [3]. Psycho-social well-being and school success may be affected as well [4–6]. Speech disorders are often described for children with clefts including the palate and might lead to diminished intelligibility of speech. In more than 50% of children with unilateral cleft lip and palate, intelligibility was rated as not normal in a study by Sell et al. [7].

In the International Classification of Functioning, Disability, and Health, intelligibility is related to Body Functions considering articulation and voice and to Activities and Participation when seen as the precondition for verbal communication. According to Schiavetti, intelligibility is defined "as the match between the intention of the speaker and the response of the listener to the speech passed through the transmission system. When all of the words in the listener's response list match all the words intended to be produced by the speaker, speech intelligibility is perfect. When none of the words in the listener's response list match the words intended to be produced by the speaker, speech intelligibility is zero. In between these extremes of perfect and zero speech intelligibility lies a continuum on which we may quantify the degree to which the response list of the listener matches the intended productions of the speaker." [8].

Articulation errors and velopharyngeal dysfunction can highly affect speech intelligibility [9,10]. Concerning patients with cleft lip and palate, articulation errors, altered articulation tension as well as altered nasal resonance and nasal emissions due to velopharyngeal dysfunction are typical speech limitations even after adequate surgical and non-surgical treatment [4,11,12]. Sometimes, mixed hyper- and hyponasality are present in incomplete velopharyngeal closure and restricted nasal air flow due to nasal obstruction [13]. Velopharyngeal dysfunction is known to be of major importance for the intelligibility of speech [14,15]. Hence, children with a cleft affecting the palate have a higher risk for reduced intelligibility. Furthermore, anterior alterations of the lip and the alveolus might lead to altered interlip coupling [16] and to articulation errors and therefore might decrease intelligibility [17]. Accordingly, there were found long-lasting subclinical effects on upper-lip kinematics and interlip coupling, but suggested that the speech production system seems to have a relatively high tolerance for this type of movement variability [16]. Hence, in average, we would expect children with cleft including the palate having a diminished intelligibility whereas isolated clefts of the lip should not have a lack of intelligibility in relation to children at the same age without clefts.

The evaluation of intelligibility has been performed with different methods [9]. One can differentiate categorical assessment using Likert scales or quasi-linear analogue scales and quantifying assessment judging word by word of a word chain [18]. An international working group founded to develop universal parameters to report speech outcomes propose a quasi-linear assessment. They set a perceptual procedure to evaluate speech outcomes [19] in conversational speech and chose the term

* Corresponding author. Tel.: +49 89 70953861; fax: +49 89 70956869.
E-mail address: maria_elke.schuster@med.uni-muenchen.de (M. Schuster).

0165-5876/$ – see front matter
doi:10.1016/j.ijporl.2011.12.010

"understandablity". Understandability should be evaluated on a 3-point Likert scale with defined scales distinguishing between 0 (speech is always easy to understand), 1 (speech is occasionally hard to understand), 2 (speech is often hard to understand), and 3 (speech is hard to understand most or all of the time). According to the UK Clinical Standards Advisory Group of 2001, a sample of conversation should be judged how intelligible a child's speech would be to a stranger on a range from 0 (normal), 1 (different from other children's speech but enough to cause comment), 2 (different enough to cause comment but possible to understand most speech), 3 (only just intelligible to strangers), and 4 (impossible to understand) [7]. For this procedure, they demonstrated a perfect inter-rater agreement for trained examiners on speech and video data of the children examined. The children's age was restricted to 5- and 12-year old; an age when speech normally should be intelligible also for strangers.

Both measures for intelligibility and understandability proposed by the UK group and the international group give a fairly precise definition of the ratings. Nevertheless, ratings might be influenced by different factors. Referring to conversational speech, understandability and intelligibility on the one hand include articulation skills but also semantic, lexical, grammatical, and speech pragmatic skills of those who speak. On the other hand, it depends on the listener. When measuring intelligibility, one has to take into account both the "sender" and the "receiver". To reduce the influence of different receivers' skills, intelligibility is often judged by a panel of listeners, experts, or non-experts. For a standardized assessment of intelligibility these judgments are averaged and the mean rating is considered to be objective.

There are many factors that can affect perceptual evaluation of intelligibility. Expertise with altered speech has been shown to have an influence on the judgment [18,20]. Even for experienced speech therapists, considerable intra-individual differences have been reported [21]. Moreover, different examination settings and perceptual rating methods complicate multi-center studies or comparisons between different centers. Discussing the restrictions on the assessment of intelligibility, one also needs to focus on age effects especially for preschool children when articulation skills are still developing. Hence, the age and the variations of normal speech development during preschool period have to be taken into account when judging speech intelligibility of children. This may result in a huge effort for scientific settings when speech is assessed perceptually by several listeners considering age effects. For all these reasons – the physiologic range of articulations skills in preschool children and the speech assessment methods – it is difficult to draw standardized comparisons of the impact of different therapy methods or types of clefts on intelligibility independently from age.

Recently, a new method for the quantification of intelligibility has been developed. It is independent from listener perception but depends on automatic speech recognition (ASR). An ASR system is a software that analyzes the acoustic properties of spoken speech. The recognition is based on a stochastic models, called Hidden Markov Models (HMM). Unlike perceptually performed tests, they allow for the consistent quantification of correctly understood words of a word sequence, called the word recognition rate (WR) without showing training effects. As it is not totally similar to intelligibility, the resulting WR is called "degree of intelligibility". Children with CLP were previously tested with this method and results demonstrated a high correlation to the single word intelligibility rating of a panel of experts [15,22,23]. When compared to perceptual evaluation of intelligibility, ASR reached 0.90 correlation with three experienced raters and is near to the inter-rater correlation between the three (0.92–0.93).

It was the aim of this study to evaluate the influence of different cleft types on the degree of speech intelligibility analyzed by an ASR system on children at the age from 3 to 10. As a precondition, we analyze the impact of age on WR on a comparison group. Cleft-type-dependent results are compared to the results of the comparison group. After age correction according to the slope of the comparison group's WR, the mean cleft group variations are presented within different age groups and for sampled cleft groups. Hence, the impact of speech development during childhood on intelligibility is taken into account.

2. Patients and methods

2.1. Patients

Participants with different cleft types (111 participants in total) (Table 1) were recorded during regular outpatient care while naming pictures of a standard test for articulation. Inclusion criteria comprised cleft lip and palate (CLP) with unilateral cleft lip and palate (UCLP), and bilateral cleft lip and palate (BCLP), isolated cleft palate (CP), isolated cleft lip (CL), cleft lip and alveolus (CLA) and Pierre Robin sequence with cleft palate with normal cognitive development. Exclusion criteria were children with learning disabilities diagnosed by using the Culture Fair Intelligence Test [24], children with hearing impairment, and non-native German speakers. About 90% of the participants spoke a local dialect.

Primary surgery was performed according to the Erlangen Cleft Center: Cleft lip is closed (Tennison-Randall) at the age of 5–6 months, cleft palate (von Langenbeck and Widmaier) at the age of 10–13 months. Osteoplasties with bone grafts, if needed, were performed at the age of 8–12 years. None of the participants had a secondary palatoplasty for velopharyngeal dysfunction before the examination. Informed consent was given by participants' parents. The study was approved by the local ethical committee (Registration No. 3473) according to the declaration of Helsinki.

2.2. Comparison group

Seven-hundred-sixty participants from preschool and school (aged 7.94 ± 1.68 years) throughout Germany served as controls (Table 2). In Germany, preschool period usually ends up at the age of six years. Recordings were performed in the same way as for the cleft groups' children with the same headset in a quiet room. Speech and hearing were not evaluated prior to the speech recordings. WR has been shown to be independent from the dialects of the children, including the local dialect (franconian), from southwest Germany (alemannic), central south (swabian), central north (standard German as spoken in Hannover), and east (saxon) [25].

Table 1
Description of 111 children with different types of clefts.

	Cleft lip CL	Cleft lip and alveolus CLA	Cleft palate CP	Unilateral cleft lip and palate UCLP	Bilateral cleft lip and palate BCLP
# total	17	9	30	44	11
# girls; # boys	7; 10	5; 4	19; 11	18; 26	5; 6
Mean age ± standard deviation	6.6 ± 2.0	7.7 ± 1.6	6.3 ± 2.1	7.2 ± 2.2	7.0 ± 2.2