An Automatic Screening Test for Preschool Children: Theory and Data Collection

Tobias Bocklet
Chair of Pattern Recognition
Martensstr. 3
91058 Erlangen, Germany
tobias.bocklet@informatik.uni-erlangen.de

Cordula Winterholler
Berufsfachschule für Logopädie
Waldstrasse 14
91054 Erlangen
cordula.winterholler@uk-erlangen.de

Andreas "Magnet" Maier
Department of Pedaudiology and Phoniatrics
Bohlenplatz 21
91054 Erlangen, Germany
andreas.maier@informatik.uni-erlangen.de

Maria Schuster
EuromedClinic
Europa-Allee 1
90763 Fürth, Germany
Maria.Schuster@euromed.de

Elmar Nöth
Chair of Pattern Recognition
Martensstr. 3
91058 Erlangen, Germany
noeth@informatik.uni-erlangen.de

ABSTRACT

In this paper the first version of a recording and evaluation system for an ongoing research project is presented: an automatic school enrollment screening test for children in preschool age. The screening test focuses on different aspects of the development of children's speech and language and has a duration of about 15 minutes. It is based on parts of already standardized tests. For a first version these tests are digitized, and adapted to an existing computer system for speech assessment so that recordings can be performed in local preschools. The recordings take place while a speech therapist is assisting the child. The described version focuses on the data collection and an subjective speech and language assessment of the attendant speech therapist. Children, that are awaiting school enrollment in fall, have been selected for this project. It is planed to collect a total amount of more than 500 children within the next 2 years. It is planed to record the children continuously over a period of 3-4 years at an interval of 12 months in order to measure their development automatically.

Categories and Subject Descriptors

H.4.0 [**Information System Application**]: General

General Terms

Languages,Human Factors,Design

1. INTRODUCTION

ICMI-MLMI'09 Workshop on Child, Computer and Interaction November 5, 2009, Cambridge, MA, USA

Every child has to undergo a screening test right before school enrollment in Germany. There have been several efforts to build up a standardized test for different aspects of speech and language development. But in general the test is performed by speech therapists of the local health authorities, which all use their own test. The goal of a recent research project is to evaluate the ability of estimating a child's speech automatically. The benefits of an automatic system are obvious: objective, time saving and reproduceable. Different requirements have to be fulfilled by an automatic system. The test has to be attractive to children and should not exceed a duration of 15 minutes, in order to keep the children interested and vigilant during the entire test. The final system has to run on a standard computer system and should be easy-to-use, so that also trained assistants can perform recordings of the test. We have also plans to create a system, that can be used by a child without the attendance of an assistant.

Speech and language development plays an important role in the infant development. It has been shown in different studies that speech and language development reflects also other aspects of a child's development [1, 2]. Parts of already existing tests were selected to test different aspects of speech and language development. Due to different stages during the development of the children these tests are all age-dependent. Only existing age-based tests were used for the automatic version of the enrollment test. The tests were selected by skilled speech therapists of the Department of Pedaudiology and Phoniatrics of the University Hospital in Erlangen. The following language skills are very important for children in school enrollment age:

- Auditory memory
- Differentiation between similar phones
- Grammatical understanding
- Active word and grammatical usage

In this paper, the test itself is described as well as some details on the recording system and the ongoing data collec-

tion. In Section 2 the theoretical background of the different subtests and their coherence with other aspects of infant development is described. In Section 3 we give some information about the recording system. A first version of the system is created in order to allow recordings and subjective evaluations. The recordings are performed by speech therapists, which are familiar with children's speech. A detailed description of the data collection is described in Section 4. The paper concludes with a summary (Section 5) and a short outlook (Section 6).

2. DESIGN OF THE SCREENING TEST

The different aspects described in Section 1 are very important to measure the stage of speech and language development of a child. For each aspect parts of different standardized tests are used. The theoretical background of these tests is described in this section. Normally the tests are performed by a speech therapist who evaluates the language/speech skills at the same time.

2.1 Auditory memory

2.1.1 Theory

Healthy two year old children have a vocabulary of about 150-180 words [9]. With this vocabulary they have the ability to produce and to understand simple sentences. With the age of three the vocabulary reaches a number of about 500 words. From three to five the child learns more complex structures of speech: semantics (meaning of sentences), morphology (word formation) and syntactics (formation of sentences). This part is called *auditory memory*. The auditory memory contains different components. These components are tested by the so-called SETK (*Sprachentwicklungstest für Kinder*). For learning different structures of speech, it is very important to keep units of speech in mind. With an ongoing development of children, the ability to represent phonemes in the auditory memory is improved, so that new words can be learned faster and more precisely. This plays an important role for the ability to learn regularities in word formation. This representation ability is most suitably tested with pseudo words. The reproduction of pseudo words aims at the capability to combine different phonemes to an unknown word. Problems in repeating these units of speech could indicate brain deficiencies of an assessed child [6]. Example of such pseudo words are *Billop* and *Ronterklabe*.
Example of such pseudo words are *Billop* and *Ronterklabe*.
Another part of the auditory memory is the memory of sentences. Here the semantic and syntactic component of the auditory memory plays an important role. The grammatical knowledge has to be represented in the memory of the child. The better this capability is developed, the more complex sentences can be repeated by a child. For testing this ability, it is important to use rather long sentence with 6 to 10 words. Long sentences can not be kept in memory completely. So the child has to make use of the linguistic characteristics of the sentences in order to repeat it. In order to fade out the general knowledge of a child, abnormal sentences are also used. These are sentences that are syntactically and morphologically correct, but do not have a meaning. An example of such a sentence is "*Bevor der Goldfisch hinfällt, frisst er aus dem Fenster*" ("*Before falling down, the goldfish gorges out of the window*").

2.1.2 Subjective Evaluation

In this subtest different aspects will be tested. The evaluation of the reproduction of pseudo words is first roughly evaluated on word level, i.e., was a specific pseudo word uttered correctly. This is done during the speech recording. A more detailed analysis focuses on an assessment on phone level. Which phonemes were uttered in a wrong way and what phoneme conjunctions caused pronunciation difficulties.
The memory of sentences is assessed in a similar way. Right during the speech recording the speech therapist evaluates the sentence completely and marks, if a sentences was repeated completely. A more detailed analysis on word level focuses on the sentences repeated incompletely: What type of sentences are repeated incompletely (real sentences or abnormal sentences) or which components of the sentence have not been repeated.

2.2 Differentiation between similar phones

2.2.1 Theory

The basic concept of this test is a relation between dysgraphia or dyslexia and the ability to differentiate between similar phones. It has been shown in [8] on 141 children that children who lack to differentiate between phones have a high risk of dyslexia and dysgraphia. The used test is the so-called HVS (*Heidelberger Vorschulscreening*) [3]. In this part of test the child has to differentiate pairs of similar words or pseudo words. This test is performed twice: The child has to tell if he or she heard two different utterances and in a second run the utterances have to be repeated by the child. It contains 10 items, i.e., pairs of words or pseudo words where only one phoneme might be different. These phonemes differ in the way of articulation, e.g., "*Pärchen/Bärchen*", "*aka/aga*" or in the place of articulation, e.g., "*Kragen/tragen*", "*ata/aka*".

2.2.2 Subjective Evaluation

This subtest is performed twice. At the first time the child has to detect differences between the pairs of word. This it is captured right during recording. The second run is accomplished after the Trog-D test is performed. This time the child has to repeat the words. During recording time the speech therapist registers if the word pairs are uttered correctly. This allows us to filter out pairs, where the child stated (not) to hear any difference in the first run but repeated them (in)correctly in the second run.
A more detailed offline evaluation focuses on the phone level. It is assessed which phonemes caused problems in particular, i.e., phonemes with different way of articulation or phonemes with different place of articulation. In this context also the class of the wrongly pronounced phone is determined, i.e., plosives, nasals, fricatives, vowels or diphthong.

2.3 Grammatical understanding

2.3.1 Theory

The language understanding is very important for a child's development. A persistent dysfunction of language understanding has serious consequences to the personality of a child. Examples are problems in school or resultant disorders of personality and behavior [10]. An early detection before school enrollment is very important. The estimation

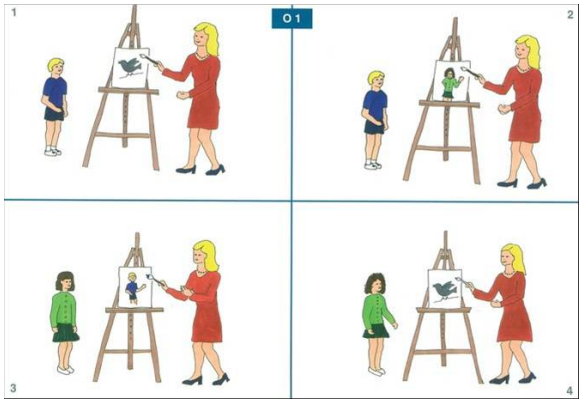

Figure 1: Testing of the language understanding by the so-called Trog-D test

of the language understanding is a difficult task. According to [10] the conclusion that a child understands what he or she is uttering is not always true. So it is useful to assess the language understanding of children without acoustic output. This is achieved by a German version of the TROG test [4], the so-called Trog-D (*test for reception of grammar -Deutsch*) [5]. TROG-D investigates the understanding of different grammatical structures, which are differentiated by flexion, syntax and function words. For this test a sentence is read to a child. The child has to choose the correct picture out of four pictures, i.e., the picture that represents the correct meaning of the sentence. For each grammatical structure a block of four sentences is presented. Structures that cause special difficulties can be found. Additionally the kind of difficulty can be captured by analyzing the chosen deflector pictures. The original test contains 84 items. For our version we took 16 of them. An example of one item is given in Figure 1. The corresponding sentence is: "*Die Frau malt dem Jungen das Mädchen*" ("*The woman paints the girl for the boy*").

2.3.2 Subjective Evaluation

In this subtest no questions have to be answered and no utterances have to be repeated by the children. That means no speech recordings are available for this subtest. The evaluation is performed online, right during recording. It is captured which of the four pictures was selected. for each item there is one main deflector, which if very similar to the correct picture and 2 deflectors, which describe a completely different meaning. It can be assessed which kind of deflector was selected and which grammatical structures caused special problems.

2.4 Active word and grammatical usage

2.4.1 Theory

The last aspect of speech that is tested, is the ability to create a short story. Therefore a comic strip with six pictures is shown to the children. The comic strip is shown in Figure 2. The children have to explain the contents of the pictures and have to form a short tale. It is important that the children can describe the pictures accurately without any help of the speech therapist. Accurately means in this case, that the plot, all characters, the scene and all involved objects are introduced and combined to a fluent text.
Right after the comic strip the children are asked to answer different questions about the plot of the different pictures. It is tested if the child is able to answer the question with an adequate phrase and if some words are known to the child, e.g., purse, cashier or (cooling) shelf.

2.4.2 Subjective Evaluation

In this test the active grammatical usage and in a way the used vocabulary is tested. The speech therapist judges if the child is able to create a flowing story with complete sentences. Therefore the beginning of the story is very important. If children are not able getting started the speech therapists intervenes and assists the child.

3. SYSTEM DESCRIPTION

The screening test is integrated into PEAKS *Platform for the Evaluation and Analysis of all Kinds of Speech Disorders* [7]. PEAKS is a client-server-based recording and assessment tool, which is developed in our lab. The client is designed as a JAVA Applet and runs in any web-browser on every standard computer system with JAVA Version 1.6.0 or above. The communication between client and server is encrypted using 256-bit SSL sockets and ensures the privacy of the children's data. In case of a recording, the speech therapist logs into PEAKS via username and password, selects the screening test and starts to record the different subtests. For each subtest instructions are given to the child in form of audio files. The subtests SETK, HLAD and TROGD require the child to repeat or to answer certain words or sentences. These sentences and the instructions have been recorded by an experienced speech therapist and are played to the child during the data collection. By these sentences each child listens to the same voice, which assures similar recoding conditions for each child. These recordings also allow recordings by trained staff in a later stage of the project. The speech therapist has the opportunity to repeat an instruction as many times as necessary. This is required, because the recordings take place directly in the kindergarten, where the noise level can be quite high sometimes. It is not possible to skip different items of subtest, but the speech therapist can skip complete subtests if necessary.
The final recordings are then transferred to the server and inserted into a database. In order to allow also offline recordings, i.e., when no Internet connection is available, a stand-a-lone recording tool was created. This is also Applet-based and transfers the conducted recordings to the server, when an Internet connection is available.

4. DATA COLLECTION

The data collection takes place in different preschools in the area of Erlangen, Germany. In cooperation with the local health authority 23 preschool have been selected, where roughly 600 children are awaiting school enrollment this year. Informed consent has been collected from the parents of the children. Together with this, also a questionnaire was handed out to the parents. The questions regarded different speech and language related aspects of the child's present development and aspects, which most likely influence the development of a child. It is asked if local dialect is spoken in the family or if the parents have different mother tongues.

Figure 2: Comic strip: Children have to tell a short tale about the content of the pictures

The recordings have been performed directly in the preschools. Skilled speech therapists tested each child separately. The duration of a test varied from 10 to 16 minutes.

5. SUMMARY

In this paper we introduced different aspects of a new project, which aims at an automatic screening for children in school enrollment age. Different aspects of the speech and language development of children are addressed within the screening test. For each aspect a different subtest is used and the theoretical background behind each test is described. These subtests had to be adapted to a computer, which required digital versions of different auxiliary material and the recording of instruction sentences. The performed subjective evaluations for each subtest are described.

6. OUTLOOK

Currently the data collection is in progress. 150 children have been recorded so far. The main challenges in this project are the evaluation on phone/sub-word level and the evaluation of the comic strip, where no predefined phrases are uttered. The next step is the automatic evaluation of the SETK subtest.

7. ACKNOWLEDGMENTS

We thank Ine Heinrich for her help in the arrangement of the different parts of the test. We also thank Milena Bäcker, Katharina Mittelhammer and Julia Kümmel for the accomplishment of most of the recordings. The authors thank Andrea Schädel for speaking the instruction and description sentences.

8. REFERENCES

[1] M. Brunner, C. Baeumer, S. Dockter, F. Feldhusen, P. Plinkert, and U. Proeschel. Heidelberg Phoneme Discrimination Test (HLAD): Normative Data for Children of the Third Grade and Correlation with Spelling Ability. *Folia Phoniatrica et Logopaedica, International Journal of Phoniatrics*, 60:157–161, 2008.

[2] M. Brunner, B. Pfeiffer, C. Heinrich, and U. Pröschel. Development and testing of the Heidelberg pre-school screening for auditory perception and speech processing (HVS). *Folia Phoniatrica et Logopaedica, International Journal of Phoniatrics*, 57:48–58, 2005.

[3] M. Brunner, B. Pfeiffer, and J. Troost. *Heidelberger Vorschulscreening zur auditiv-kinästhetischen Wahrnehmung und Sprachverarbeitung (HVS).* Hogrefe-Verlag für Psychologie, Göttingen, Germany, 2001.

[4] D.V.M.Bishop. *TROG. Test for the Reception of Grammar.* Thames Testing Company, 1989.

[5] A. V. Fox. *TROG-D. Test zur Überprüfung des Grammatikverständnisses. Handbuch.* Das Gesundheitsforum. Schulz-Kirchner Verlag, Idstein, Germany, 2006.

[6] H. Grimm. *SETK 3-5. Sprachentwicklungstest für drei - bis fünfjährige Kinder. Diagnose von Sprachverarbeitungsfähigkeiten und auditiven Gedächtnisleistungen.* Hogrefe-Verlag für Psychologie, Göttingen, Germany, 2001.

[7] A. Maier, T. Haderlein, U. Eysholdt, F. Rosanowski, A. Batliner, M. Schuster, and E. Nöth. PEAKS - A system for the automatic evaluation of voice and speech disorders. *Speech Communication*, 51(5):425–437, 2009.

[8] L. Möhring, H. Schöler, M. Brunner, and U. Pröschel. Zur Diagnostik struktureller Defizite bei Lese-Rechtschreib-Störungen in der klinischen Arbeit: Beziehungen zwischen verschiedenen Leistungsindikatoren. *Laryngorhinootologie*, 82:83–91, 2003.

[9] L. Rescorla, J. Mirak, and L. Singh. Vocabulary growth in late talkers: Lexical development from 2,0 to 3,0. *Journal of Child Language*, 27:293–311, 2000.

[10] P. Schmitz and A. V. Fox. Sprachverstehenstests im Deutschen unter besonderer Berucksichtigung des TROG-D.(Theorie und Praxis). *Forum Logopaedie*, 2007.