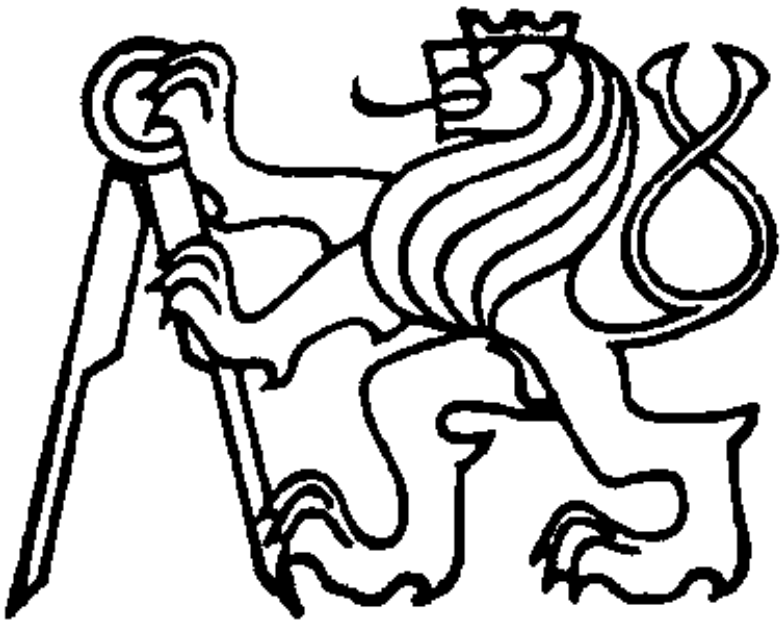
CZECH TECHNICAL UNIVERSITY IN PRAGUE



DOCTORAL THESIS STATEMENT

Czech Technical University in Prague

Faculty of Electrical Engineering

Department of Circuit Theory

Pavel Grill

SPECIFIC LANGUAGE IMPAIRMENTS AND POSSIBILITIES OF CLASSIFICATION AND DETECTION FROM CHILDREN'S SPEECH

Ph.D. Programme: Electrical Engineering and Information Technology

Branch of study: Electrical Engineering Theory

Doctoral thesis statement for obtaining the academic title of "Doctor", abbreviated to "Ph.D."

Prague, August 2015

The doctoral thesis was produced in full-time~~/part-time/combined~~ manner *

Ph.D. study at the department of Circuit Theory of the Faculty of Electrical Engineering of the CTU in Prague

Candidate: Pavel Grill
Establishment: Department of Circuit Theory
Address: Obětí okupace 612, 271 01 Nové Strašecí

Supervisor: prof. Ing. Jana Tučková, CSc.
Department of Circuit Theory
Faculty of Electrical Engineering of the CTU in Prague
Technická 2, 166 27 Prague 6

Opponents:

The doctoral thesis statement was distributed on: …….

The defence of the doctoral thesis will be held on …….. **at** …… …… **a.m./p.m. before the Board for the Defence of the Doctoral Thesis in the branch of study** ……. **in the meeting room No.** …… **of the Faculty of Electrical Engineering of the CTU in Prague.**

Those interested may get acquainted with the doctoral thesis concerned at the Dean Office of the Faculty of Electrical Engineering of the CTU in Prague, at the Department for Science and Research, Technická 2, Praha 6.

Chairman of the Board for the Defence of the Doctoral Thesis
in the branch of study *(to be specified)*
Faculty of Electrical Engineering of the CTU in Prague
Technická 2, 166 27 Prague 6.

*) *leave out as appropriate*

CONTENTS

CURRENT SITUATION OF THE STUDIED PROBLEM

Introduction

The ability to communicate via spoken language is one of the most important human attributes. Although there are several other means of communication, speech is difficult to replace in everyday life. Inabilities to communicate using speech can isolate individuals from society. Isolation resulting from speech impairments is significant for children with specific language impairments.

This work is part of an ongoing research project that integrates results from the fields of neurology, psychology, logopedics, MRI tractography and speech processing. The aim of this research is to further advance the diagnosis of specific language impairment in children and to develop efficient treatments for this disorder.

Our Laboratory of Artificial Neural Network Applications (LANNA) [1] in the Czech Technical University in Prague collaborates on a project with the Department of Paediatric Neurology, 2nd Faculty of Medicine of Charles University in Prague and with the Motol University Hospital, which focuses on the study of children with SLI. One aim of this project is to record children's speech (the normal and pathological speech of SLI patients) and to obtain data about SLI and speech disorders using automatic utterance analyses with self-organizing neural networks. The goal of this research is to determine which parameters correlate across the results generated from diagnostics (from a number of different specialists, e.g., speech therapists and specialists, psychologists, neurologists, and from EEG and MRI tractography) and tests. Our laboratory uses methods based on computer speech analysis to diagnose children with SLI.

This thesis addresses the issue of identifying specific language impairments in children on the basis of their speech. The aim is to develop methods and procedures that could be utilized to unambiguously classify these children. The procedures described in this thesis are intended to be used as part of a software tool for evaluating treatment progress and assisting physicians in clinical praxis and to create a simple tool for identifying this disorder that can be used anywhere by anyone.

Existing methods and their value for the diagnosis of specific language disorders were examined. The current practice uses various tests to assess children's language skills. This approach is comprehensive but certainly not simple and quick, and it cannot be used to quantitatively test children with SLI. These limitations were a significant factor in the decision to develop a simple test for detection SLI in children. Our method is based on the number of pronunciation errors in the child's utterances. The main idea to create a diagnostic tool that is accessible and easy to use. The use of a mobile device allows extreme flexibility in where the tests are conducted; the test can be performed anywhere (e.g., at home, in kindergarten, at a park, …) and not just in a speech therapist's clinic. The

users of this application will be primarily the parents of the children with possible language impairments.

An alternate approach uses methods that diagnose children with SLI using acoustic analysis of the speech signal. The methods for detecting SLI are based on auditory signals that are specific to the acoustic features of speech. These features can be easily obtained, and calculations can be performed without complicated modifications of the speech signal. This method seeks to apply modern techniques to the diagnostic approach. Modern techniques allow the calculation of as many different acoustic features of an audio recording as possible. This approach reduces difficult decisions about which features and methods are relevant to the task but adds requirements for optimization methods and classification methods. The main benefit of this method resides in the possibility of developing an automatic detection system as its foundation, without the need for lengthy preparations of input data or labeling the children's utterances. Labeling utterances is difficult because of mispronunciation and various artefacts caused by fidgeting and because not all errors in the utterances clearly indicate specific language impairments.

The methods and procedures introduced in this thesis are based on different algorithms. Both approaches have one thing in common: the use of artificial neural networks, specifically Kohonen Self-Organizing Maps (KSOMs) [2], for the final classification. KSOMs determine the decision-making criteria used to classify children. Both methods, error analysis and feature analysis, were developed with intent of creating simple and reliable methods for classifying children with SLI and to help determine treatments.

Descriptions of classification experiments are an integral part of the thesis. The experiments involved the construction of evaluation methods that take into account the specific assignment of grant NT11443-5/2010, to which this dissertation also applies. It was necessary to develop such methods to classify children with SLI into three categories (mild, moderate and severe) depending on the degree of SLI without prior knowledge of this factor. The classification into three categories was carried out de facto blindly. To verify the classification results, a speech therapist's assessment was used. The algorithms are based only on the internal dependencies of data obtained from analysis without external knowledge about the children's severity.

To investigate the speech problems of children with SLI, it was necessary to create a speech database. The creation of a speech database designed for research on children with specific language impairments (especially developmental dysphasia) was one of the goals of this thesis. It was a very lengthy process. In my case, it represents almost ten years of constant work. In contrast to other speech databases, data acquisition was much more complicated because I work with children, very often preschoolers; this can create specific problems related to maintaining the children's attention and to parents' reluctance to agree to the research. Currently, the database contains 289 speech recordings from 188 children. The completion of the speech database required manually processing a total of

19,115 words. A precise description of the speech signal was created using the LABELING program [3] to segment the speech signal. Correctly processed data are critical for analyses.

State of Art

The ability to communicate via spoken language is one of the most important human attributes. Although there are several other means of communication, speech is difficult to replace in everyday life. Inabilities to communicate using speech can isolate individuals from society. Isolation resulting from speech impairments is significant for children with specific language impairments.

Specific Language Impairment (SLI) [4], [5], [6], [7], [8] is given as a diagnosis when a child has delayed or disordered language development for no apparent reason. It is described as a language disorder that delays the mastery of language skills in children who have no hearing loss or other developmental delays. Other names for this disorder can be developmental language disorder, language delay or Developmental Dysphasia (DD). The problem of developmental language disorder is one of the most common childhood learning disabilities. Almost 7 percent of all children aged 4 to 12 years have this disorder. The impact of the disorder in real life is that a child does not have the same speech skills as other children at the same age because the speech skills are delayed. Children with SLI thus experience the creation of a type of social barrier that separates these children from their contemporaries and disrupts their social lives. The most important factor for the development of a child's language skills and subsequent treatment is the timely diagnosis of SLI. In the following text, we describe a simple rule for classifying SLI based on speech analysis. The sooner treatment begins, the sooner child will reach a so-called "normal" level of language skills for children of their age, even if their language is delayed. According to recent studies, approximately 10 percent of all children experience some type of speech problems, with approximately 7 percent suffering from SLI and between 1 and 2 percent having serious problems with speech that require special attention [9].

SLI is characterized by a specific development of speech with signs of delay and aberration. Children with SLI have problems with the syntactic, semantic and grammatical aspects of their native language. They do not speak with a rich vocabulary, and they have problems making sentences with correct word order. SLI is a primary disease affecting other activities of a given individual (e.g., brain damage, hearing impairment), and secondary expression is reflected in speech and speech execution. Other manifestations of this defect can be a worse short-time memory, problems with attention and concentration, or problems with painting and perception of music. Additionally, these children can show abnormal EEG activity and defects in fine motor skills. The child usually knows that he has problems with speech, despite also experiencing any of these symptoms. The situation can escalate, such that the child stops using speech.

One of the more citate publication is [9], which reflecting findings and interpretations based on the hundreds of studies that have appeared since the publication of the first edition in 1997. Topics include linguistic details (descriptive and theoretical), word and sentence processing findings,

genetics, neurobiology, treatment, and comparisons to such conditions as autism spectrum disorders, ADHD, and dyslexia. The epidemiologic study estimated the prevalence of specific language impairment (SLI) in monolingual English-speaking kindergarten children is describes in [10]. [11] is the chapter from comprehension Child Psychology and Psychiatry: Frameworks for Practice (Second Edition), which introduces the research literature on specific language impairment (SLI). Authors characterize the key features, causes and the major types of SLI, and how children with SLI may be identified and differentiated from children with other developmental disorders. Fewer publications can be found under the name of Developmental Dysphasia. It is for example: A conceptual approach of the Theory Diagnosis and Treatment of Developmental Dysphasia from Speech Language Therapist A. G. Beesems is the transcript of a lecture given in Turkey at the: "disabled 07 Congress" [11]. In the Czech sources are found both names. Speech therapist and phoniatrist more use SLI (speech language impairments), but clinical neurologists rather DD (developmental dysphasia). In this thesis I will (with some exceptions) to use the name of Developmental Dysphasia. In most cases, Czech user can get an information which are available on the websites of organizations associating parents of children with SLI [12], [13].

Research groups have developed various tests for SLI, which are based on grammatical and reading skills. One of them is the Grammar and Phonology Screening (GAPS) test [14]. This test is a quick screening tool that can be administered by a professional or amateur. The grammatical abilities and key pre-reading skills of children at the age of 3.5 to 6.5 years are tested. The GAPS test was standardized on 668 children from across the UK. The goal of this test is to determine whether the children have sufficient knowledge concerning the use of grammatical rules to create sentences and whether they know the rules underlying how to add sounds together to correctly make words.

Speech analyses have already been used in patients with a number of speech difficulties. A large number of works discuss the acoustic characteristics of these speech issues. In particular, paralinguistic analysis, such as the recognition of speech emotion [15], [16], [17], [18], [19], [20], [21], and pathology[22], [23], [24], [25], [26], [27], [28] is increasingly becoming a mainstream topic in speech and language processing. The Interspeech 2013 Autism SubChallenge [29] addresses two developmental disorders with speech manifestations: autism spectrum disorders and specific language impairment. The article [30] aims to provide a broad overview of the paralinguistic analysis as a complex whole. This article describes a broad overview of the constantly growing field by defining the field, introducing typical applications, presenting exemplary resources, and sharing a unified view of the chain of processing.

1. AIMS OF THE DOCTORAL THESIS

The main motivation of this thesis is to find approaches and methods for classifying children with specific language impairments that provide relevant information to guide speech therapists'

decisions. These methods should also provide information about the level of specific language impairments.

This doctoral thesis has the following interrelated goals:

- **The creation of the Children speech database:** The task was to create a specific speech database of children with SLI, with complete data processing and the processed recordings of utterances from the control group (healthy children). For further research, only part of the database, namely, the data related to children aged 6 to 11 years, was used.

- **The identification of children with SLI:** The goals were to propose a method for identifying and classifying children with specific language impairments and to develop a novel classification approach that could help speech therapists to assess children. The method should work using utterances recorded in a speech therapist's office; thus, it must be robust enough to disregard artefacts and noise present in the signal.

- **Proper formant detection:** The motivation for the proper detection of formants from the continuous speech signal is to overcome problems that might hinder the wider application of these methods. The complexity of formant calculation, along with their possible misclassification (i.e., with the Burg algorithm in PRAAT), created a tool that was unsuitable for research purposes. Thus, a neural network was implemented for the proper classification of formants.

- **Discussion of results:** We compare the results obtained from the methods with the findings of a clinical speech pathologist and discuss eventual discrepancies. The data obtained might provide a way to collect a new data and participants in research. The methods should be able to distinguish between healthy children and children with SLI. Additionally, the method is assumed to have potential for further extension to perform more precise classification.

- **Application for classification:** We propose a simple tool for assessing a person's speech skills. The results obtained from the above-described experiments and tests should lead to the development of a practical application.

2. WORKING METHODS AND RESULTS

a) Speech Databases

To investigate the effect of speech problems of children with SLI, it was necessary to create a speech database [104]. The main of the criteria in creating the corpus were the selection of a suitable text and participants for speech recording. The words and phrases are directly selected for research of children with this disorder and they take into account the physiological and mental development of the child. It is necessary to select suitable words, phrases and sentences for recordings speech of children. It is essential to select utterances that children of a given age will be able to articulate. I worked with children aged from 39 to 132 months. Younger children cannot yet read, and they repeat spoken

utterances as a result, so it is necessary to maintain the same conditions for all enrolled children. A suitable text for recording contains different types of words and phrases, which were obtained from speech therapists and clinical psychologists who were using common, professionally recognized testing procedures, while also making use of their own knowledge and expertise in the subject area.

The recordings were created in a doctor´s office in the faculty hospital and private speech therapist's office. Only the speech therapist and the participants were present at the recording because it is very difficult to keep a child's attention until the end and the presence of another person would be disturbing. The session took place in the following manner: the speech therapist read a chosen text and the child repeated it. The sessions were structured in this way to standardize the conditions for both groups of children (healthy or SLI children). Both of the speakers are recorded by lapel microphone into the computer. The sound recordings are saved in the standardized vaw format. The sampling frequency is set to 44.1 kHz with 16-bit resolution in mono mode.

One software packs was created to simplify the creation and management of this speech database, and it is designed for research; Create Database Structure (CDS)

The entire database contains three subgroups of recordings of children's speech from different types of speakers. The first subgroup (healthy children, or H-CH) consists of recordings of children without speech disorders; the second subgroup (SLI-CH I) consists of recordings of children with SLI and the third subgroup (SLI-CH II) consists of children who have SLI of different degrees of severity (1 – mild, 2 – moderate, and 3 – severe). The description of all databases is provided in Table 1.

	H-CH		**SLI-CH I**	**SLI-CH II**
	Healthy	**With Defect**		
Girls	45	16	13	26
Number of recordings	45	16	22	45
Boys	25	17	33	46
Number of recordings	25	17	64	88
All children	70	33	46	72
All recordings	70	33	86	133
All utterances	4620	2178	5676	8819

Table 1 Description of All Databases.

b) **Classification Based on The Error Analysis (transcriptional analysis)**

Many research groups have developed various tests for SLI based on grammatical and reading skills. One is the Grammar and Phonology Screening (GAPS) test [14]. The goal of this test is to determine whether a child has sufficient knowledge concerning the use of grammatical rules to create sentences and whether they know the rules about how to add sounds together to correctly make words.

In my case, I attempted to analyze the words pronounced by children with SLI and compare them with those pronounced by healthy children (control group). I focused only on the description of errors

in individual words. During my studies of children with SLI, I found that their utterances include many more errors than those of healthy children. These errors occur across all age categories (our research includes children aged 39 to 131 months).

I developed an error analysis based on the number of pronunciation errors in the utterances. A requirement of pronunciation is the ability to distinguish the sounds of the spoken language by hearing. The basic criterion for developing this error analysis includes a process of creating three matrices (namely the reference (RM), testing (TM) and confusion (CM) matrices) and two parameters for speech (text reading by speech therapist (ut1) and repeating text by children (ut2)).

- **Classification of the Children with SLI**

The results of the analyses of utterance errors are displayed in Figure 1, which presents all of the participants. The healthy children are displayed in blue, and the children with SLI are displayed in red. Pronunciation errors are displayed in the upper graph. A higher value indicates a higher number of errors. The utterances of the children with SLI have a higher total number of errors compared with the utterances of the healthy children. The distributions of errors for the healthy children and the children with SLI are displayed in the middle graph. The distribution of the healthy children's errors was clustered around lower values compared with the error distribution of the children with SLI. Box plots representing the distributions of errors for the healthy children and the children with SLI show clear differences between these groups. The children with SLI had a higher number of errors in their utterances than did healthy children of the same age. Table 2 shows the difference in the average number of errors between these two groups; p_sli and p_h.

Healthy Children vs. Children with SLI - Participants					
Age Category	**Average Error**		**Difference p_sli vs. p_h**	**Comparison**	**Difference [%]**
	p_sli (2)	**p_h (1)**			
All	38,89	4,93	33,96	2 vs. 1	688,84

Table 2 Error Analysis - Comparison both groups - p_sli and p_h.

Table 3 presents the methods used to distinguish these two groups and their percent success rate. I achieved a 93.81% success rate using the best method. Because it was necessary to determine the limits value for every category, I used two different approaches to classify the participants into these groups. In the first method, HM ("hand-made"), I determined the threshold values for each group using the minimum and maximum values. The values located outside these limits were identified as misclassifications (misclass in Figure 2). These values were then evaluated according to the penalty score (PS) criterion. The PS is the sum of all errors for each word. The criterion for classification into each group was as follows: for the SLI group, (seven or more words contained an error), and for the healthy group, The other three methods are based on Artificial Neural Networks (ANN) [71], specifically Self-Organizing Maps (SOMs) [2], which are a standard part of the Neural Network Toolbox in MATLAB.

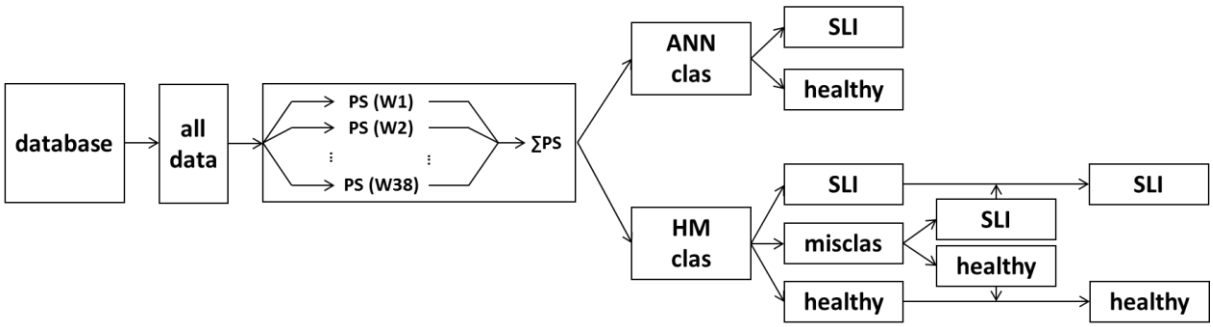


Figure 2 Error Analysis - Overview of the classification through ANN and HM classifier.

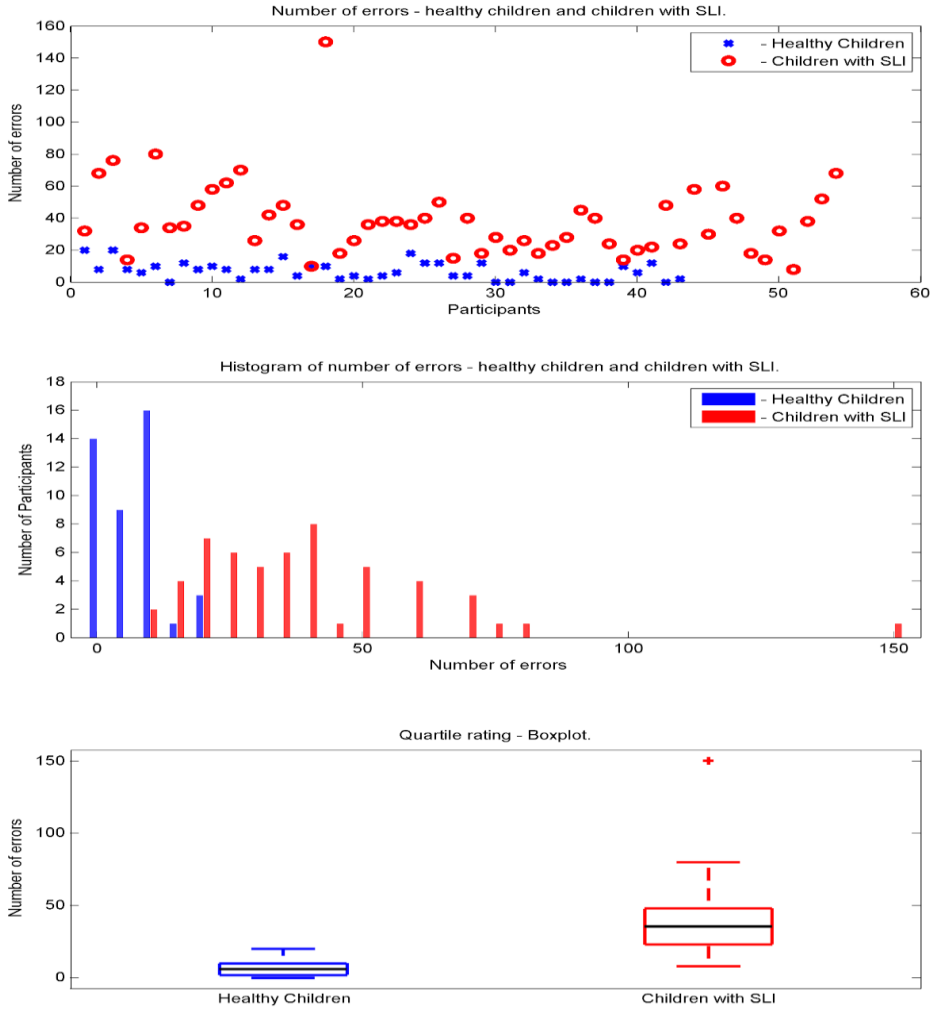


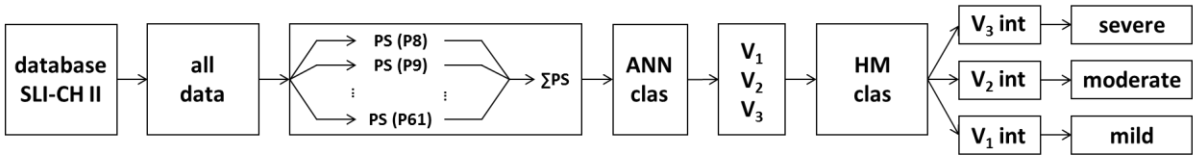
Figure 1 Evaluation of the error analysis.

Both classifiers, ANN and HM, offer approximately the same success rate for classification. For further development and utilization of this method (especially for parents), the HM classifier is easier to use. This result proved the premise that children with SLI have a higher number of speech errors compared with healthy children of the same age.

- **Differentiate children with SLI under three categories**

This experiment focused on identifying and classifying three degrees of specific language impairment: specifically mild, moderate and severe. To determine severity, we used only the error rate (i.e., a greater number of errors indicated a higher degree of SLI severity). The results were compared with a speech therapist's assessment. The aim of the experiment was to determine whether there is a relationship between the error rate and the degree of SLI.

The procedure was as follows: Only the data from the group of children with SLI (SLI-CH II) were classified. An error rate was assigned to each word. Then, the resulting penalty score (the sum of the error rates for all words) was calculated for each participant. To determine the particular degree of SLI, we used ANN and HM classifiers. An overview of the process of classifying the children according to the three degrees of SLI, including intervals, is provided in Figure 3.



$$V_3 \text{ int} : V_3 - \frac{V_3 - V_2}{2} < PS$$

$$V_2 \text{ int} : V_2 - \frac{V_2 - V_1}{2} \le PS \le V_3 - \frac{V_3 - V_2}{2}$$

$$V_1 \text{ int} : PS < V_2 - \frac{V_2 - V_1}{2}$$

Figure 3 Overview of the classification into three types of degree of SLI.

The resulting classifications into three categories depending on the degree of SLI are displayed in Figure 4. There are 2 box plot graphs. The left graph is for classification based on error analysis, and the right graph is for classification based on the assessment of a speech therapist. The mild category is displayed in blue, the moderate category is displayed in red, and the severe category is displayed in black. Both graphs have approximately the same layout. Box plots representing the distributions of utterance errors for all three classes show clear differences among these groups.

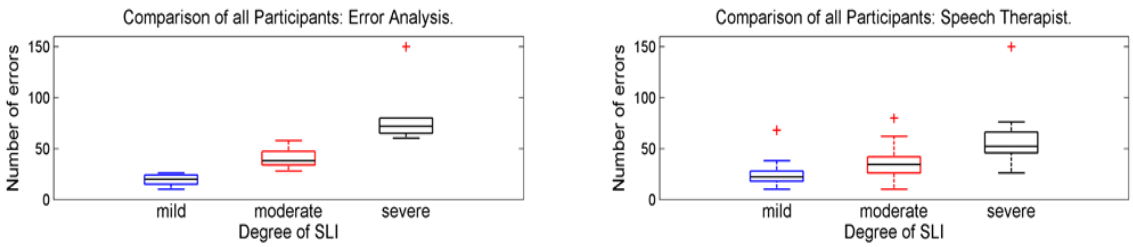


Figure 4 Comparison classifications - error analysis and assessment of speech therapist - depending on the error rate.

c) Classification Based on The Feature Analysis

This thesis focuses on children with SLI. One inspiration was achievements in acoustic analyses related to the recognition of emotions. An analogy between the analysis of emotions and the analysis of SLI can be made: Children without pathological changes in speech and no diagnosis of any disease can be compared with a neutral emotion, while children with SLI can be compared with some unspecified emotion, e.g., fear or anger.

In my case, I attempted to compare the words spoken by children with SLI with those spoken by healthy children (control group). I focused on acoustic speech parameters, which were extracted from individual words without using the Labeling program [3]. This method of classification could provide a basis for the future development of a completely automatic detection system. The aim of this chapter and the use of this method is to identify acoustic features that can be used to uniquely identify children with SLI.

Both of the examined issues (the classification of children with SLI and the ability to distinguish between different degrees of SLI) followed the same implementation structure. The implementation work can be divided into four parts whose respective key components can be described as follows:

1. Input data: The data used to identify children with SLI were selected from our speech database, particularly from the H-CH and SLI-CH II subgroups.

2. Feature extraction: The feature extraction methods are among the main components of the detection system used to identify children with SLI based on their speech. To extract the features, we used the OpenSMILE toolkit [32] (freely available under the terms of the GNU General Public License).

3. Feature selection: Presents methods and procedures used to reduce of the number of features.

4. Classification: Presents the decision-making procedures used for the final classification. The assignment of the children to a specific class, i.e., healthy vs. SLI, was relatively simple. Each participant was evaluated on the basis of the selected features of several words (usually 38). Thus, each participant was evaluated by of all the selected features from all the words. The resulting classification was based on the number of classifications, i.e., 1 for a correct classification and 2 for a misclassification, and a class is assigned based on the number of classifications. The winning class was based on a larger number of classifications.

The entire process used to identify the children with SLI is shown in Figure 5.

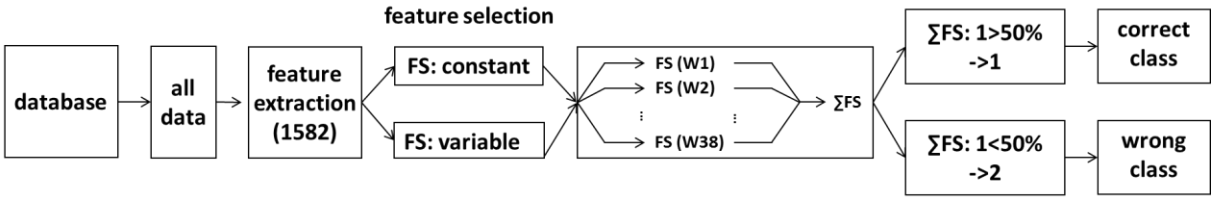
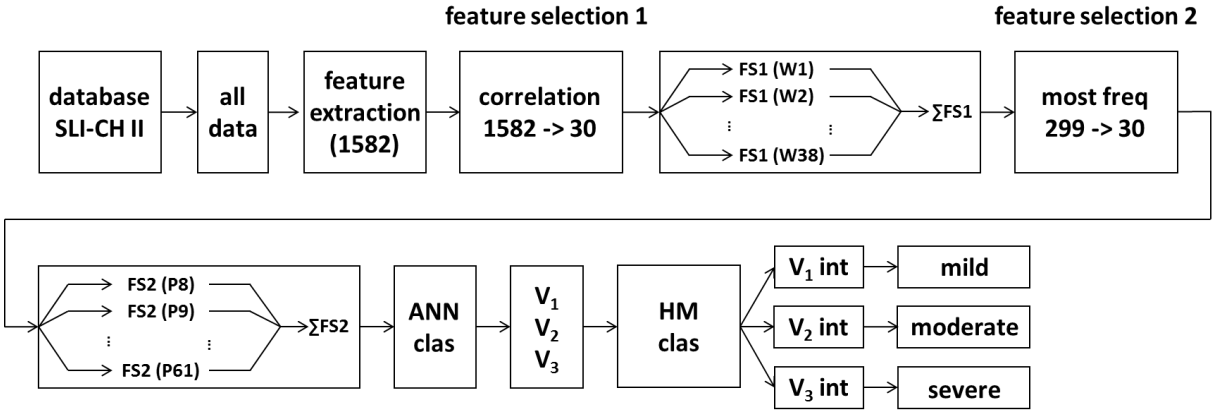


Figure 5 Overview of the classification children with SLI through acoustic features.

For the final classification of the degree of SLI, the same evaluation method used for error analysis, the ANN and HM classifiers, was used. SOM was used to determine three values (V1, V2 and V3) around which most of the other values occurred. The entire process used to classify the three degrees of SLI is shown in in Figure 6.



$$V_3 \text{ int} : V_3 - \frac{V_3 - V_2}{2} < PS$$

$$V_2 \text{ int} : V_2 - \frac{V_2 - V_1}{2} \le PS \le V_3 - \frac{V_3 - V_2}{2}$$

$$V_1 \text{ int} : PS < V_2 - \frac{V_2 - V_1}{2}$$

Figure 6 Overview of the classification into three types of degree of SLI through acoustic features.

The method described in this chapter is being developed to determine whether a speaker has a specific language impairment. It is based on the acoustic speech parameters extracted from individual words without using the Labeling program. This allows the creation of an automatic detection system based on this method.

- **Classification of the Children with SLI**

The results of the feature analyses for all of the participants are displayed in Figure 7. The "healthy" classifications are displayed on the left, and the "SLI" classifications displayed on the right.

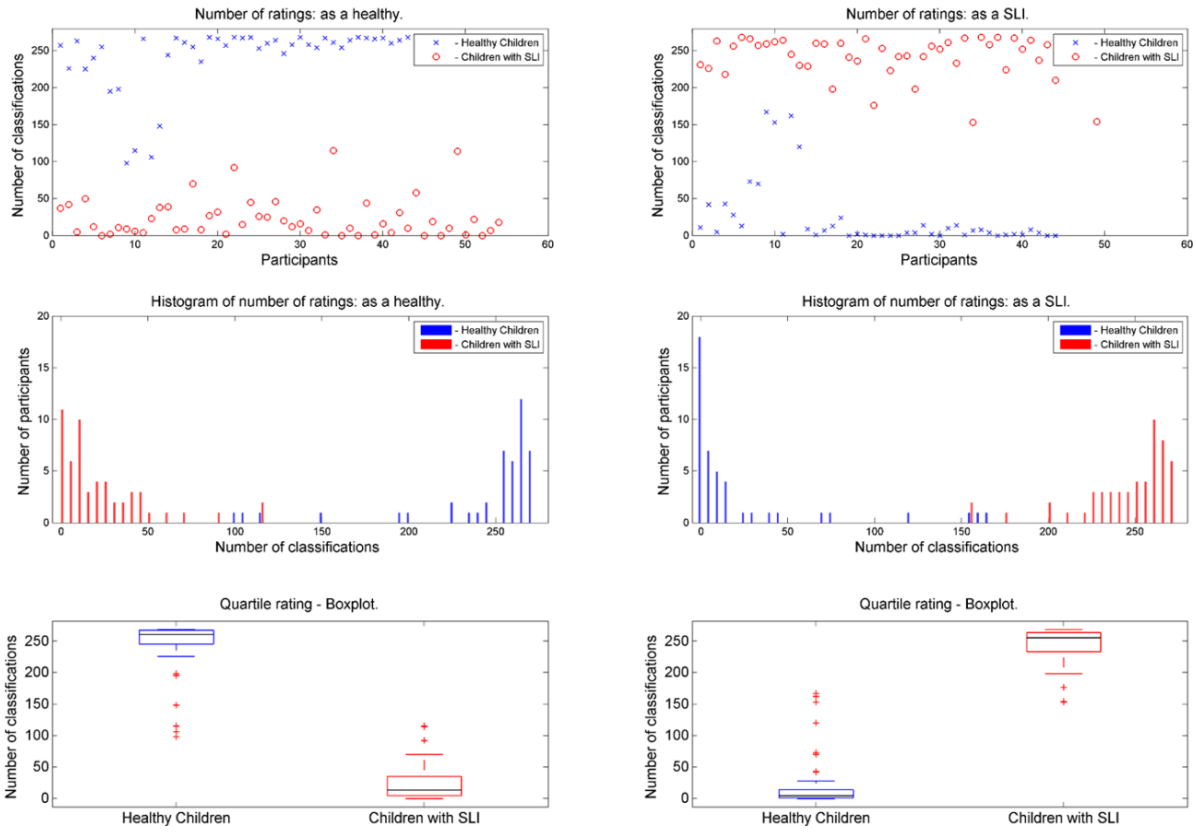


Figure 7 Evaluation of the feature analysis. Data from healthy children are shown in blue colour, and data from children with SLI are shown in red colour. Samples with a higher number of errors are in a higher position in the first graph. A histogram of the errors of each of the participants is shown in the second graph. Box plots represent the distributions of the number of utterance errors of healthy children and children with SLI.

Table 2 presents the final results used to distinguish the two groups and their percent success rate. I achieved a 96.94% success rate for both groups. Only 3 participants out of 98 were misclassified. These participants were from the p_h group, i.e., the healthy children. This result proved that it is possible to find acoustic features that can distinguish healthy children from children with SLI with high accuracy.

Healthy Children vs. Children with SLI – Final Percent Success Rate.				
Age Category	**Classification of Participants**			Success Rate [%]
	Group	Correct	Wrong	
All	P_SLI (2)	54	0	100
	H-CH (1)	44	41	93,18
	∑(1 + 2)	**98**	**95**	**96,94**

Table 2 The success of classification. The percentage rates of correct classification of the method used to distinguish the two groups.

- **Differentiate children with SLI under three categories**

The resulting classifications into three classes depending on the degree of SLI are displayed in Figure 8. There are 2 box plot graphs. Left graph is for classification through the feature analysis and

right graph is for classification based on the assessment of speech therapist of values obtained from feature analysis.

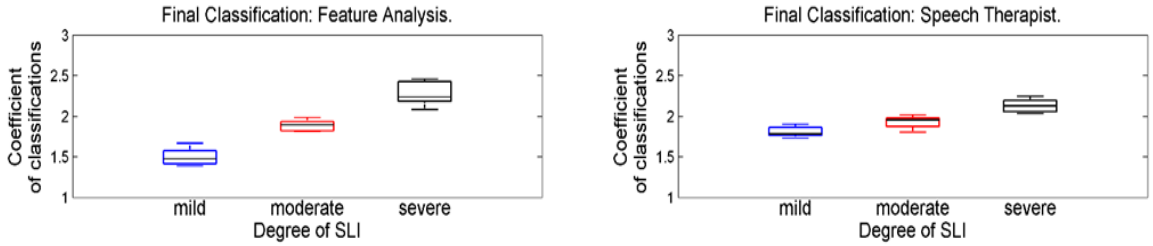


Figure 8 Comparison classifications - feature analysis and assessment of speech therapist - depending on the coefficients of classification.

d) Detection of Formants

The ability to produce and perceive speech originates in certain parts of the human brain. SLI is described as a neurological disorder of the brain [33], [34], [47]. The formant is a parameter that possesses a physical dimension (i.e., the presence of acoustic energy across the spectrum of speech sounds). The existence of formants can be related to human brain activity and to the movement of the articulatory system. This fact satisfies a condition for the use of formants in the classification of children with SLI. One of the prerequisites for the use of formants as descriptive parameters is the ability to obtain formant values with a minimum number of errors. To acquire suitable formants, we used the software tool FORANA [45]. The development of this software mainly driven by the need to correctly complete a formant analysis. Originally, formant frequencies were extracted from speech signals using PRAAT [30] acoustic analysis software. However, because the use of the PRAAT, specifically the use of Burg's algorithm to compute formants (method "To Formants (burg) ..."), the software produced formant classification errors in the course of the analysis, and results obtained using this approach could not considered relevant. More about this issue can be found in [40], [49].

A new formant analysis method using a continuous signal solves problems with the frequency of analysis. The method can by very simply described in the following four steps:

1. Dividing the signal into n equally long consegments. Each consegments contains more segments.
2. Multiple formant analysis. Each consegments is subjected to more formant analysis. ¨
3. Selection of the appropriate frequency.
4. Compositions of the formants.

A graphic illustration of the FF algorithm is shown in Figure 9.

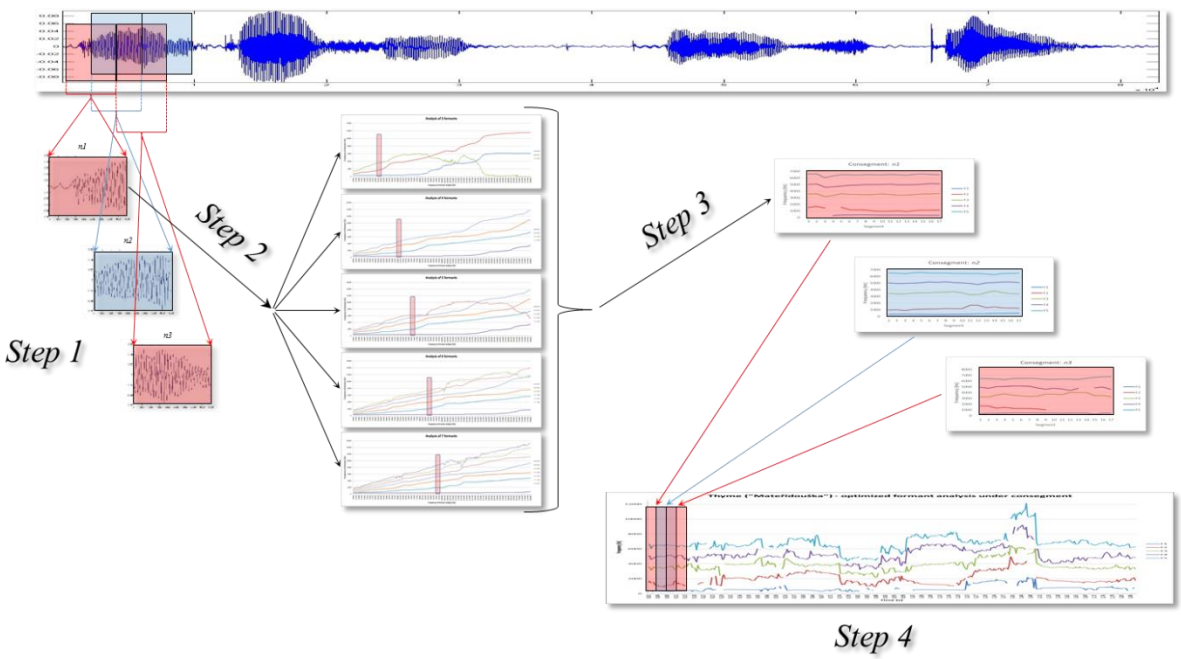


Figure 9 Graphic illustration of Folding algorithm of formants

e) SLIt Tool

The Test of Specific Language Impairments (SLIt Tool) is a tablet application that calculates very simple test for detection of children with SLI on iOS platform (Apple, Inc.), specifically for using on iPad (iPad 3rd generation or newest). The application is based on the Error Analysis, which was presented in previous chapters and is the original method of testing children. The main idea was the creation of the application that is very friendly and easy to use. A mobile device enables extreme flexibility in where the tests are carried out, making it possible to perform the test anywhere (e.g., at home, in kindergartner, in the park, …) and not just in speech therapist clinic. The users of this application will be primarily parents of children. The uniqueness of this application is the possibility to perform timely diagnostic test for determining SLI; in the result this means faster and more successful speech therapy. Functional design of SLIt Tool is in Figure 10.

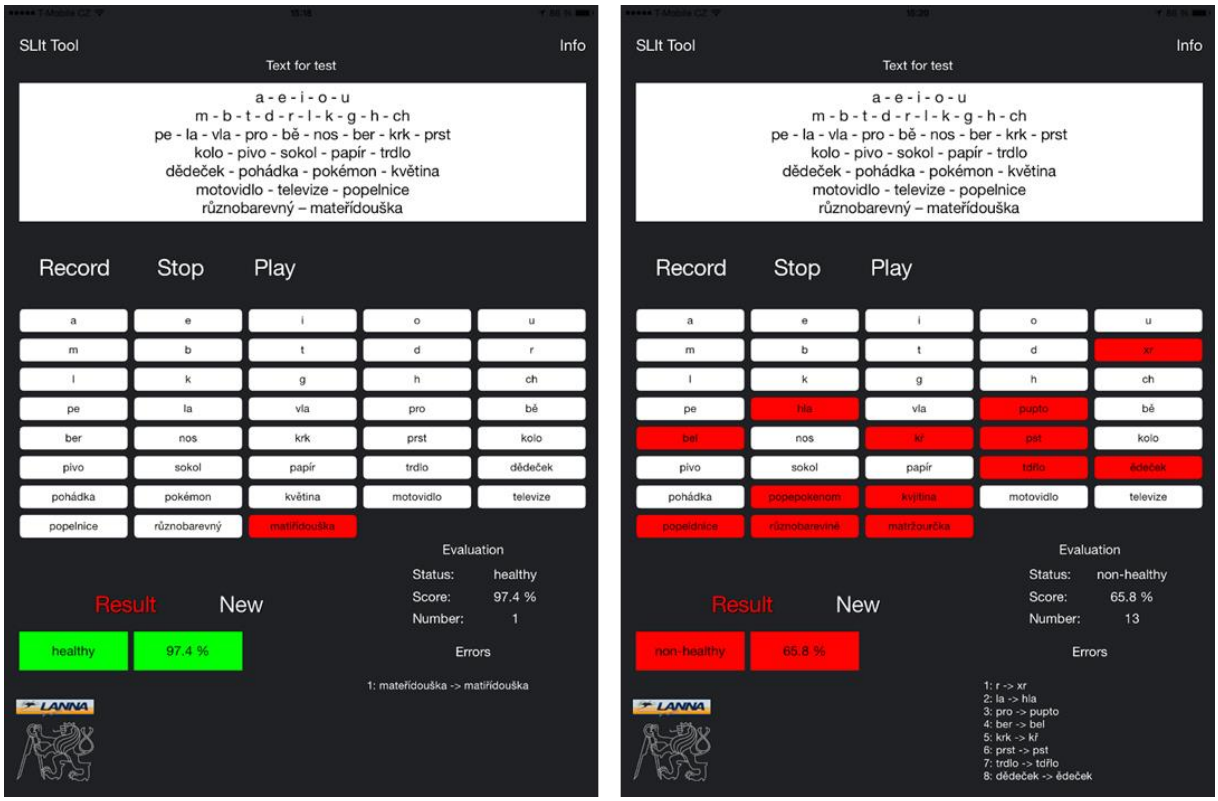


Figure **Chyba! V dokumentu není žádný text v zadaném stylu.**.1 Screen shot of SLIt Tool application on iOS platform (a) A part in the left: In this case we tested a healthy children with this results; error analysis is equal to 97.4 % success rate and the final state is healthy. (b) A part in the right: This case shows a real test of child with diagnosis of SLI with this result; error analysis is equal to 65.8 % success rate and the final state is non-healthy.

3. CONCLUSION

A method described in this thesis was developed to analyse disordered children speech. I focused on children with specific language impairments and believe that the utilization of this method in clinical practice will bring more insight into progression and treatment of the disease and help to think out efficiently treatment of the disease.

This work is a part of the research project that is focused on treatment of SLI. The project integrates results from the fields of neurology, psychology, logopedics, MRI tractography and speech processing. In cooperation with the department of Paediatric Neurology in 2nd Faculty of Medicine of Charles University in Prague we develop methods for utterance analysis that further advance diagnosis of the children with SLI and help to find the most efficient therapy.

Speech database

The creation of a speech database designed for research focusing on children SLI (especially developmental dysphasia) is a very lengthy process. In contrast to the other speech databases, data acquisition is much more complicated, because I work with children, very often of preschool age, and this can create specific problems. The most complex problem is how to maintain the children's attention throughout the recording. Children are very often distracted and may not focus on the recording: they may speak during the speech therapist’s speech, asking him about various things or

they may play with the microphone. Therefore, the recordings contain many unwanted sound remnants. Another problem is the reluctance of parents, of both healthy and sick children, to agree to the research. The research process involves many examinations of their children and parents have to fill out a medical questionnaire, which is one reason why the dataset is not larger. It is important to note that these are biological data which have their own particular specificity and the quantity of the data is not the most significant of their properties. The last problem is related to the technique and location of the recordings. The rooms were not soundproofed and normal urban noises can be overheard in recordings. The same applies to the quality of the recordings. At the beginning of our research study we had a different equipment for recordings than at the end of the research. The recordings are with a variable quality and with a different level of background noise. It is very complicated to compare recordings of low quality and those of high quality. Despite these problems and shortcomings, I managed to find several methods for identifying children with SLI.

Error analysis

Error analysis is based on the number of errors in the utterances. The advantage of this access is that its function does not require complex computational methods and can be performed by anyone. This approach provides complete information about the most common errors and substitutions between speech sounds in the utterances of the children with SLI. The results of the error analysis show that the number of errors in utterances is an important cue that can be used as a border line to determine healthy children and children with SLI.

This approach enabled the creation of a simple test to detect children with SLI, specifically mobile application SLIt Tool. The initial goal of creating a test originated when we solved the problems with the creation of a children's speech database. The main problem is in the quality of the speech recordings, which is variable because the recording was carried out in real environment of speech therapy clinics. All of these problems have been described in [35]. The criteria used to create the simple detection test included noise immunity in the recordings, and the newly developed SLIt Tool for testing that we presented in this article fulfills this condition.

The result of the output from application is the recommendation for children which obtain a status as a "non-healthy" to visit a speech therapist for deeper classification of their speech problems. The indisputable advantage is the simplicity of this method, which can be used anywhere by anyone. To use this test, no special technical knowledge that would prevent its wider use by speech therapists or other language therapists is required.

Feature analysis

This method, called feature analysis, for identifying children with SLI based on the auditory signal features specific to the acoustic features of speech. These features can be easily obtained and can be calculated without complicated modifications of the speech signal. This method seeks to apply

a modern techniques to this issue. Modern techniques permit the calculation of as many different acoustic features of audio recording as possible. This approach reduces the need for difficult decisions about which features and methods are relevant to the task but adds the need for optimization and classification methods. The main benefit of this method is its foundation for the possible development of an automatic detection system. Similar to the results for error analysis, these experiments show that feature analysis can distinguish healthy children and children with SLI and also can distinguish different degrees of SLI (mild, moderate and severe). These results were in agreement with the assessments of a speech therapist.

Formants

Formant analysis provides information about the individual vowels in the frequency spectrum. Each vowel has a clearly defined location in the vocalic triangle when two conditions are fulfilled. First, the formants must be correctly classified. Second, the utterance must be properly spoken. The whole point of using formants in our research is to verify the correctness of the spoken utterances. Children with SLI have problems with correctly speaking difficult (multisyllabic) words. It is supposed that developmental dysphasia can influence a shift of formant frequencies in spectral characteristics compared with the formant frequencies of healthy children. Formant analysis clearly verifies whether the vowels contained in utterances are pronounced correctly. Otherwise, if there are any errors in the analyzed vowel in the utterance, there is a shift in the frequency spectrum. This fact means that the speakers have the articulatory organs in a bad position and the distribution of articulatory cavities has the wrong shape for a vowel. In the final result, it leads to the malfunction of speech control in the brain. The disadvantages of using formant analysis include a notable susceptibility of the quality of the recordings and the demand for correct classification of the calculated formants.

Conclusion

These results prove that it is possible for children with specific language impairments to be clearly identified and distinguished from healthy children, based on their speech and speech skills. I combined traditional and alternative approaches to this issue and obtained a resistance tool that is not dependent on the quality of the captured recordings. I found several different classification methods for children with SLI. All of these methods can be used separately for the classification of these children. Each method yields a high level of success in classifying children with SLI, but each has its particular limiting factors and shortcomings. Using these methods together, we are able to eliminate these shortcomings and obtain a powerful tool for diagnosing children with SLI.

The possibility of the classification of children with SLI using their speech analysis is very significant. The first factor is the opportunity to detect more children with this disorder. For these children, therapeutic treatment can be performed earlier. Another important factor is the financial

aspect. Diagnostics that use speech are less demanding on hardware, and are more financially feasible compared with other specific medical examinations.

List of literature used in the thesis statement

[1] LANNA - Laboratory of Artificial Neural Network Applications, Department of Circuit Theory at the FEE-CTU in Prague. http://ajatubar.feld.cvut.cz/lanna/

[2] Kohonen, T. (2001). Self-Organizing Maps. Third, extended edition. Springer.

[3] Tučková J, Bártů M, Zetocha P (2009) Aplikace umělých neuronových sítí při zpracování signálů. (lecture notes in the Czech) Prague: Czech Technical University in Prague. ISBN 978-80-01-04400-1.

[4] Leonard L B (1998). Children with specific language impairment. Cambridge, MA: MIT Press. ISBN 9780262621366

[5] Bishop, D V M (2014). Uncommon Understanding (Classic Edition): Development and Disorders of Language Comprehension in Children. England Psychological Press. ISBN 9780203381472.

[6] Klein, L.T., Amato, V. (2013). Languages and Linguistics. Nova Science Publishers, Incorporated. ISBN 9781629483337.

[7] Zwitserlood, R. L. M. (2014). Language Growth in Dutch School-Age Children with Specific Language Impairment (Proefschrift). LOT. Netherlands. ISBN 9789460931383.

[8] Specific Language Impairment. Available:
http://www.nidcd.nih.gov/health/voice/pages/specific-language-impairment.aspx
Accessed 03 November 2014

[9] Tomblin, J.B., Records,N.L., Buckwalter,P., Xuyang Zhang, Smith,E. and O'Brien, M.: Prevalence of Specific Language Impairment in Kindergarten Children. Journal of Speech, Language, and Hearing Research, December 1997, Vol. 40, 1245-1260. doi:10.1044/jslhr.4006.1245

[10] Conti-Ramsden, G. and Durkin, K. (2011) Specific Language Impairment, in Child Psychology and Psychiatry: Frameworks for Practice, Second Edition (eds D. Skuse, H. Bruce, L. Dowdney and D. Mrazek), John Wiley & Sons, Ltd, Chichester, UK. doi: 10.1002/9781119993971.ch29

[11] Beesems, A.G.: Foundation Developmental Dysphasia. Theory Diagnosis and Treatment. Speech Language Therapist, Developmental Dysphasia Foundation Amsterdam.

[12] Bočková, B.: The Support of Pupils with Specific Language Impairment (in Czech: Podpora žáků se specificky narušeným vývojem řeči). 1ed, Brno: Masarykova univerzita, 2011, ISBN 978-80-210-5609-1.

[13] http://www.logopedonline.cz/vady-reci/vyvojova-dysfazie.html

[14] van der Lely HKJ, Payne E, McClelland A (2011) An Investigation to Validate the Grammar and Phonology Screening (GAPS) Test to Identify Children with Specific Language Impairment. PLoS ONE 6(7): e22432. doi:10.1371/journal.pone.0022432

[15] C.-N. Anagnostopoulos, T. Iliou, and I. Giannoukos, "Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011," Artificial Intelligence Review, pp. 1–23, 2012.

[16] A. Stuhlsatz, C. Meyer, F. Eyben, T. ZieIke, G. Meier, and B. Schuller, "Deep neural networks for acoustic emotion recognition: Raising the benchmarks," in ICASSP, 2011.

[17] T. Vogt and E. Andre, "Comparing feature sets for acted and spontaneous speech in view of automatic emotion recognition," in ICME, 2005.

[18] A. Batliner, K. Fischer, R. Huber, J. Spilker, and E. Noth, "Desperately seeking emotions: Actors, wizards, and human beings," in Proc. ISCA Workshop on Speech and Emotion, 2000.

[19] A. Kazemzadeh, J. Gibson, J. Li, S. Lee, P. Georgiou, and S. Narayanan, "A sequential bayesian dialog agent for computational ethnography," in Interspeech, 2012.

[20] D. Bone, C.-C. Lee, and S. S. Narayanan, "A robust unsupervised arousal rating framework using prosody with cross-corpora evaluation," in Interspeech, 2012.

[21] E. Mower, M. Mataricc, and S. Narayanan, “A framework for automatic human emotion classification using emotion profiles,” Audio, Speech, and Language Processing, IEEE Transactions on, vol. 19, pp. 1057–1070, 2011.

[22] J. McCann and S. Peppe, “Prosody in autism spectrum disorders: A critical review,” International Journal of Language & Communication Disorders, vol. 38(4), pp. 325–350, 2003.

[23] J. van Santen, E. Prudhommeaux, L. Black, and M. Mitchell, “Computational prosodic markers for autism,” Autism, vol. 14, pp. 215–236, 2010.

[24] D. Bone, M. P. Black, C.-C. Lee, M. E. Williams, P. Levitt, S. Lee, and S. S. Narayanan, “Spontaneous-speech acoustic-prosodic features of children with autism and the interacting psychologist,” in Interspeech, 2012.

[25] C. Kaland, E. Krahmer, and M. Swerts, “Contrastive intonation in autism: The effect of speaker- and listener-perspective,” in Interspeech, 2012.

[26] R. Lunsford, P. A. Heeman, and J. P. H. van Santen, “Interactions between turn-taking gaps, disfluencies and social obligation,” in Interspeech, 2012.

[27] M. Swerts and C. de Bie, “On the assessment of audiovisual cues to speaker confidence by preteens with typical development (TD) and atypical development (AD),” in Interspeech, 2012.

[28] G. Kiss, J. P. van Santen, E. Prudhommeaux, and L. M. Black, “Quantitative analysis of pitch in speech of children with neurodevelopmental disorders,” in Interspeech, 2012.

[29] B. Schuller, S. Steidl, A. Batliner, A. Vinciarelli, K. Scherer, F. Ringeval, M. Chetouani, F. Weninger, F. Eyben, E. Marchi, M. Mortillaro, H. Salamin, A. Polychroniou, F. Valente, and S. Kim, “The interspeech 2013 computational paralinguistics challenge: Social signals, conflict, emotion, autism,” in Interspeech, 2013.

[30] Schuller B., Steidl S., Batliner A., Burkhardt F., Devillers L. , Müller Ch., Narayanan S., Paralinguistics in speech and language-State-of-the-art and the challenge, Computer Speech and Language, v.27 n.1, p.4-39, January, 2013, ISSN 0885-2308.

[31] PRAAT: doing phonetics by computer. Available: http://www.fon.hum.uva.nl/praat/. Accessed 11 August 2015

[32] Eyben F, Wöllmer M, Schuller B (2010)"openSMILE - The Munich Versatile and Fast Open-Source Audio Feature Extractor", In Proc. ACM Multimedia (MM), ACM, Florence, Italy, ACM, ISBN 978-1-60558-933-6, pp. 1459-1462, October 2010. doi:10.1145/1873951.1874246

[33] Swierkoszová J (1990) Špecifické poruchy učenia. In Lechta,V. a kolektív: Logopedické repetitorium. Bratislava: SPN. ISBN 80-08-00447-9.

[34] Tučková J, Komárek V (2008) Effectiveness of Speech Analysis by Self-Organizing Maps in Children with Developmental Language Disorders. In: Neuroendocrinology Letters. Ed.: Peter G. Fedor-Freybergh. Society of Integrated Sciences, vol. 29, No.6, Nov/Dec 2008, ISSN 0172-780X.

List of candidate’s works relating to the doctoral thesis

Publications in impact journals

[35] Grill P, Tuckova J (2014) A database of normal and dysphatic children's speech. In: PLoS One Manuscript draft with manuscript number: PONE-D-14-47426 “Send to press”

Publications excerpted in SCOPUS

[36] Tučková J, Bártů M, Zetocha P, Grill P (2011) Self-Organizing Maps as Data Classifier in Medical Applications. Proceedings of the 3rd International Joint Conference on Computational Intelligence - International Conference on Neural Computation Theory and Applications, Paris, France, 2011. SciTePress , Madeira.

Others publications

[37] Vavřina J, Grill P, Olšanský V, Tučková J (2012) A phoneme classification using PCA and SSOM methods for a children disorder speech analysis. Lékař a technika, 2012, vol.42, no.2, pp.85-88.

[38] Grill P, Vavřina J, Tučková J (2013) Databases and their applications for diagnosis of developmental dysphasia. Proceedings of the 11th international workshop IEEE, ECMSM 2013, Toulouse, France, 2013

[39] Grill P, Tučková J (2011) Identification the speaker from continual signal by formants analyses. In Proceedings of the 8th International Conference on Digital Technologies 2011 [CD-ROM]. Žilina: Slovenská elektrotechnická společnost, 2011, p. 127-131. ISBN 978-80-554-0437-0.

[40] Grill P, Tučková J (2010) Formant analysisis - vowel detection of children with developmental dysphasia. In Digital Technologies 2010. Žilina: TU v Žilině, 2010, ISBN 978-80-554-0304-5.

[41] Grill P, Tučková J (2008) Formant Analysis - The Appropriate Frequency of Formant Analysis. Proceedings of the 5th International Conference on Digital Technologies 2008, Žilina, Slovakia

[42] Tučková J, Grill P, Zavadil O, Wasserbauer V (2007) Formant analysis – FORANA. In Proceedings of the 4th International Conference on Digital Technologies 2007, Žilina, Slovakia, 2007.

[43] Tučková J, Grill P (2009) Určování formantů pro analýzu emotivně zabarvené řeči . Proceedings of the 8th Czech-Slovak Conference Trends in Biomedical Engineering, Bratislava, Slovensko.

[44] Grill P (2012) Odhad formantů v patologické řeči. In LETNÍ DOKTORANDSKÉ DNY 2012. Praha: ČVUT, 2012, s. 82-85. ISBN 978-80-01-05050-7.

[45] Grill P, Tučková J (2009) FORANA. In Technical Computing Prague 2009. Praha: Humusoft, 2009, s. 32-39. ISBN 978-80-7080-733-0.

[46] Grill P (2012) The database of the children with developmental dysphasia. In BioDat 2012 - Conference on Advanced Methods of Biological Data and Signal Processing. Prague: Czech Technical University in Prague, 2012, vol. 1, p. 32-34. ISBN 978-80-01-05153-5.

[47] Vránová M, Tučková J, Kynčl M, Grill P, Komárek V, et al. (2011) MRI abnormality řečových drah. In Celostátní mezinárodní logopedická konference Tábor, 4. 10. 2011

[48] Vránová M, Grill P (2011) MRI abnormality řečových drah a počítačová analýza promluv dětí s vývojovou dysfázií. Celostátní konference LSMS , Praha, 20. 5. 2011

[49] Grill, P. - Tučková, J.: Formants application to diagnose of children developmental dysphasia. 9th Czech-Slovak Conference Trends in Biomedical Engineering. Ostrava: VŠB, 2011, s. 98-101. ISBN 978-80-248-2475-8.

SUMMARY

Předložená dizertační práce se zabývá tématem detekce neuro-vývojové poruchy u dětí. Vývojovou dysfázií (VD), jenž se jinak označuje specificky narušený vývoj řeči, trpí přibližně 5 - 7 % dětské populace. V současné době se jedná o velmi často diagnostikovanou poruchu v rámci SLI, která koreluje i s jinými poruchami např. se specifickou poruchou v učení. Jedná se o vrozenou komplexní poruchu osvojení řečových dovedností a zasahuje celou osobnost člověka. To znamená zrakové a sluchové vnímání, paměť; prostorovou orientaci; motoriku - hrubou, jemnou, grafomotoriku; řeč. Není ale ustáleným pravidlem, že musí být při VD zasaženy všechny tyto oblasti. Velmi důležitým faktorem pro úspěšnou terapii této poruchy je její včasná diagnostika.

Tato dizertační práce byla součástí Grantu NT11443-5/2010, který zahrnuje a porovnává výstupy dosažené z oblastí neurologie, psychologie, logopedie, EEG, MRI traktografie a počítačového zpracování dětské řeči.

V rámci řešení dizertační práce bylo stanoveno několik konkrétních cílů a to:

1) Vytvoření specifické řečové databáze obsahující řečové nahrávky od dětí s vývojovou dysfázií.
2) Identifikace dětí s VD. pomocí několika různých metod.
3) Vyřešení problému chybné klasifikace formantů s možností jejich automatické detekce.
4) Vytvořit jednoduchou aplikaci pro klasifikaci dětí s VD.

V této dizertační práci jsou popsány dva různé přístupy zaměřené na klasifikaci dětí s VD. První klasifikuje tyto děti pomocí chybovosti promluv (error analysis). Bylo statisticky ověřeno, že existuje závislost dětí s VD na chybovost v promluvách a byl stanoven postup, který byl konzultován s klinickým logopedem, na detekování dětí s VD. Tento postup byl implementován do podoby mobilní aplikace (SLIt Tool), jenž je primárně určena pro rodiče dětí a umožňuje jim provést test, zda jejich děti netrpí touto neuro-vývojovou poruchou. Výstupem z této aplikace je doporučení, v závislosti na dosaženém výsledku, zda se obrátit (test vyjde negativně) nebo neobrátit (test vyjde pozitivně) na řečového specialistu, aby provedl odborná vyšetření a stanovil přesnou diagnózu. Tento postup umožní včasnější diagnostiku této poruchy a spolu s ní i rychleji a úspěšněji provedenou terapii. Druhý přístup byl založen na akustických parametrech řečového signálu (feature analysis) s důrazem na možnost využití této metody pro automatickou klasifikaci těchto dětí.

Výsledky dosažené v této dizertační práci potvrdily hypotézu, že lze klasifikovat děti s VD pomocí analýzy jejich řeči.