Effectiveness of speech analysis by self-organizing maps in children with developmental language disorders

Jana Tuckova[1], Vladimir Komarek[2]

1. Faculty of Electrical Engineering, Czech Technical University in Prague;
2. Department of Child Neurology, 2nd Faculty of Medicine, Charles University and Medical Hospital Motol; Czech Republic.

Correspondence to: Doc. Ing. Jana Tuckova, CSc.
Faculty of Electrical Engineering, Czech Technical University in Prague,
Technicka 2, 166 27 Prague 6, Czech Republic.
tel.: +420721502556, fax: +42023339805, e-mail: tuckova@fel.cvut.cz

Submitted: 2008-08-16 *Accepted: 2008-10-06* *Published online: 2008-12-29*

Key words: **developmental language disorder; supervised self-organizing maps; children speech analysis; artificial neural networks; EEG monitoring.**

Neuroendocrinol Lett 2008; **29**(6):939–948 **PMID:** 19112409 NEL290608A07

Abstract

At present, more and more researchers are applying tested mathematical-engineering methods into different domains of life. One of these areas is helping people with different forms of disabilities. Research in this area is focused on searching for the relation between clinical and electrophysiological symptoms of children with developmental dysphasia. Sleep EEG and speech analyses are the primary areas under discussion, while the finding of methods acceptable for improvement of the diagnosis and determination of therapeutic procedures is the research topic. It is possible to reduce fundamentally, or to cure optimally these disorders in advanced diagnosis. Therefore it is important to search for new methods and to combine what has been used separately till now.

INTRODUCTION

Developmental language disorder (DLD), alternatively named specific language impairment (SLI) or developmental dysphasia (DD) is defined as a disruption in the ability to acquire normal language skills adequately to age, while at the same time displaying intact peripheral hearing, normal intelligence and absence of behavioral disorder or negative social factors (Menkes & Moser, 2006). Sufficient evidence exists to confirm that the incidence of interictal epileptiform activity in children with DLD is higher than that of normally developing children. We assume that the only way to determine the possible negative effect of interictal discharges is the direct detection of cognitive deficits during interictal discharges by means of computerized neuropsychological testing and synchronous video-EEG monitoring. Evidence that suppression of discharges can improve neuropsychological functioning is essential to justify the use of antiepileptic treatment in „nonepileptic" neurodevelopmental disorders. The epileptiform discharges in N-REM sleep are similar to those found in children with Landau Kleffner syndrome (LKS), alternatively named acquired epileptic aphasia. In contrast to dysphasia, aphasia is considered to be "acquired" when it results from a condition that occurs after language development has begun, generally after the age of 2 years. Children with LKS typically develop normally until they are about 4 years old, when they rapidly and quite unexpectedly cease to understand what is said to them and soon also lose the ability to express themselves in words. At its onset, the language regression is associated with continuous spikes and waves in sleep

Abbreviation and units:

DLD	– developmental language disorders
SLI	– specific language disorder
DD	– developmental dysphasia
LKS	– Landau-Kleffner syndrome
SOM	– self-organizing maps
KSOM	– Kohonen's self-organizing maps
SSOM	– supervised self-organizing maps
ANN	– artificial neural networks
SEGANN	– segmentation by artificial neural networks
MFCC	– mel-frequency cepstral coefficients

EEG. As in many DLD children with receptive prominent disorder, epileptiform activity simultaneously exists in wakefulness and non-REM sleep EEG; a possible hypothesis can be proposed for the amelioration of the language processing via antiepileptic drugs, specially benzodiazepines and steroids. Linkage between the intellectual or behavioral deficit has not yet been clearly established. Duvelleroy-Hommet et al. (1995) found epileptiform sleep EEG in 9 children from 24 patients with developmental expressive dysphasia.

Tuchmann et al. (1991a, b) analyzed EEG in 85 children with developmental language disorders. He found epileptiform activity in 58% of children in combination with seizures and in 9% of children without epilepsy. We found epileptiform sleep EEG in about 40% of our series of patients with developmental dysphasia, most of whom have never experienced clinical seizures. The results of our previous research provided evidence to support the theory that not only epileptic seizures but also subclinical epileptiform EEG discharges have a negative effect on brain development and on long-term prognosis. We believe that the quality of life of the child and its family could be improved if the diagnosis of epileptiform cognitive/behavioral deficit was correctly made, and the appropriate therapeutic intervention started as soon as possible.

The relationship between this disorder, audio signals, phonetic and phonologic characteristics is not well understood. Partial problems are being addressed in many research establishments, among others the ASHA (American Speech-Language-Hearing-Association), ATA (Alliance for Technology Access), LDA (Learning Disabilities Association of America), NIDCD (National Institute on Deafness and Other Communication Disorders) in USA, at Charles Stuart University in Australia, at the University of Manchester in GB, at the Brain Research Institute of the University of Vienna in Austria, RCSLT in London (Royal College of Speech and Language Therapists) for example. Georgopoulos at al. (2003) described the methodology tool based on fuzzy cognitive maps.

We can establish a relation between developmental dysphasia (Hrncir & Komarek, 2004; Pospisilova, 2005) and the assessment of the degree of perception and impairment of the speech. From the point of view of the characterization of language, determination of relevant and irrelevant speech information, and its connection with a searched target, is very complicated. For this reason, a part of project is addressed by artificial neural networks (ANN), which are trained in the knowledge of phonetics. Another argument for using the ANN is its robustness.

Many research teams around the world are engaged in the medical applications of ANN. They are, for example the Artificial Neural Networks in Medicine and Biology Society (ANNIMAB-S).

Our method is clustering of the pattern characteristics visible by the allocation of the vowels respectively by changes in allocation of the vowels pronounced by the patients. In the first experiments, above all, the vowels were chosen for easy labeling (start/end determination). Another added reason has been that developmental dysphasia can influence a shift of formant frequencies in spectral characteristics compared with the formant frequencies of healthy children. Generally, a relatively small number of studies have involved children. The limited information availability concerning this area is a reason for the preparation of a comparable speech database of healthy children.

A preference for self-organizing maps (SOM) has been assumed from the nature of our problem. For many real problems, the target values for all the patterns of the database are not known. Nor do we know all the characteristics of the patterns. In such a case, a selection of one form of unsupervised learning – clustering is suitable. One of the symptoms of the children with a disorder indicating developmental dysphasia is a malfunction of perception and impairment of speech, yet we cannot say which one is affected, nor to what degree. We decided to use a supervised self-organizing maps described by Kohonen, 2001. He was inspired by the self-organizing procedure in the human brain, by its adaptations and learning ability. Ngan et al. (2002) applied the Kohonen SOM procedure in a medicine research also in another area. One of the goals of our research is to create a software pack with a user-friendly interface for medical doctors or other medical staff.

MATERIAL AND METHODS

Population

35 children were referred for suspected DLD by our outpatient child neurology centre, community neurologists and speech therapists. 28 of them (age: 3 years 5 months - 9 years 1 month, average 5 years 6 months) fulfilled the diagnostic criteria of DLD. All children underwent an overnight sleep video-EEG and logopedic/psychological evaluation. 72 healthy children (age 4-10 years) served as a control speech database.

EEG evaluation

We have performed standard EEG assessments with the use of 19 EEG channels in a standard montage system 10-20. For this reason, we had to utilize another method of simplified sleep structure evaluation: using a videorecording, we had marked the whole sleep period defined as a time from first falling asleep in the evening till final awakening in the morning. During this period, we marked all time intervals where the child was in deeper sleep stages, i.e. in NREM 2 up to NREM 4. These stages were detected by expertise visual inspection of EEG signals with standard EEG classification according Rechtschaffen-Kales scoring rules. Detailed marking for the whole night was performed only in two selected EEG recordings and partial marking with several first NREM stages from evening sleep was done in few other recordings. Seeing that this method in extraordinary arduous and time-consuming, we have used another method of marking deeper NREM stages in the rest of EEG recordings. From the previous investigations (see e.g. Duvelleroy-Hommet et al. (2000)) it is known that nonlinear EEG analysis called CER (Course-grained Entropy Rate = one of 31 computed analyses) shows (by level of its values) conditions where EEG signals present deeper NREM states (i.e. under NREM1 stage). With help of these CER analyses, we have marked all continuous time intervals (in lengths of at least 15 minutes) during the sleep period, which are supposed to correspond to these deeper NREM states.

Psychological evaluation

The children were assessed by a clinical psychologist, using the Gessel developmental scale, SB IV and an additional test standardized for the Czech language: sound differentiation test, world differentiation test, auditory analysis and synthesis test. Spontaneous talk was evaluated as well. Inclusion criteria: PIQ $\geq$ 70, disturbed phonemic discrimination, disturbed language on various levels – phonologic, syntactic, lexical, semantic and pragmatic.

Speech signal evaluation

The character of a speech signal is a very complex system based on technical, human, physiological, phonological and phonetic properties. Human physiological properties affect the ability of good pronunciation very strongly. Another factor is a case of longer segment of time (polysyllabic words). A specific language can exercise influence on this ability; e.g. one major is the clustering of consonants in the Czech language (strniště-stubble, srnka-roe deer).

The age of the child influences expression in the pronunciation ability for consonant clusters or polysyllabic words. If we analyze the children's speech, it is very difficult to determine whether the speech disability is due to cerebral dysfunction or infantility arising from a mental disease/disorder. In the case of small children, a likely cause can also be collection decay or mental intention. Account must be taken of these problems when we record the speech sentences of the children.

The speech corpus has to be composed from a children's speech recorded at kindergartens and on the first level of elementary school, and continuously complemented speech of children's patients recorded at hospital. Voice recording was done in real settings with high noise level. The noise reduction necessary for conventional methods could bring some irreversible information loss, though this problem is minor when artificial neural networks are applied.

The second problem deals with the speech evolution of children. The speech quality was strongly influenced by children's emotional tone: the children were afraid or they were shy.

All these characteristics and problems associated with children‘s speech are possible to solve by artificial neural networks. Methods based on the ANN are robust enough for the minimization of these effects. Concretely, the standard unsupervised Kohonen's self-organizing maps (KSOM) were using as starting experiments. They are based upon a mapping of input features onto different areas of the cerebral cortex in a topologically ordered manner. Later, the new variant – supervised self-organizing map (SSOM) (Kohonen, 2001) was used. The model combines aspects of vector quantization with a topology-preserving ordering of the quantization vectors.

Database Creation and Speech0 Pre-processing

In order to evaluate the degree of these modifications, it is necessary to have a comparative voice analysis of healthy children for the possibility of an assessment of the rate modification. Currently, a children‘s voice database is not available. Our team has created a speech database of children with developmental dysphasia and a comparative database of healthy children (44 female and 28 male). The utterance texts are compiled in a paediatric neurological clinic by neurological specialists as related to medical therapy. The same text is used for the healthy children for comparative analysis. The text (phonemes, syllables, words and sentence) is read aloud by an assistant (for healthy children) or by a psychologist (for patients) and the children repeat the text. Identical conditions for all age categories must be addressed. The speech of each child is recorded as a wav-file and is subsequently segmented. Utterances which create the speech database are divided into 11 parts (see Table 1).

Our choice of patterns takes special characteristics of the Czech language into account. Using the syllables configuration were: CV, CCV, CVC, CCC and CCCC in the part 3. Symbol "C" represents a consonant and symbol "V" represents a vowel. Consonant clusters CCC and CCCC are specific for Czech pronunciation. Only text parts from 1 to 7 were used in described experiments. No ill child repeated text in parts 8 to 10. From the point of view of listeners, parts from 1 to 4 did not induced problems for patients.

From the point of view of the impairment of speech, differences between healthy and ill children were registered by speech signal analysis only. The success of the process of analysis is definitely dependent on the preci-

Table 1. Speech database - structure

No.	Type of part	patterns
1	vowels	5
2	consonants	10
3	syllables	9
4	2-syllable words	5
5	3-syllable words	4
6	4-syllable words	3
7	5-syllable words	2
8	doubled words	3
9	augmentation of word order	4
10	compound sentence	1
11	acoustic differentiation	10

Table 2. Age group of child patients

Treatment by medication		Control group	
Patient	Age	Patient	Age
d	3.9	k	3.5
e	5.0	f	3.9
n	5.4	h	4.1
g	5.7	b	5.2
m	6.4	c	5.4
a	6.7	p	8.8

sion of the labeling of the natural speech signal in the database. By labeling, we understand the determination of all speech units, i.e. the decision where phonemes begin and end.

The speech signal was labeled by hand by means of the Cool Edit 2000 programmed in the first case, but now the automatic process by means of original software SEGANN (Segmentation by Artificial Neural Networks) is being explored (Tuckova & Zurek, 2007).

For all speech signals, the following pre-processing was used: sampling frequency 16 kHz, quantization 16 b/per sample, length of segments 20 ms (320 samples), overlap 50%, Hamming window, preemphase 1. In the first experiments, we used a signal parameterization for an acquirement of inputs data into SSOM, which are mel-frequency cepstral coefficients (MFCC). These coefficients approximate the human auditory system better then the other parameterizations (Psutka J. et al. (2006); Zwicker E.& Fastl H.(1990). Every segment was represented by 16 coefficients of MFCCs. Every vector of input data set was also completed with auxiliary information about the type of vowels.

Self-Organizing Map Applications

Multidimensional input data are transformed into a decreasing dimensional space during the iterative learning procedure, which is one of the important characteristics of the KSOM. This network contains one executive layer, which coordinates particular input vector elements (created by investigate properties or characterizations) with all executive neurons. The search of the minimal distance between the input vectors and coordinates of the neurons in the map is given by a basic mathematic formula, with the Euclidean distance or its square most frequently used. The neuron with minimal distance is named winner of the competitions (competitive learning); the winner is nearest from coded patterns.

A neighborhood of the winner is created using Kohonen's algorithm. All similar input vectors are updated in them. The cluster of all input vectors with common properties is created. They are allocated on the map and point to the number of dominant properties in one training epoch; it can to point a movement in the input data and „regrade" any characterization into different groups in the course repetition. Invariably, it is a classification.

The unified distance matrix or U-matrix is a representation of the KSOM that visualizes the distance between the neurons and its neighbors (Kohonen, 2001). The KSOM neurons are represented by hexagonal cells. The distance between the adjacent neurons is calculated and presented with different colors. Dark colors between neurons correspond to a large distance and thus represent a difference between the values in the input space. Light colors between the neurons means that the vectors are close to each other in the input space. Light areas represent clusters and dark areas represent cluster boundaries. This representation can make clusters better visible (this means has been using for our experiments – see paragraph IV).

The supervised self-organizing map (SSOM) (Kohonen, 2001) combines aspects of the vector quantization method with the topology-preserving ordering of the quantization vectors. The algorithm of the SSOM represents a very effective way of classification, but only for well-known input data or for well-known classes of input data (in our case we know text which is pronounced and thereby phonetic classes).

The SSOM consists of m units located on a regular, low-dimensional grid of map units. The map unit positions on the regular grid are fixed; each map unit is connected to a number of neighboring map units with a neighborhood relation. Supervised learning means that the input vector is formed of two parts, $\boldsymbol{x_0}$ and $\boldsymbol{x_c}$, where
$\boldsymbol{x_0}=[\boldsymbol{x_{01}}, x_{02},\ldots,x_{0n}]^T$,
$\boldsymbol{x_0}\in\Re^{\,n}$ is an original input vector of dimension n and
$\boldsymbol{x_c}=[\boldsymbol{x_{c1}}, x_{c2},\ldots,x_{ck}]^T$,
$\boldsymbol{x_0}\in\Re^{\,k}$ is assigned as known class of $\boldsymbol{x_0}$

(supervisor) in a training set (indication of vowels in our experiments). Each element of vector $\boldsymbol{x_c}$ represents one of k classes. A new vector $\boldsymbol{x}=[\,\boldsymbol{x_0}, \boldsymbol{x_c}\,]^T \in \Re^{\,n+k}$ will have a dimension $n+k$, which is valid for a prototype vector $\boldsymbol{m_i}=[m_{i1}, m_{i2}, \ldots, m_{i(n+k)}]^T$, $\boldsymbol{m_i} \in \Re^{\,n+k}$ as well. During the classification of an unknown input vector $\boldsymbol{x}$, only its $\boldsymbol{x_0}$ part was compared with the corresponding part of the prototype vectors. The class of each unit (neuron) is found by taking maximum over these added elements, and a label is give accordingly.

Spoken speech is a time-dependent sequence of phonemes. It is necessary to process the input data to ANN in a batch form. It is significantly faster and does not require any specification of a learning-rate factor (in comparison with the incremental learning algorithm, which is a commonly used algorithm in ANN training). New prototype vectors are calculated as a weighted average of the input vectors, where the weight of each input vectors is the neighborhood function value $h_{i,m^*(j)}$ at its winner $\boldsymbol{m}^*(j)$:

$$m_i(t+1) = \frac{\sum_{j=1}^{N} h_{im^*(j)}(t)\, x_j}{\sum_{j=1}^{N} h_{im^*(j)}(t)} \qquad (1)$$

where t is the number of iteration, $\boldsymbol{x_j}$ is the input vector, N is the number of input vectors. The most usual neighborhood function is the Gaussian one:

$$h_{ij}(t) = \exp\left(\frac{-\left\| r_j - r_i \right\|^2}{2\sigma^2(t)} \right) \quad \text{for } \left\| r_j - r_i \right\| \le \sigma(t) \qquad (2)$$

$$h_{ij}(t) = 0 \quad \text{for } \left\| r_j - r_i \right\| > \sigma(t) \qquad (3)$$

r_j, $r_i \in \Re^{\,2}$ are the location vectors of units j and i in map for 2-D. A parameter $\sigma(t)$, the neighborhood radius, defines the width of the kernel. It is some usually a smoothly decreasing function of time.

The Batch Map

The batch map (Kohonen, 2001) is an iterative process in which a number of input vectors $\boldsymbol{x}$ are classified into the respective V_i regions first. Secondly, new prototype vectors $\boldsymbol{m}$ are computed as weighted averages of all training samples:

$$\mathbf{m}_j = \frac{\sum_{i=1}^{n} h_{b_i j} \mathbf{x}_i}{\sum_{j=1}^{n} h_{b_i j}} = \frac{\sum_{i=1}^{m} h_{ij} N_i \bar{\mathbf{x}}_i}{\sum_{j=1}^{n} h_{ij} N_i} \qquad (4)$$

where $\mathbf{x}_i$ is an input vector, n is the number of input vectors, m is the number of units, N_i is the number of input vectors in Voronoi set V_i :

$$V_j = \left\{ \mathbf{x} \mid \left\| \mathbf{x} - \mathbf{m}_j \right\| < \left\| \mathbf{x} - \mathbf{m}_k \right\| \ \forall k \ne j \right\} \qquad (5)$$

$\bar{\mathbf{x}}_i = \sum_{\mathbf{x} \in V_i} \mathbf{x} / N_i$ is the mean of the vector $\boldsymbol{x}$ in Voronoi set V_i. The value of neighborhood function between map units $\boldsymbol{m_j}$ and $\boldsymbol{b_i}$ is sign as $\boldsymbol{h_{bij}}$, the winner (denoted also as the best-matching prototype - BMU) to the input vector $\boldsymbol{x_i}$ is computed by equation (6):

$$\mathbf{b}_i = \arg\min_j \left\{ \left\| \mathbf{x}_i - \mathbf{m}_j \right\| \right\} \qquad (6)$$

RESULTS

When we used the SSOM, we must divide data into specific classes. In our case, there are 3 or 5 classes, according to the phonetic category (Palkova, 1994). We are searching for similar acoustic properties which depend on the manner of their formation in vocal organs.

The ability of vowel classification and allocation in the map as a vocalic triangle is investigated in neural network applications. We have started from the hypothesis that it involves the disorder of movement of vocal organs in articulation in the case of developmental dysphasia, influencing the formant generation. The vowel mapping of patients is different in comparison with the vowels mapping of healthy children. The standard SOM and the supervised SOM for the training maps by healthy children were compared. Experiments with the speech differentiated by age or gender present dramatically better results for SSOM – see Figure 1.

Consequently, SSOM mapping has been used henceforth. Five clusters represent the Czech vowels *a*, *e*, *i*, *o*, *u*. The number of the unit in the SSOM of the healthy children is proportional to the presence of vowels in the examined text.

Parameters of Supervised SOM

Additionally, the following parameters were used for maps in this contribution:

- Map initialization – a 2-D map contains 24 x 24 units in a hexagonal grid, a random initialization of the prototype vectors.
- Map training – the batch map algorithm, the Gaussian neighborhood function decreased uniformly in training steps from 24 down to 1. The number of training steps was 5000.

Result Classification – example

In the experiments described here, we analyzed the vowel mapping. By this time 46 patients had been observed, of them 14 patients did not display other symptoms of developmental dysphasia.

Speech analysis has been performed for 12 children (3 girls and 9 boys) in the course of three- or four-month periods. After each period, the same utterances are recorded and analyzed. Six of them are on medica-

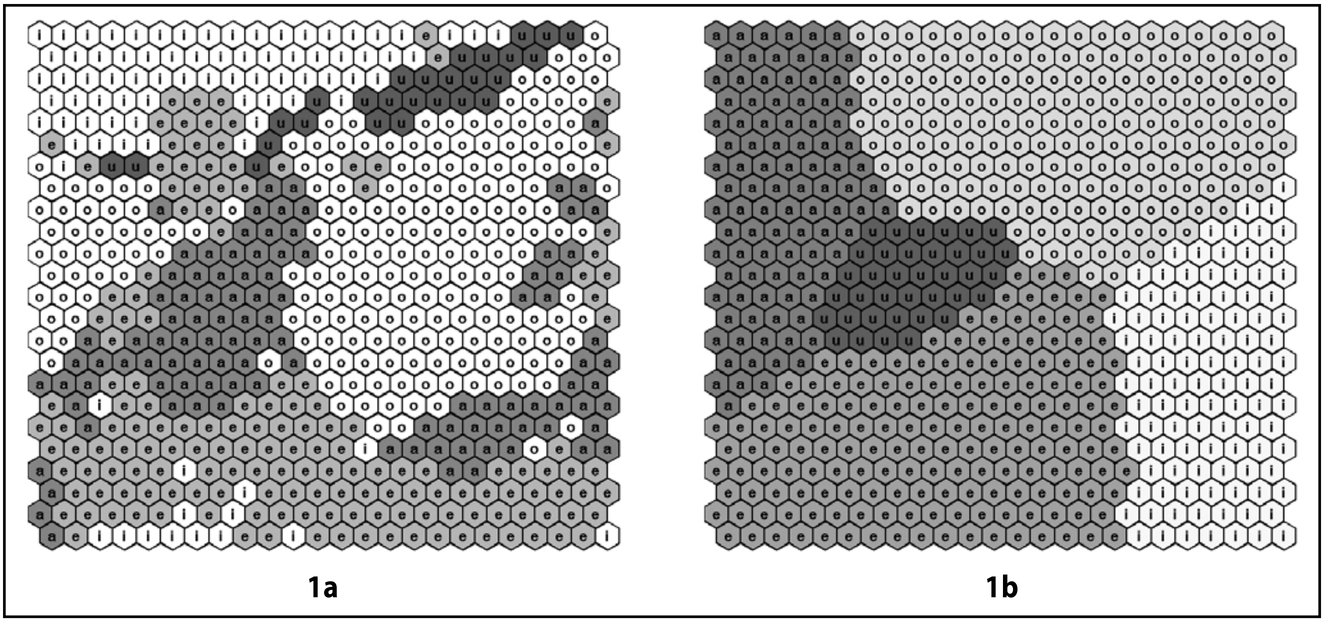


Figure 1. Comparison of vowels classification: a) by SOM, b) by SSOM.

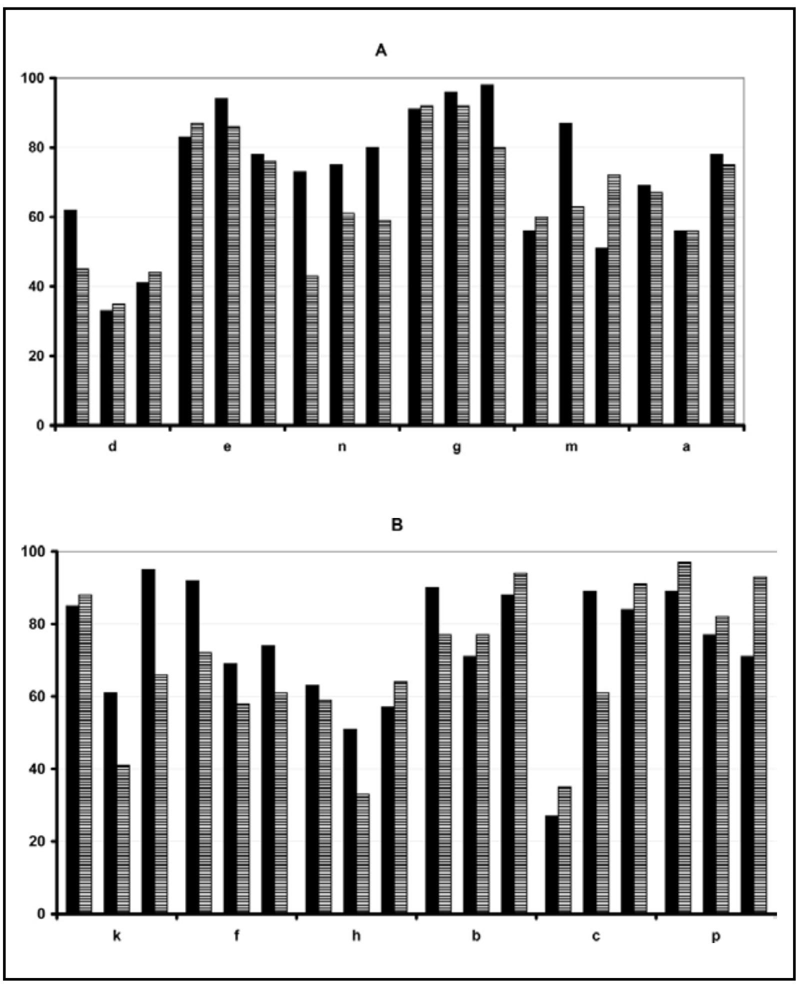


Figure 2. The behavior of the change of vowel classification of children with developmental dysphasia. a) patients with medication, b) patients - control group.

tion, while the others have only visited a speech therapist and a control clinical examination. Figure 2 shows results of vowels classification by SSOM. A solid column presents the success coefficient of classification to maps trained by speech of healthy 6 years old children, a line-hatched column presents the same for maps trained by speech of healthy children in age from 7 to 10 years. The patients indicated as 'a, d, e, g, m, n' have been treated with medication.

An example of the cluster visualization in the maps which represents the vowels distribution of an alternative child patient in comparison with the vowel distribution of 55 healthy children (35 girls and 20 boys) is showed in Figures 3 and 4. The children indicated for vowel classification in this contribution are between the ages of 7 and 10. The patient designated as "m" (Figure 3) was older than the patient designated as "n" (Figure 4). As is visible in figures 2, 3 and 4, a dependence on children's age is evident (see Table 2).

White units are the successful classifications from the map trained by speech data of healthy children, black units represent classification errors. Their number and location in the map is changed after each recording. An ability for good pronunciation depends on age too. The patient "m" was capable of pronouncing more speech patterns than the patient "n".

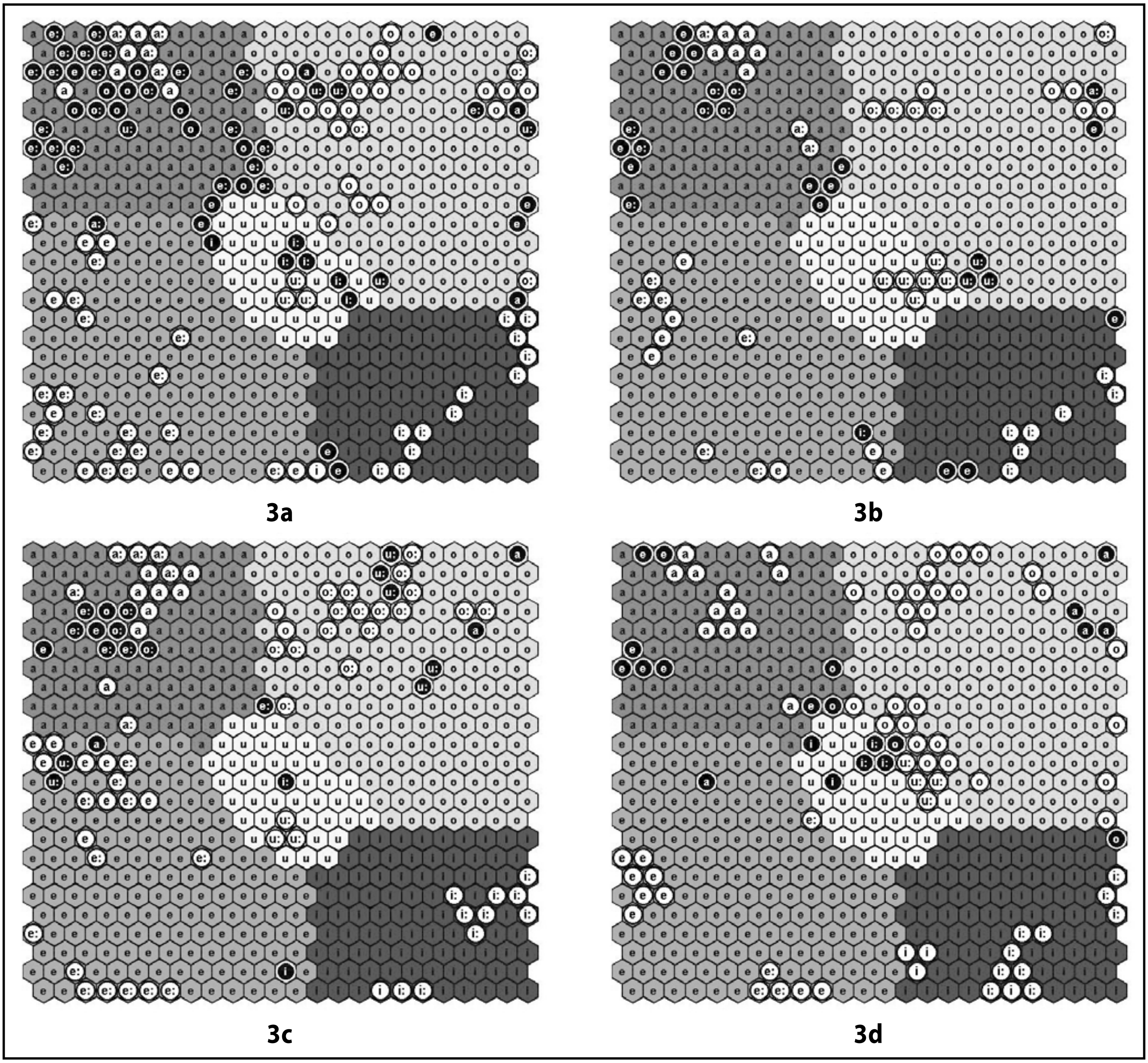


Figure 3. Classification of the vowels of the healthy and ill children indicated as "m". a) before therapy - 1st recording, b) after first part of therapy - 2nd recording, c) after second part of therapy - 3rd recording, d) after third part of therapy - 4th recording.

Histograms in Figure 5 show a relative rate

$$RR = 100\,\frac{Y}{Y+N} \qquad (7)$$

where Y is the number of successfully and N is the number of wrongly classified vowels in %. Our aim is to achieve a minimum of wrong classifications. Data analysis is aggravated by the following fact: ill children are not able to pronounce some vowels (monitored children have displayed problems with the pronunciation 'e', 'i', at some times with 'u').

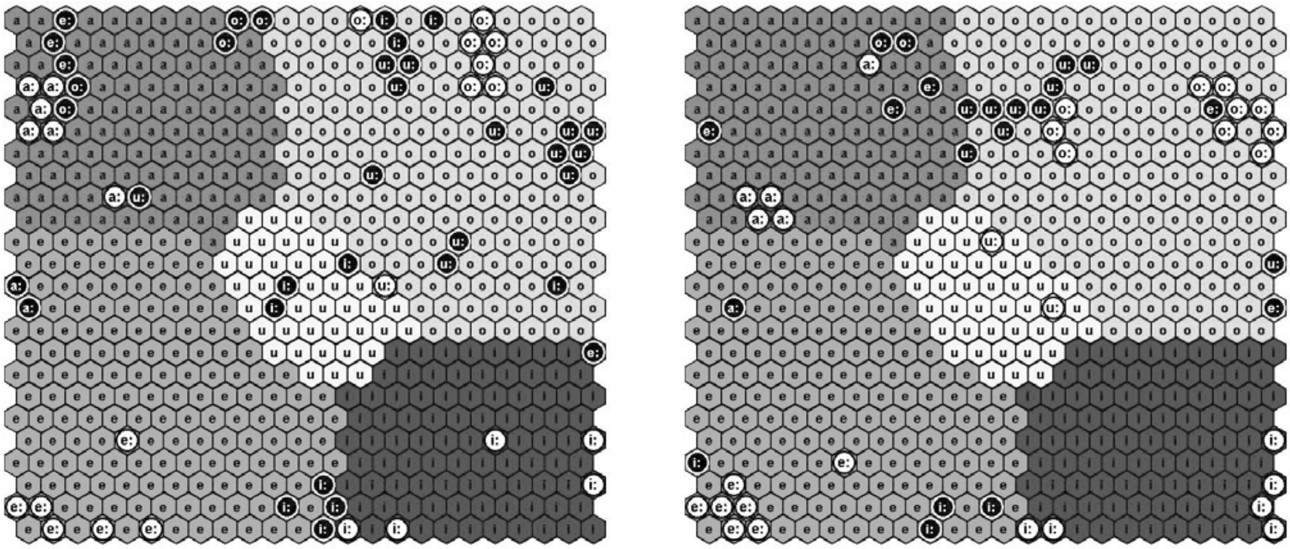
The success in vowel classification in percentage for patients who are on medication is shown in Figure 5 a), the same relationship for patients from the control group is showed in Figure 5 b). Each group is formed by six patients.

The values in particular bars are computed according to equation (7). The solid column presents 1st recording (before therapy), the line-hatched column presents 2nd recording (after first part of therapy), the cross-hatched column presents 3rd recording (after second part of therapy).

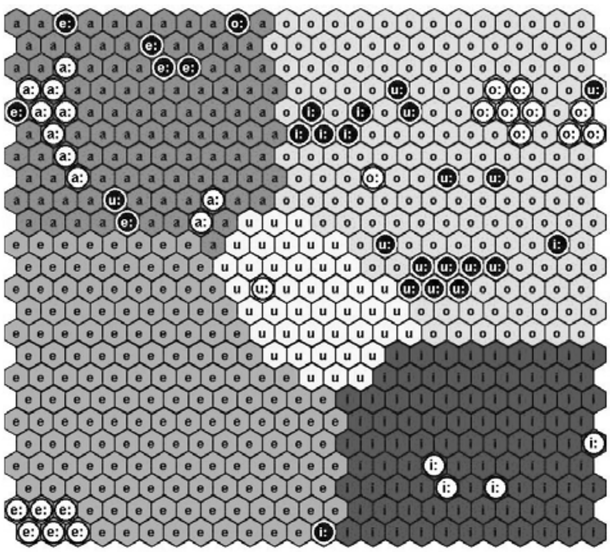
The obtained results are confirmed by psychological evaluation of patients 'n', 'g' for the present.

Software

All analyses and experiments were performed by the computational system MATLAB 7, Release 14. The software, called SOM Toolbox©, was applied in our experi-



4a 4b



4c

Figure 4. Classification of the vowels of the healthy and ill children indicated as "n". a) before therapy – 1st recording, b) after first part of therapy – 2nd recording , c) after second part of therapy – 3rd recording.

ments. SOM Toolbox was developed in the Laboratory of Information and Computer Science (CIS) in the Helsinki University of Technology and it is built using the MATLAB script language. The SOM Toolbox contains functions for creation, visualization and analysis of the Self-Organizing Maps. The Toolbox is available free of charge under the General Public License from (Vesanto et al. ; Kohonen et al. 1995). For the project, new special M-files, which should be a part of supporting program package, were created (Tuckova & Zetocha, 2006).

DISCUSION

This contribution describes an original method for the intensity of speech defect monitoring in child patients with developmental dysphasia. Despite considerable research effort, the diagnosis of neurodevelopmental disorders is still often made too late not only in autism (Oslejskova, 2007) but also in developmental language disorders. Binding process of synchronized and distributed activity is necessary for the mechanism of comprehension and seems to be altered not only in several psychiatric disorders (Bob, 2007) but also in neurodevelopmental disturbances including dysphasia.

We were using the knowledge of phonetics, acoustics and ANN applications. The SOM were chosen for solving part of the project. New variants of the SSOM were tested theoretically and experimentally after the first experiments with Kohonen's SOM. Also, we focused on the SOM implementation onto the gate array (Bartu et al., 2006). Besides the previously cited advantages, this type of ANN is favorable for persons without an engineering background, primarily for the ability to visualize high-dimensional data samples in a low-dimensional display.

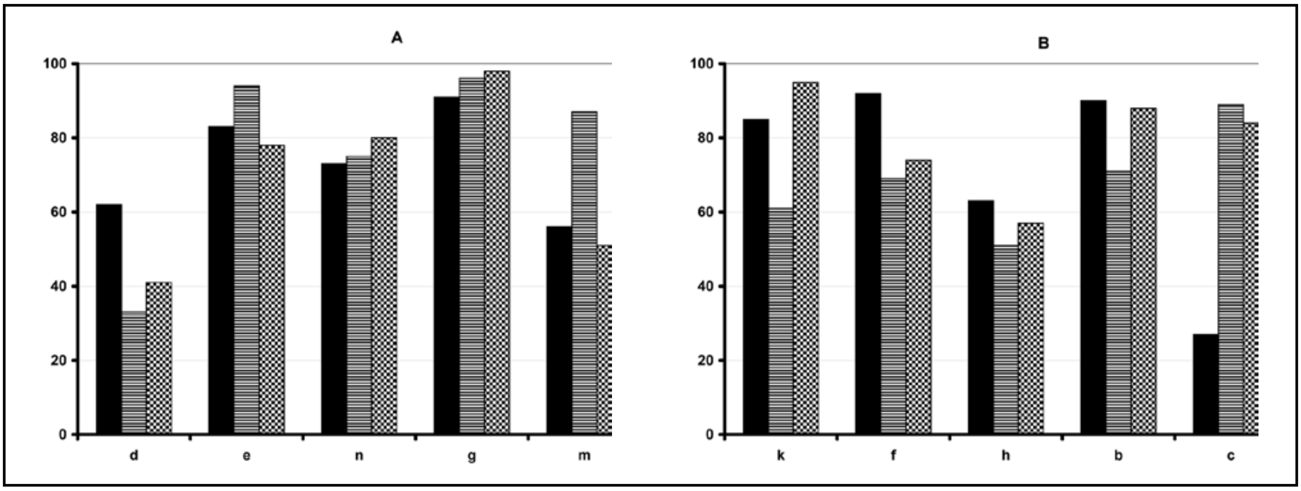


Figure 5. The histogram of the success in classification of the child patients. a) patients treated by medication, b) patients from a control group. The solid column presents 1st recording (before therapy), the line-hatched column presents 2nd recording (after first part of therapy), the cross-hatched column presents 3rd recording (after second part of therapy).

Correctness of the results from neural network output is dependent on the scope and quality of the training set, on the quality of speech units labeling, on the selection of relevant parameters (markers) for neural network training and on the neural network architecture.

Our experiments with SSOM promise a good chance for the creation of an automatic software pack including data preparation, neural networks training and evaluation and visualization module for clinical practice.

Our effort in future work will be focused on longer speech units (syllables, multi-syllable words). Incapability to formulate multi-syllable words (three and four syllables) or phoneme overlap faults, which are the other symptoms of the developmental dysphasia. Also, verbal dyspraxia, i.e. an obvious clumsiness in word repetition, is mentioned in Kohonen, 2001.

The processing of speech signals is complicated by the effect of the real environment (non-professional speakers, high noise in the environment - speech was recorded in ordinary rooms). The second problem that we have to solve is the fact that we are analyzing the children‘s speech. Often, its own development is not terminated for some age group or the quality of the utterances is strongly emotionally influenced. Also, we have at our disposal only a small amount of speech data, especially for patient. Though a database of child speakers is permanently kept. The size of the database of healthy child speech is limited also by the possibilities of data recording in the preschool and primary school institutions, especially concerning with parent permissions. We assume that it would be necessary to open a sizable screening project during preventive medical check-ups of small children.

The labeling (an extraction of particular phonemes from the utterances) is predominantly hand-operated for the present. An automatic system based on the ANN is being developed.

Acknowledgement

This research is a joint project of the Dept. of Circuit Theory in FEE-CTU in Prague and the Dept. of Paediatric Neurology in the 2nd Faculty of Medicine of Charles University in Prague. This research was partially supported by grants IGA MZ CR-NR 8287-3 / 2005 (clinical part) and by grants GACR 102/05/0278 (experiments).

REFERENCES

1 Bob P. (2007). Consciousness and co-consciousness, binding problem and schizophrenia. Neuroendocrinol Lett; **28**: 723–726.

2 Bartu M, Tuckova J, Stastny J. (2006). An Accelerator for Kohonen Self-Organizing Maps. Proc. of the Digital Technologies Intern. Workshop, Žilina, ISBN80-8070-624-7 (Book of Abstracts), ISBN 80-8070-637-9 (CD ROM).

3 Duvelleroy-Hommet C, Bilard C, Lucas B,Gillet, P, Barthes M.A, Santini M.J. et al.(1950). Sleep EEG and Developmental Dysphasia: Lack of a Consistent Relationship with Paroxysmal EEG Activity During Sleep. Neuropediatrics, **26**:14–18.

4 Georgopoulos V.C, Malandraki G.A, Stylios C.D. (2003). A fuzzy cognitive map approach to differencial diagnosis of specific language impairment. artificial Intelligence in Medicine. **29/3**: 261–278. ISSN: 0933–3657.

5 Hrncir Z, Komarek V. (2004). Analyses of EEG Recordings. Neural Network World, Int. Journal on Non-Standard Computing and Artificial Intelligence. Ed. M. Novak. **14**, 1/2004: 21–25.

6 Kohonen T, Hynninen J, Kangas J, Laaksonen J. (1995). SOM_PAK: The Self-Organizing Map Program Package. Helsinki University of Technology, Lab. of CIS. Available via anonymous ftp at internet address cochlea.hut.fi (130.233.168.48).

7 Kohonen T. (2001). Self-Organizing Maps. Ed.:Huang T. S, Kohonen T, Schroeder, 3rd ed. Springer-Verlag Berlin, ISBN 3-540-67921-9.

8 Menkes J.H, Moser F.G. (2006). Neurological examination of the child and infant. Text book of child neurology. 7th ed, Lippincott.

9 Ngan S-C, Yacoub ES, Auffermann W.F, Hu X. (2002). Node merging in Kohonen's self-organizing mapping of fMRI data. Artificial Intelligence in Medicine. **25/1**:19-33. ISSN: 0933–3657.
10 Oslejskova H, Kontrova I, Foralova E, Dusek L, Nemethova D. (2007). The course of diagnosis in autistic patients: the delay between recognition of first symptoms by parents and correct diagnosis. Neuroendocrinol Lett; **28**: 895–900.
11 Palkova Z. (1994). Fonetika a fonologie češtiny. [Phonetics and phonologics of the Czech language (in Czech)]. Univerzita Karlova-Praha, ISBN: 80-7066-843-1.
12 Pospíšilová, L. (2005). Diagnostické otázky u vývojové dysfázie. [Diagnostics Questions of Developmental Dysphasia. (in Czech)]. Vox pediatrie, journal of general practitioner for children and young. **1/5**:25–27.
13 Psutka J, Müller L, Matousek J, Radova V. (2006). Mluvíme s počítačem česky. [Talking with Computer in Czech. (in Czech)]. Academia Praha, ISBN 80-200-0203-0.
14 Tuchman F.R, Rapin I, Shinnar S. (1991a). Autistic and Dysphasic Children I: Clinical Characteristic. Pediatrics. **88**: 1211–1218.
15 Tuchman F.R, Rapin I, Shinnar S. (1991b). Autistic and Dysphasic Children II: Epilepsy. Pediatrics. **88:** 1219–1225.
16 Tuckova J, Zetocha P. (2006). Speech analysis of children with developmental dysphasia by Supervised SOM. International Journal on Neural and Mass-Parallel Computing and Information Systems "Neural Network World", Ed. M.Novak, ICS AS CR and CTU,FTS. **16/6**: 533-545. ISSN 1210-0552.
17 Tuckova J, Zurek M.(2007). Automatic Segmentation of the Speech Signal by Artificial Neural Networks. 8th Int. WSP ECMS 2007 & Doctoral School (EDSYS, GEET), Liberec, Czech Republic.
18 Vesanto J, Himberg J, Alhoniemi E, Parhankangas J. SOM Toolbox for Matlab 5, SOM Toolbox Team, Helsinki University of Technology, Finland, www.cis.hut.fi/projects/somtoolbox.
19 Zwicker E, Fastl H. (1990). Psychoacoustics. Facts and Models. Springer Series in Information Sciences. Ed.: Huang T.S, Kohonen T, Schroeder M.R, Springer-Verlag Berlin, Heidelberg, New York, ISBN 3-540-52600-5.