Features extraction for the automatic detection of ALS disease from acoustic speech signals

Maxim Vashkevich, Elias Azarov, Yuliya Rushkevich, Alexander Petrovsky*

Belarusian state university of informatics and radioelectronics (BSUIR), Minsk, Belarus

Computer Engineering Department

**Republican research and clinical center of neurology and neurosurgery, Minsk, Belarus*

Aim of the work

Goal

Development of feature extraction methods for detection of pathological changes in acoustic speech signal for the diagnosis of the bulbar form of Amyotrophic Lateral Sclerosis (ALS).

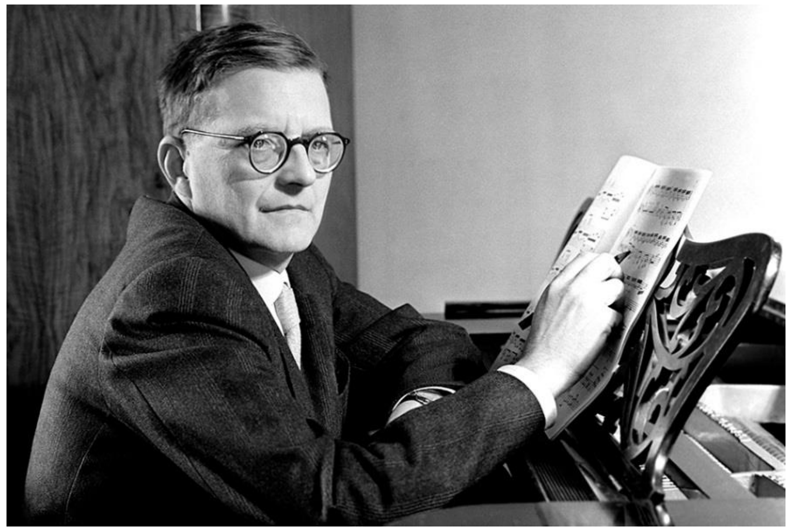
Actual problems

1) early diagnosis of ALS (bulbar changes (i.e. difficulty with speech or swallowing) are the first symptoms in approximately 30% of persons with ALS);
2) monitoring of the ALS progression;
3) optimization of the efficacy of medicinal treatment of ALS.

Amyotrophic Lateral Sclerosis

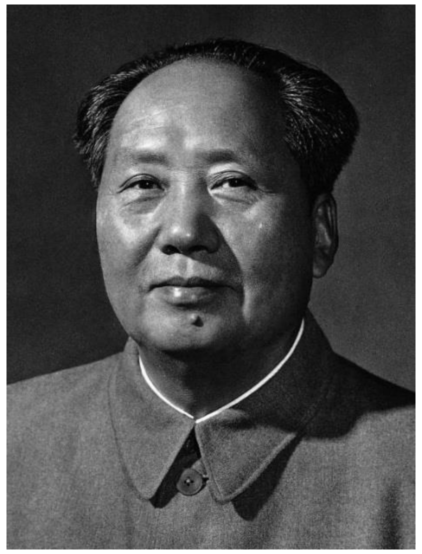
ALS – degenerative progressive disease of the central nervous system, accompanied by muscle weakness of limbs and trunk, swallowing and speech disorders (pronounced words are broken, speech is slowed, sonority changes).

Famous persons suffering from amyotrophic lateral sclerosis

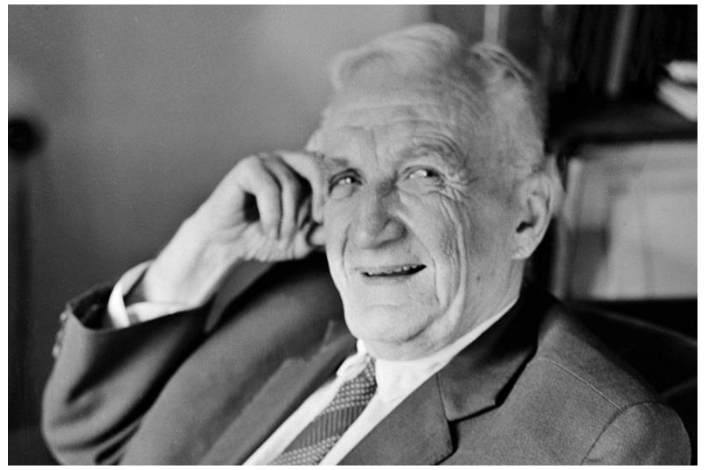


Dmitry Shostakovich (1906-1975)

Stephen Hawking (1942-2018)



Mao Zedong (1901-1972)



Igor Tamm (1895-1971), Soviet physicist (Nobel Prize, 1958)

1. Speech subsystems

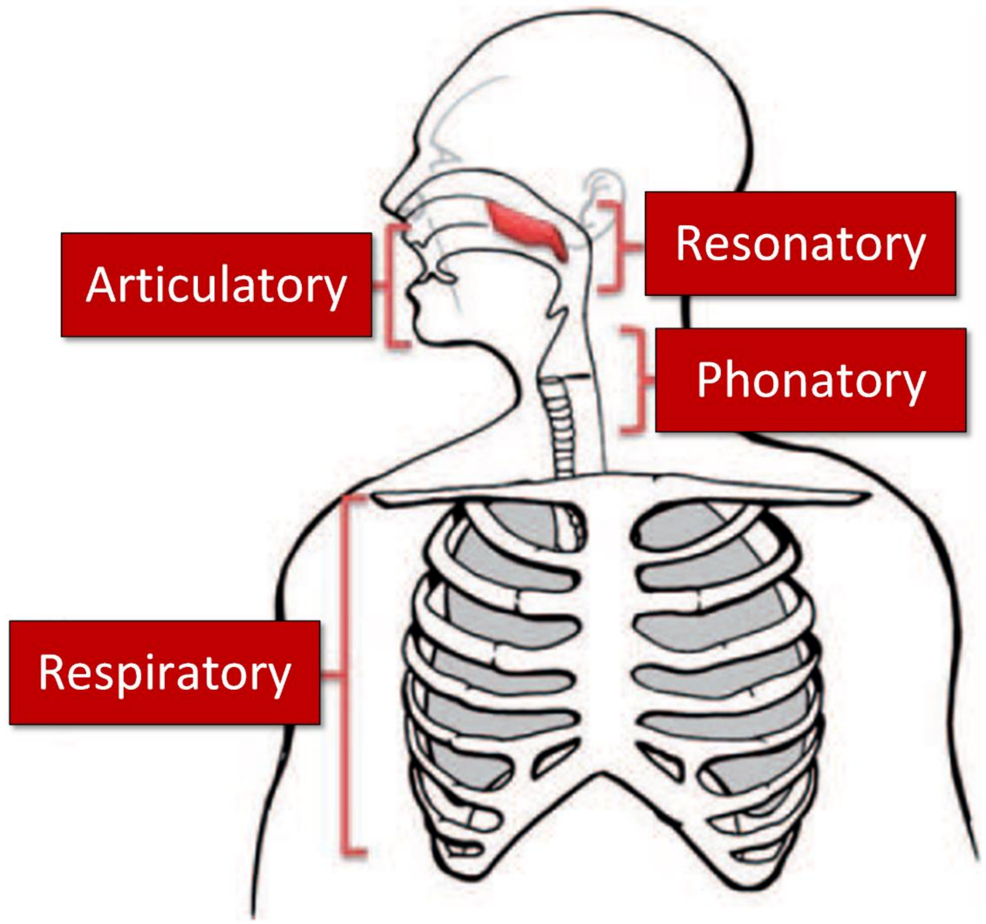
Speech production network comprised of 4 subsystems:

1) ***respiratory***

2) ***phonatory***

3) ***articulatory***

4) ***resonatory***



ALS impacts to all of this subsystems and it is leads to

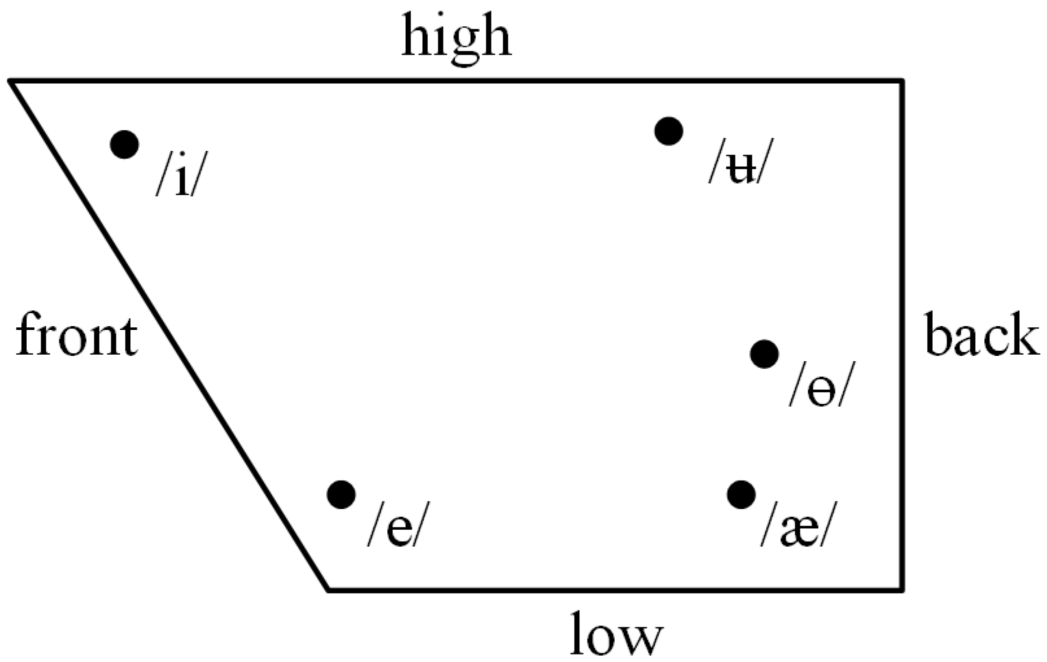
– ***dysphonia*** (harshness of sound);

–***dysarthria*** (imperfect articulation);

– ***decreasing speaking rate.***

2. Foundation of vowels selection

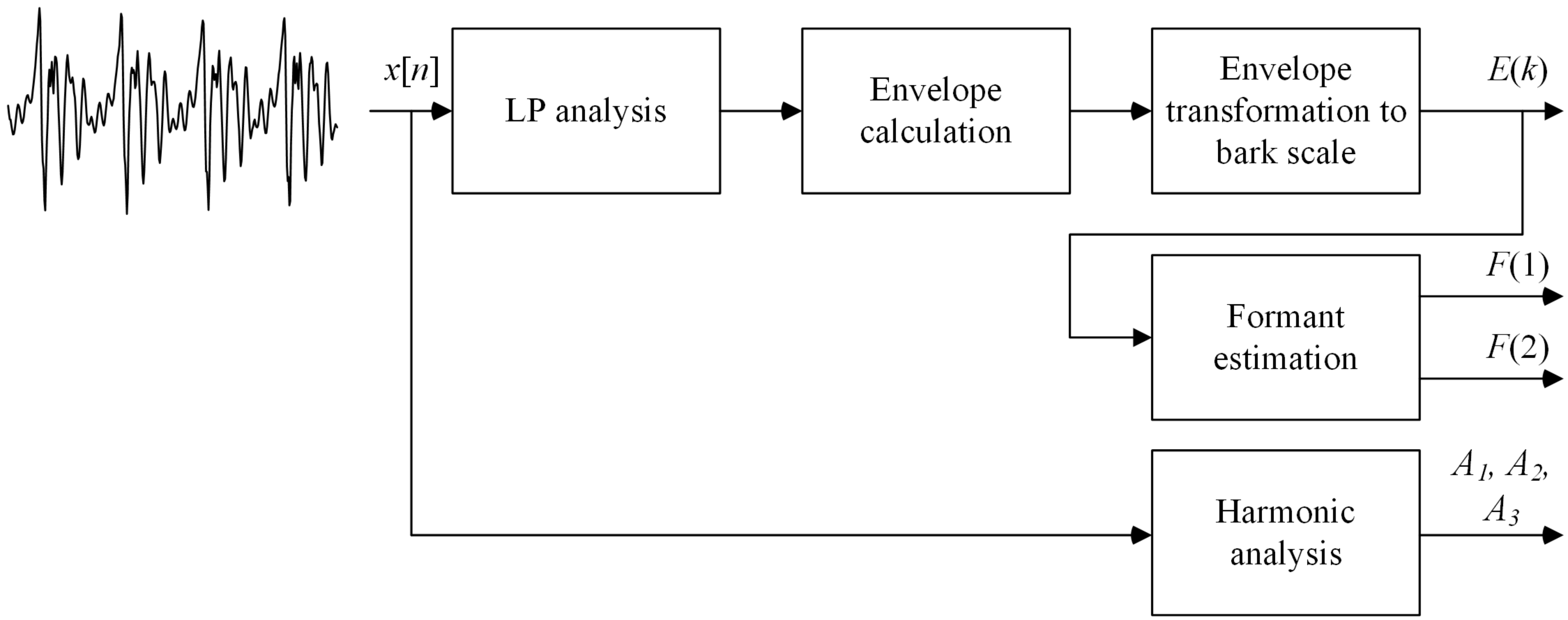
– schematic characterization of Russian vowels (/i/, /ʉ/, /ɵ /, /æ /, /e /) in terms of relative tongue surface (*front* ↔ *back* and *high* ↔ *low*) positions



– For detecting symptoms of ALS it is advisable to select the vowels /**æ**/ and /**i**/.

– Records with counting from 1 to 10 (in Russian) were used as a material for experiments; As a rule, sounds were selected from the words "*od**i**n*", "*dv**æ***", "*tr**i***".

3. Scheme of speech signal analysis



– As a result following parameters extracted:

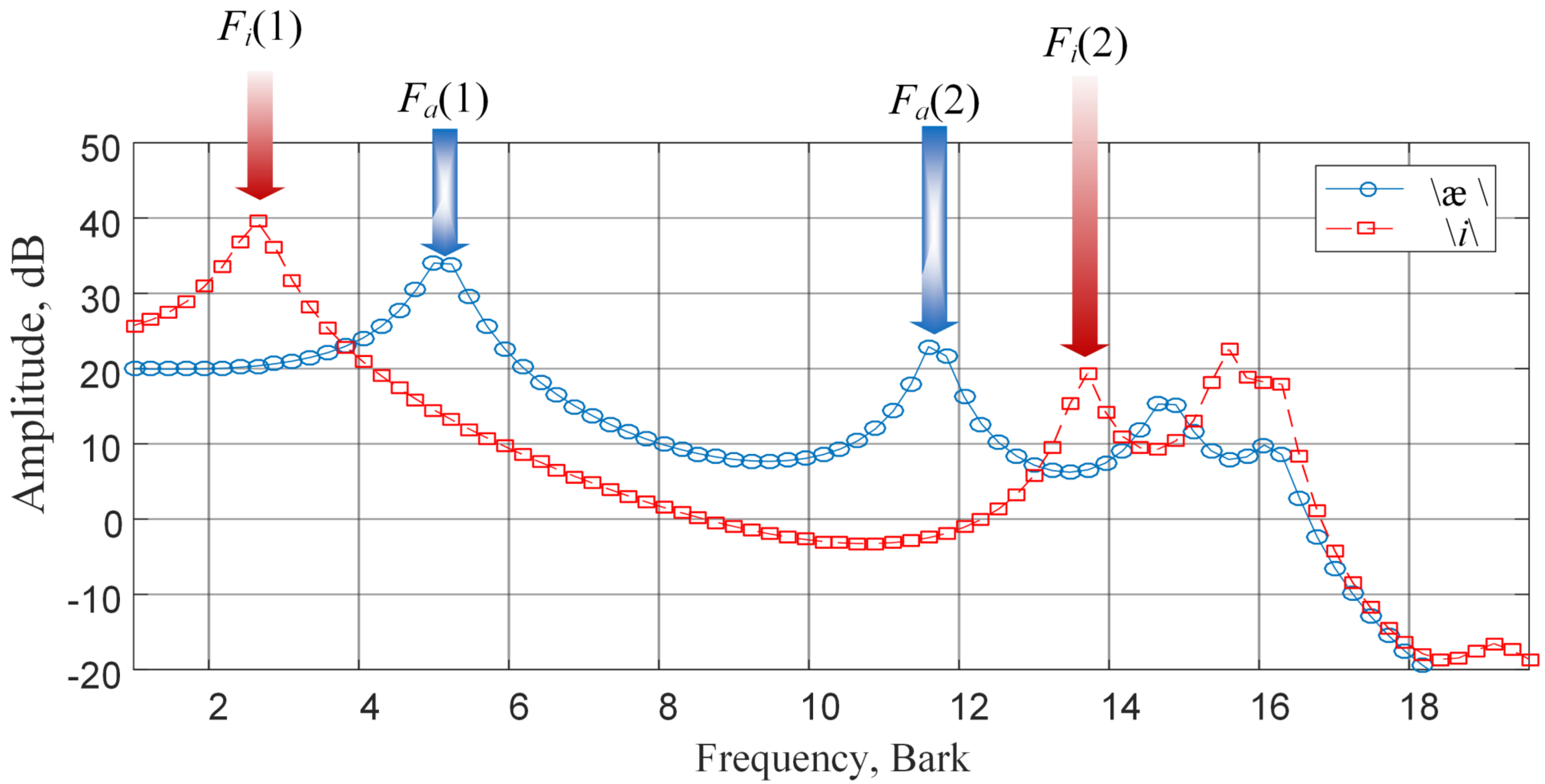
1) Spectral envelope (in Barks) $E(k)$;

2) Formants $F(1)$ и $F(2)$;

3) Amplitude of firsts three harmonics A_1, A_2 and A_3.

4. Mutual location of the formant frequencies (1)

– ***Peculiarity***: the patient's speech is compared not with the "standard", but with his own speech.

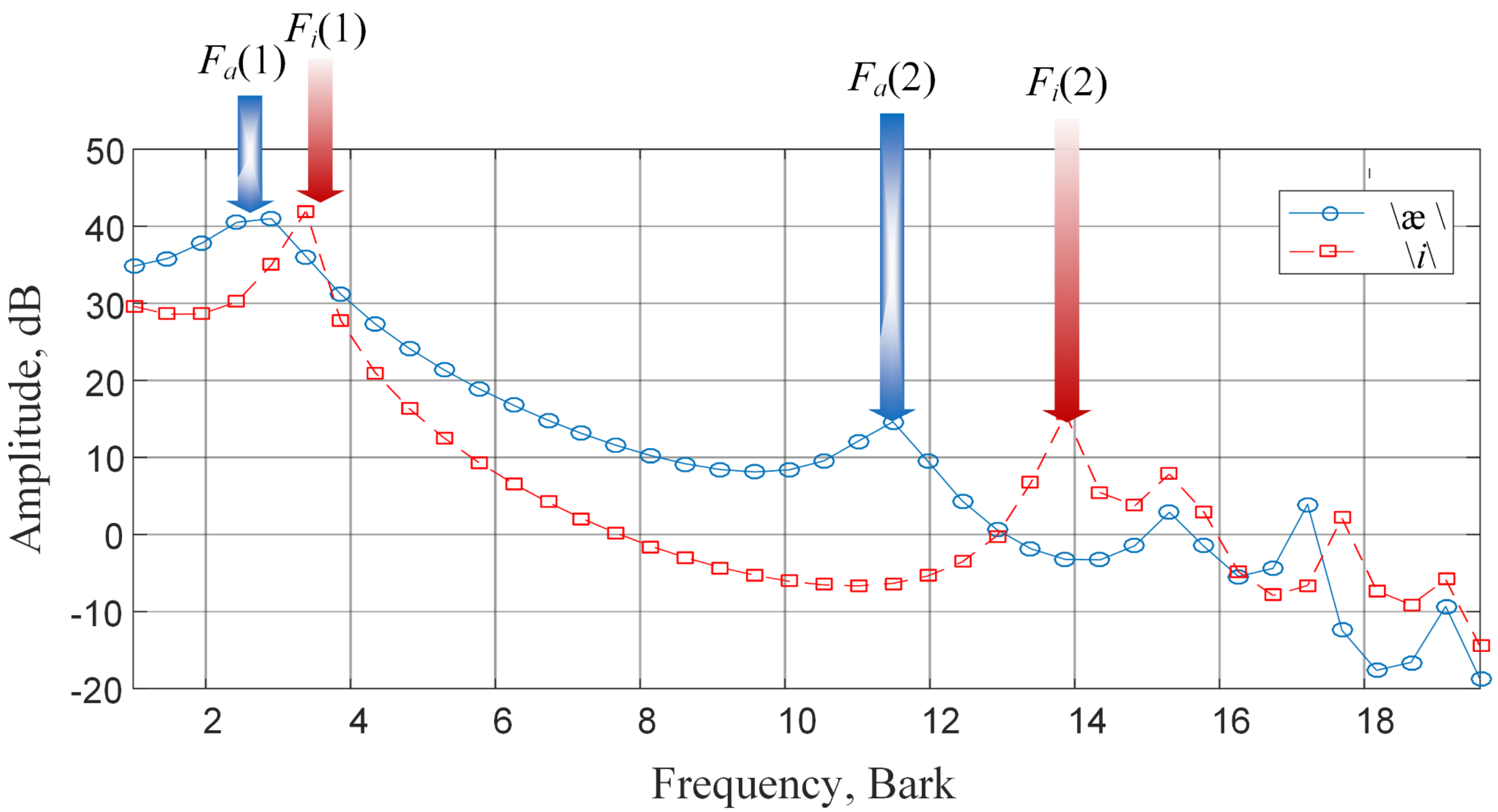


Formant structure of the vowels /æ/ and /i/ (healthy person)

– ***Formant order***: $F_i(1) < F_a(1) < F_a(2) < F_i(2)$.

5. Mutual location of the formant frequencies (2)

For patients with ALS mutual location of the formant frequencies can be violated.

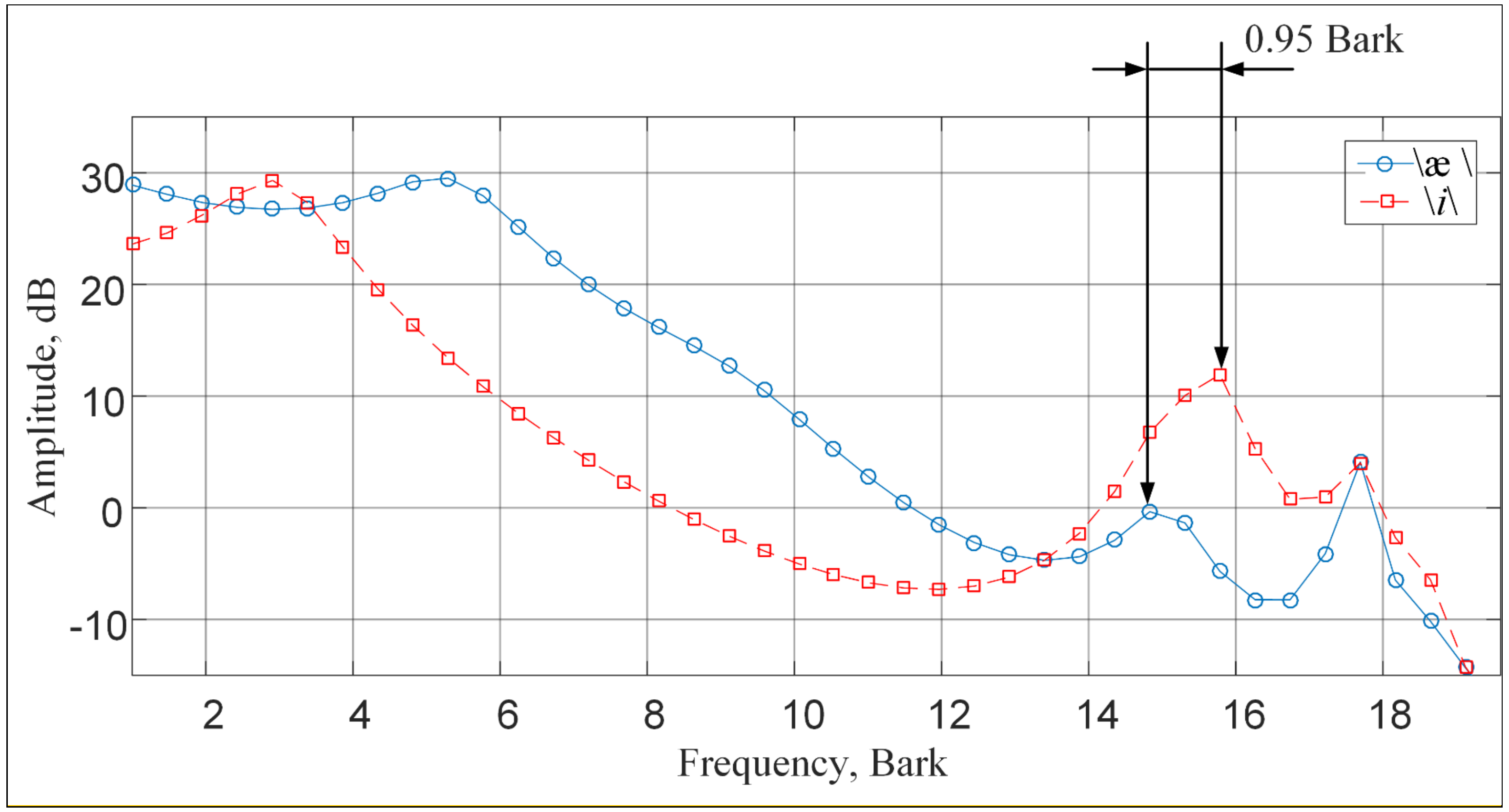


Abnormal mutual location of the formant frequencies (patient with ALS)

– ***Formant order (norm)***: $F_i(1) < F_a(1) < F_a(2) < F_i(2)$.

– ***Formant order (pathology)***: $F_a(1) < F_i(1) < F_a(2) < F_i(2)$.

6. Convergence of formant frequencies



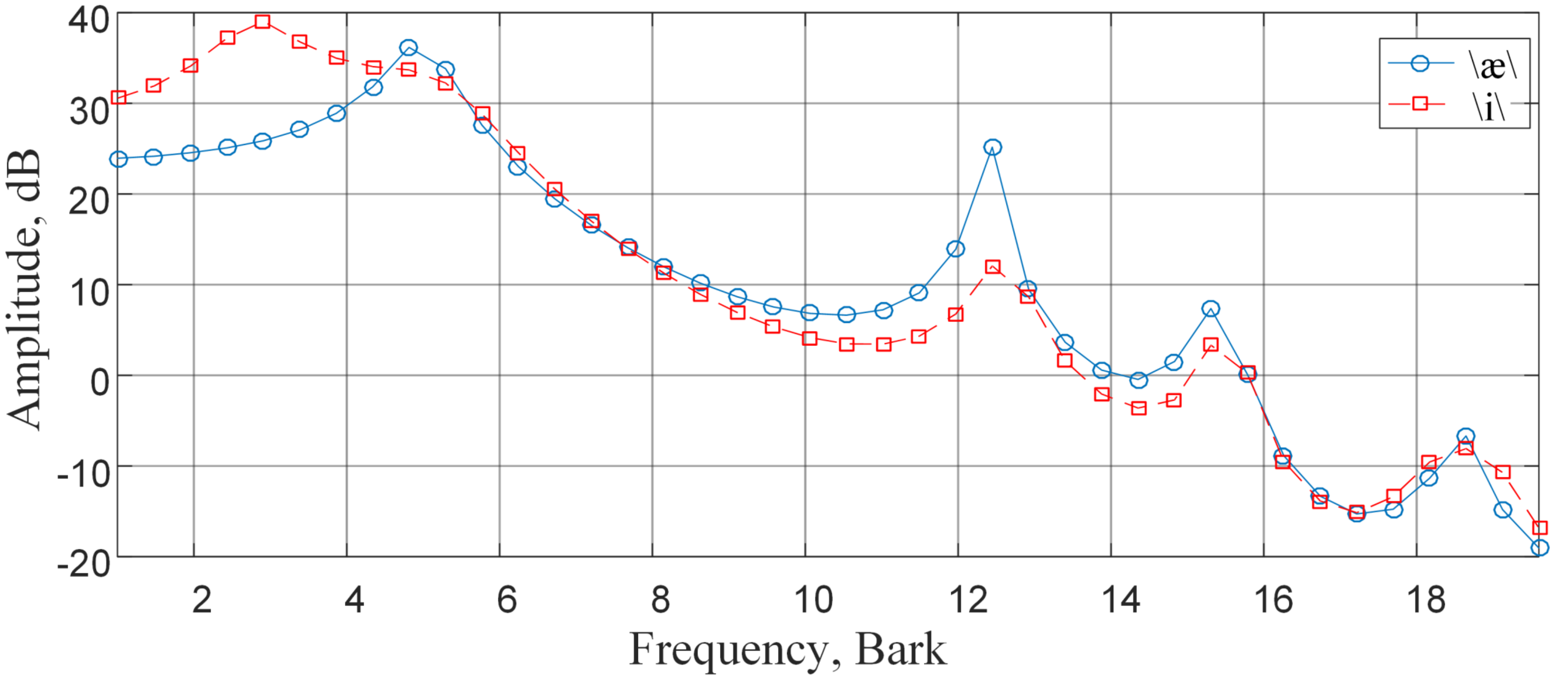
In the case when the normal order is not violated, there is often a significant convergence of the formant frequencies.

To quantify the degree of violation of the mutual formant structure following measure is proposed

$$fmt_{err}(F_i, F_a) = \begin{cases} 2, & \text{if} \quad F_i(1) > F_a(1) \quad or \quad F_a(2) > F_i(2) \\ 2 - \dfrac{F_a(1) - F_i(1)}{2} - \dfrac{F_i(2) - F_a(2)}{2}, & \text{if} \quad F_i(1) - F_a(1) < 2 \;\; and \;\; F_i(2) - F_a(2) < 2 \\ 1 - \dfrac{F_a(1) - F_i(1)}{2}, & \text{if} \quad F_a(1) - F_i(1) < 2 \\ 1 - \dfrac{F_i(2) - F_a(2)}{2}, & \text{if} \quad F_i(2) - F_a(2) < 2 \\ 0, & \text{othrewise} \end{cases}$$

7. Similarity of the envelope sounds /æ/ and /i/

– Joint analysis of envelopes of vowels /æ/ and /i/ of persons with ALS have revealed an increased similarity between the shapes of these envelopes.

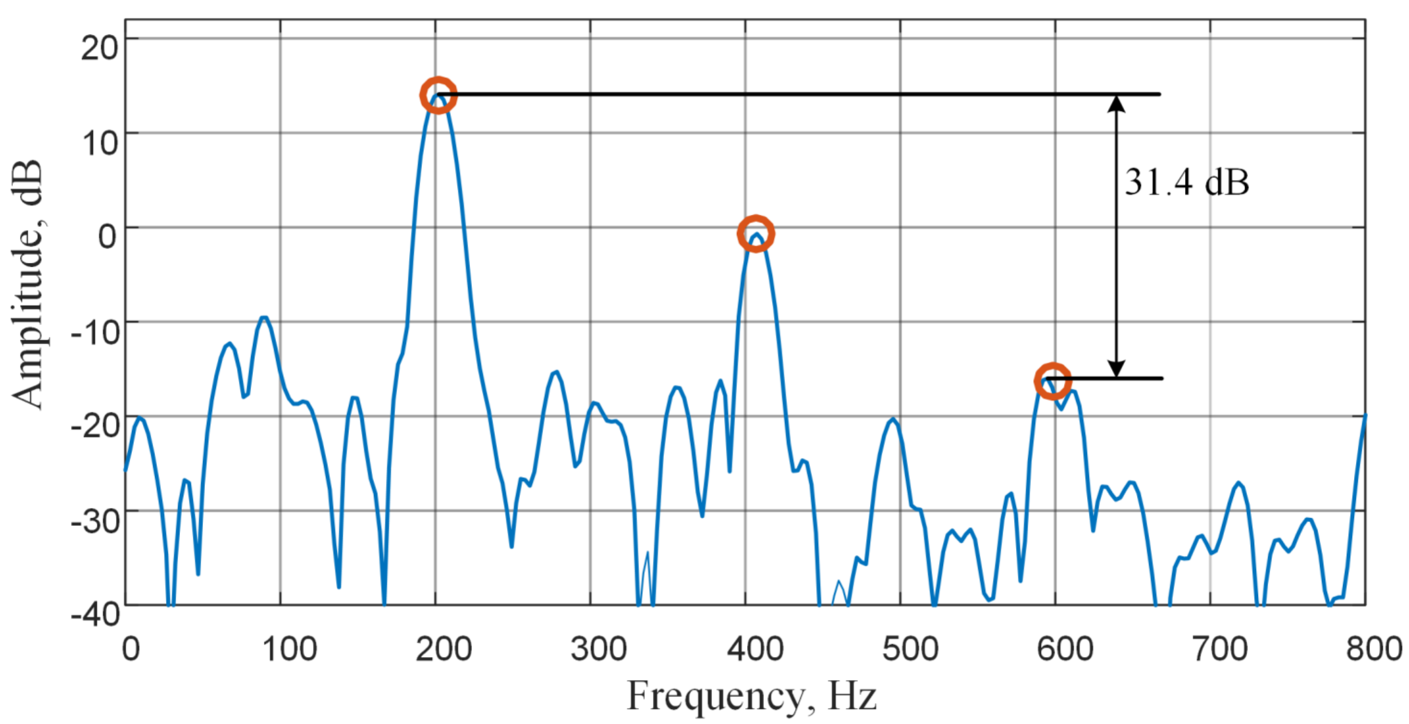


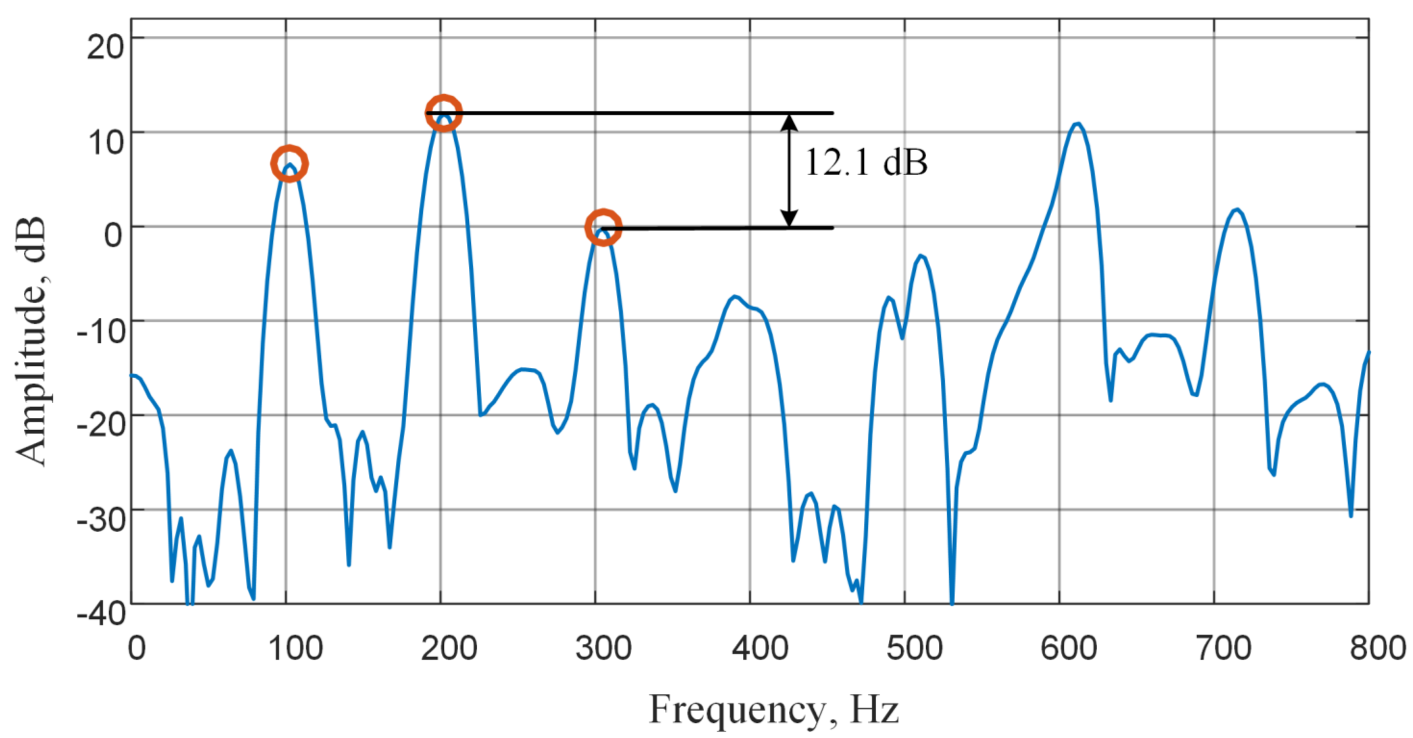
– To quantify the differences between the envelopes of vowels /æ/ and /i/ it is suggested to use the l_1-norm distance measure:

$$d_1(E_i, E_a) = \sum_{k=1}^{P} |E_i(k) - E_a(k)|.$$

8. Difference in the amplitudes of the harmonics (/æ/)

– Analysis of harmonic structure of vowel /æ/ of ALS speech have revealed that dysphonic disorder affects first three harmonic components.





– To quantify the degree of deviation in amplitude structure the following measure is proposed:

$$harm_{diff}(A_1, A_2, A_3) = \max(A_1, A_2) - A_3.$$

9. Data collection

– ***Organization***: Republican Research and Clinical Center of Neurology and Neurosurgery (Minsk, Belarus).

– ***Contingent***: 48 speakers

Category	male	female	total	Mean age	SD age
Healthy	15	7	22	36.3	9.5
ALS	14	12	26	56.5	10.5

– ***Equipment***: The samples recorded at 44.1 kHz using smartphone with a standard headset and stored as 16 bit uncompressed PCM files.

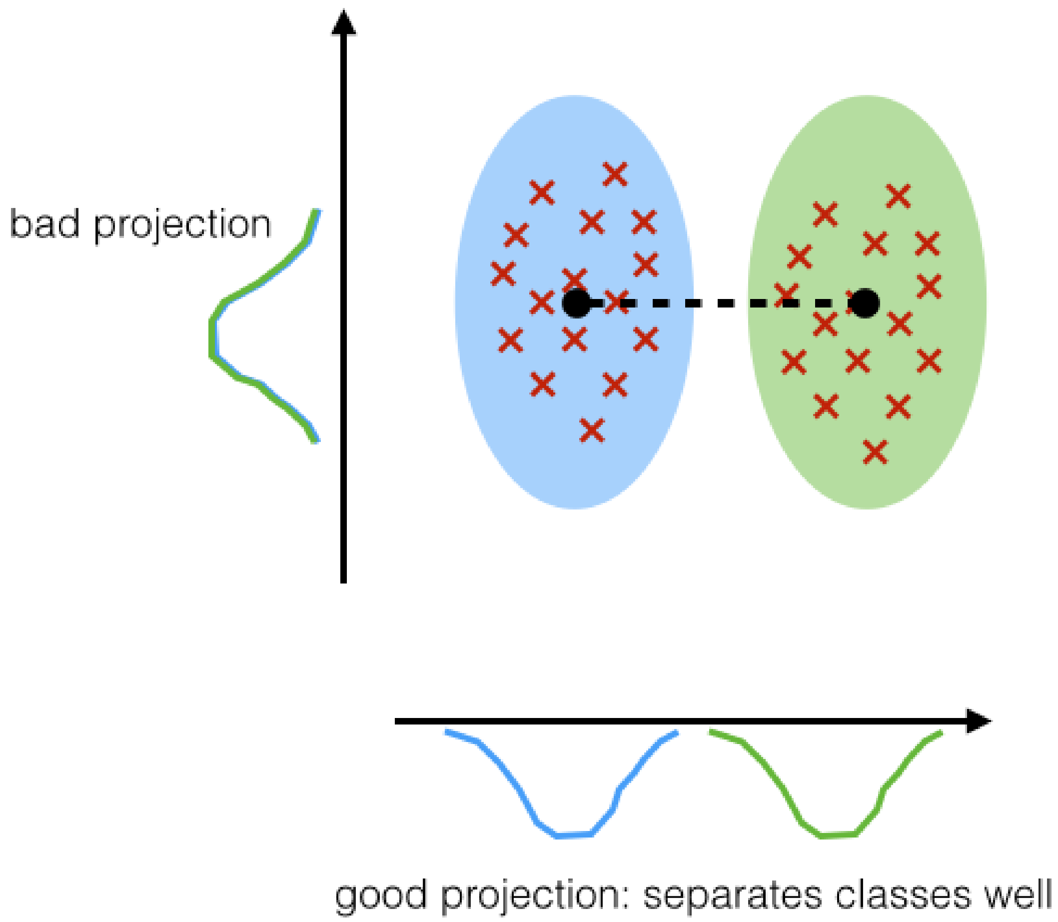
– ***Data***: For classification purpose 106 pairs of vowels /æ/ and /i/ were manually pre-segmented prior to feature extraction (61 – healthy, 45 – pathology).

10. Linear discriminant analysis

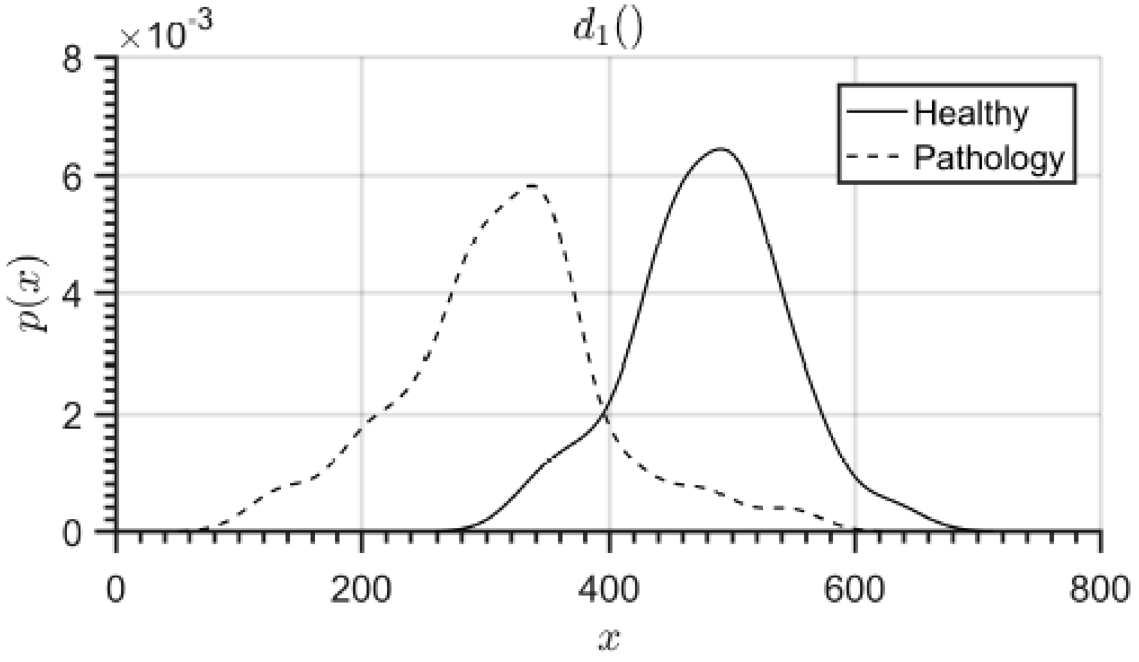
The idea of ***linear discriminant analysis (LDA)*** lies in the search for such a hyperplane $\mathbf{w}$ in the feature space, so that the projection of all training vectors onto it *minimizes the within-class variation and maximizes the between-class variation*.

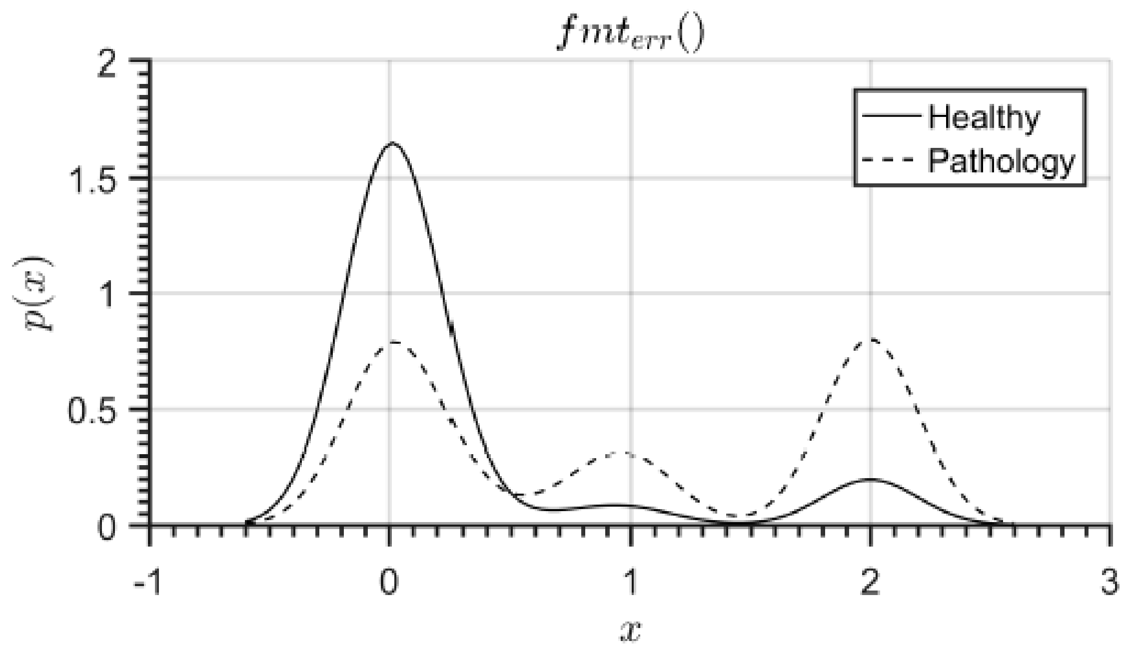
$$\mathbf{w} = \underset{\mathbf{w}}{\operatorname{argmax}} \frac{\mathbf{w}\mathbf{S}_B\mathbf{w}^T}{\mathbf{w}\mathbf{S}_W\mathbf{w}^T}$$

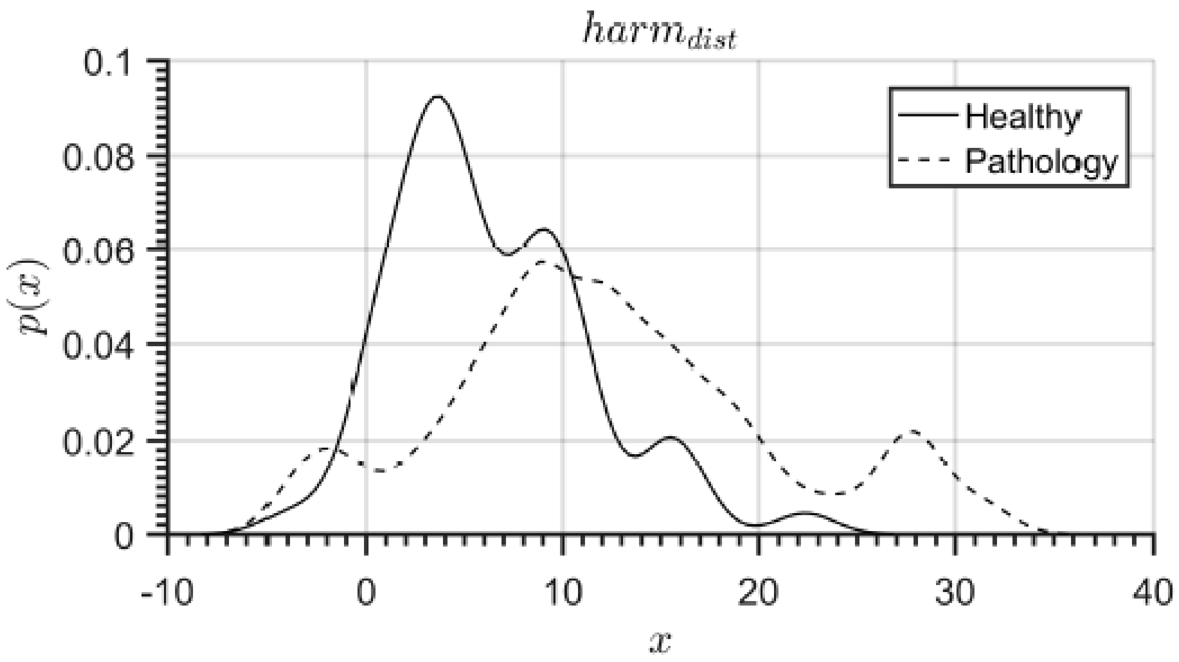
where $\mathbf{S}_B$ – between class scatter matrix, $\mathbf{S}_W$ – within class scatter matrix.



11. Statistical analysis of features

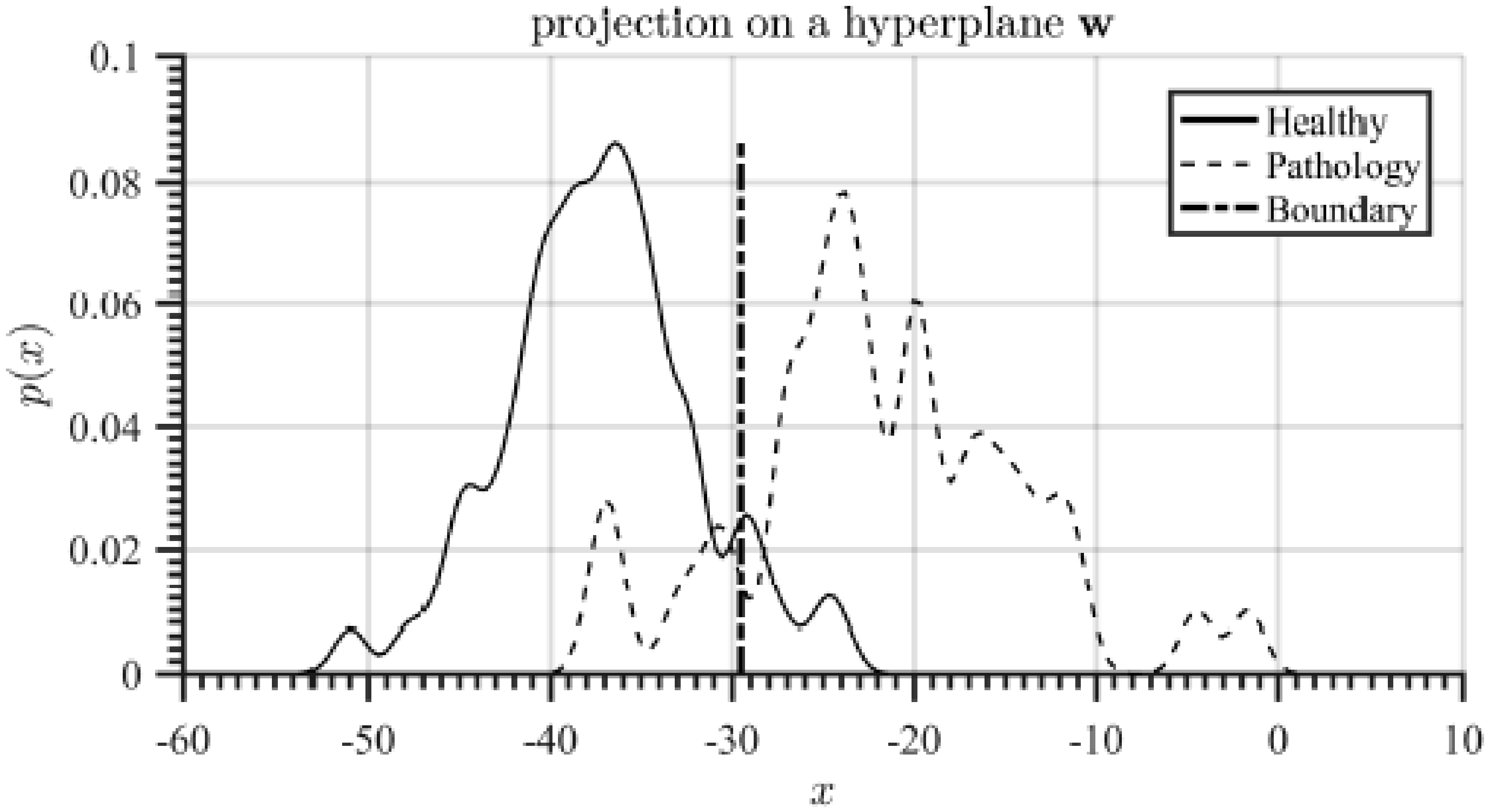






12. Experimental results

Kernel density function for projection on hyperplane of all train vectors.



Overall classification accuracy	88:0%
true positive	90:5%
true negative	84:6%.

13. Conclusion

– New features are based on analysis of

1) envelopes of vowels /æ/ and /i/;
2) mutual formant structure of vowels /æ/ and /i/;
3) harmonic structure of vowel /æ/.

– Usage of presented feature with LDA-based classier allows to achieve overall classification accuracy of 88%.

– Further work is necessary to improve classification results.

14. Gratitude

Thank you for attention!