Finding the Most Uniform Changes in Vowel Polygon Caused by Psychological Stress

Miroslav STANEK, Milan SIGMUND

Dept. of Radio Electronics, Brno University of Technology, Technicka 12, 616 00 Brno, Czech Republic

mirek.stanek@phd.feec.vutbr.cz, sigmund@feec.vutbr.cz

Abstract. *Using vowel polygons, exactly their parameters, is chosen as the criterion for achievement of differences between normal state of speaker and relevant speech under real psychological stress. All results were experimentally obtained by created software for vowel polygon analysis applied on ExamStress database. Selected 6 methods based on cross-correlation of different features were classified by the coefficient of variation and for each individual vowel polygon, the efficiency coefficient marking the most significant and uniform differences between stressed and normal speech were calculated. Using the mean of cross-correlation values received for area difference with vector length and angle can be classified as the best method for observing generated differences. Generally, best results for stress detection are achieved by vowel triangles created by /i/-/o/-/u/ and /a/-/i/-/o/ vowel triangles in formant planes containing the fifth formant F5 combined with other formants.*

Keywords

Speech processing, emotion recognition, psychological stress, vowel polygons

1. Material and Methods

Generally accepted meaning of the term stress is tension, pressure and strain. By this reason, stress can be briefly defined as the state of organism during which the subject is faced to extraordinary conditions and classified as an emotion leading to impact the human behavior. Basically, two types of stress are recognized [1]. The first type is so-called eustress stimulating the subject to better performance as the reaction on positive load. Conversely, distress is the second stress type known as the negative reaction on the overload leading to disease, damage or subject destruction. Previous statement gives the testimony of stress generation caused by external objects, so-called stressors, further divided into five main groups: psychical, physic, social, traumatic and children's. Differences within stressor types and their description can be found in [2].

The main motivation of this paper is to present a novel method to psychological stress detection in speech by using vowel polygons, the set of chosen formants grouped into various formations, which can be further applied on other emotions for reaching possibly useful results.

Recently in this field, various tools are utilized for stress detection as well as approach based on the similarity of speech feature, e.g. set introduced by Kurniawan [3] using pitch, MFCCs, Relative Spectral Transform-Perceptual Linear Perception (RASTA-PLP), other biomedical features and Support Vector Machine (SVM) classifier. Kurniawan also points the efficiency between using MFCCs and MFCCs together with pitch is more or less equal. Another method for speech under stress classification is presented by Johari et al., where variances of possibly deployed wavelet filters are used for energy and entropy achievement, which is further classified by SVM and Linear Discriminant Analysis (LDA) [4] applied on SUSAS database [5]. The description of another interesting emotion, including stress, classifier developed for call centers can be found in [6], where the best developed classifier is based on SVM and uses so-called Pearson Correlation relevant to the set of selected features. Further publications describing the set of features containing LPC spectrum of residual and auxiliary muscle tension ratio [7], spectrograms and Sigma-pi Neuron [8], autocorrelation envelope, fundamental frequency, formants and MFCCs [9] present possible methods to psychological stress detection in speech. Vowel polygons have not been used yet as the speech feature in this field not even for Czech language. Recently, only the determination of formant feature depending on actual emotional state was observed to vocal tract description in Czech and Slovak language [10] which is related to presented topic.

Generally, a short survey oriented on used speech features and classifiers for stress and emotion recognition can be found in [11]. As it can be seen by this review, mostly used classifiers are Hidden Markov Models, SVM and Gaussian Mixture Models. The list of mostly used speech features is also unchanged, thus LPCs, MFCCs, energy, formants and pitch. Quite huge review oriented on psychological stress included in speech is written by Gid-

dens et al. [12]. This review provides perfectly processed presentation of recent work in the field of stressed speech and the survey of recently used speech parameters. Stress patterns are also described in detail as well as achieved results by various authors, their final summarization and conclusion of further work. Another survey of used methods, databases and mined results can be found in [13], but this publication is mainly oriented on emotion recognition in speech, so the stress topic is described briefly.

In the case of changes in speech caused by psychological stress, the pitch variation depending on stressed speaker's mood is described in more details in [14], where variable increase in speaker's pitch was investigated depending on stress level. Similar experiment was made by Tse et al. [15] where relation between fundamental frequency and its standard deviation under psychological stress was observed by two experiments. Firstly, the presentation was performed by voluntary speakers in their self-comfort mood, but the second experiment had the condition of minimal pitch variation. By this experiment the fact was proved, that the speech and its parameters can be successfully self-controlled by speaker despite the psychological stress influence. Speech fundamental frequency was also used in another experiment for obtaining the interaction between pitch of stressed speech and its long-term averaged spectra for validation support of a reactivity dimension in schizophrenia [16]. By another speech under stress analysis, the differences between lower and higher stress level were observed [17], exactly higher word productivity is occurred in speech under higher stress level as well as more rests during speech [18].

2. Stress Influence

Obvious signs of vowel polygon changes depending on normal and stressed state of speaker are observed in two criteria. Firstly for each vowel polygon, the area differences between actual (stress) and original (normal) are observed for investigating the possible uniform behavior of this parameter as well as the direction and length of vector facing from original to actual Centre of Gravity (CoG). Figure 1 shows generated vectors for AEI vowel triangle observed in formant plane F3-F4 for high level psychological stress.

For the majority of all possible vowel polygons, the same effects are occurred as well as for illustrated example (see Fig. 1). Firstly, created vectors are mostly uniform in
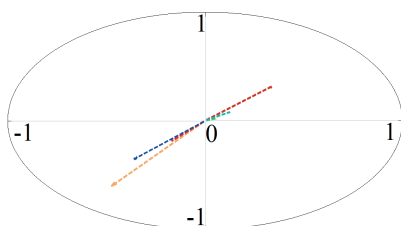

Fig. 1. Differences between AEI34 created vectors' length and direction.

their direction for high stress influence, and their angle reaches approximately value ±π/4. Generally, stress influenced vectors are not occupied in the second and fourth quadrant. By these statements and previous research [19], the increasing direction uniformity of created vectors can be assumed with increasing stress level which leads to erasing the deviations between speakers.

Following observations are focused on getting the cross-correlation values between vowel polygon area difference and one parameter of created vector. These values are also further statistically analyzed by coefficient of variation R defined as follows

$$R_x = \left| \frac{\sigma_x}{\bar{x}} \right| \qquad (1)$$

where σ_x is standard deviation of observed parameter x (e.g. cross-correlation values of selected vowel shape over all formant planes) and $\bar{x}$ is its mean value. This statistical pointer shows higher uniformity of received results by lower number leading to more reliable and significant results [20].

3. Applied Methods

Presented research was applied on previously described database ExamStress [21], exactly on randomly selected 10 male Czech native speakers telling the same text during and after final exam, which means that two identical records differing only in emotional state are received for each speaker.

These records represent the input of developed and further used software system generating and analyzing vowel polygons [22]. Briefly, each input sound record is resampled to $f_s = 8$ kHz, and further vowels are recognized from fluent speech by using two-level recognition system (Mahalanobis distance, Forward-feed Neural Network), retroactively checked [23] and the values of all occupied formant frequencies in each vowel are saved for further processing. In the case of used sampling frequency, at most five formants can be observed in LPC spectrum, which leads to the total number of ten possible formant planes. As it was mentioned, presented research is oriented on Czech language containing five vowels /a/, /e/, /i/, /o/, /u/ and their so-called long equivalents differing only in duration not in pronunciation. The total number of five Czech vowels leads to sixteen different shapes (ten triangles, five tetragons and one pentagon) which can be investigated. These shapes situated in formant planes are called vowel polygons and their generation, marking and other information can be found in [24]. Recently, vowel polygons, mainly called as vowel spaces, were used in other fields of speech processing, e.g. achieving children age differences [25], whisper analysis [26] and observation of the Parkinson disease [27], but not applied on stressed speech. The presented method can be also possibly useful to active hypoxia level detection [28].

Differences between normal and stressed vowel polygons are based on previously described formant behavior [29]. The core of provided experiments is created by cross-correlation of chosen vowel polygon's parameters couples for achievement of obvious relations between them. Nowadays, the cross-correlation is ordinary used in the speech processing in the field of emotion recognition [30], speech [31] and speaker identification [32]. Following results are obtained for six different couples of cross-correlated parameters. For simplification in the following text, these couples are represented as used experimental methods. The first couple signs the cross-correlation of difference area value and vector length (Method 1), the signum of area difference and vector length (Method 2), the area difference value and vector angle (Method 3), the signum of area difference and vector angle (Method 4). Method 5 is defined as the mean of method 1 and method 3. Method 6 represents the mean of method 2 and method 4.

4. Experimental Results

4.1 Cross-correlation

Used stress-influenced records were spoken by master students and captured before trying to pass oral final exam. Generally, experimental results and processes presented in this subsection are captured for 10 male Czech native speakers before and after master thesis defense faced to examination board. Due to possible option which can lead to striking failure of current situation, the stressor's pressure is very intensive on observed subject leading to high stress level situation [33].

Table 1 contains experimentally achieved values for each parameter by Method 5 as an illustration of reached ratios depending on selected formant plane over all possible formant polygons. Obviously, this method is characterized by more or less stable values of all parameters and very satisfactory R values. The formant plane F3F4 can be selected as the most suitable for psychological stress detection due to reaches almost the highest mean value of calculated cross-correlation (slightly significant positive dependency), the smallest standard deviation value leading

to the most uniform results over all vowel polygons in this plane. The worst results have been reached by formant plane F1F2 which can be explained by the importance of the first and second formant (F1 and F2) to vowel, not to emotion or speaker characterization.

Average values of observed parameters reached for each method are summarized in Tab. 2. By the comparison of all average results, both mean methods (Method 5 and Method 6) can be classified as the most stable in formant criterion, i.e. over all possible vowel polygons. This fact is based on the smallest values of standard deviation and coefficient of variation leading to very uniform results in each individual formant plane independent on selected vowel polygon. On the other hand, Method 4 seems to be absolutely useless because the highest R value signs the highest dependency on the selected formant plane in the case of psychological stress detection.

Method number	Cross-correlation [-]		$\bar{R}$
	$\bar{\mu}$	$\bar{\sigma}$	[-]
1	0.2953	0.3218	1.2058
2	0.2830	0.3016	1.1690
3	-0.2434	0.2880	1.7907
4	-0.1490	0.3108	11.7649
5	0.3690	0.1539	0.4257
6	0.3336	0.1582	0.4756

Tab. 2. Experimental results over all vowel polygons further averaged over all formant planes.

Similarly to results contained in Tab. 1, results depending on selected vowel polygon (over all formant planes) are listed in Tab. 3 for Method 5. Obviously, the format plane-independent criterion gives more stable results than in the previous case which is characterized by lower R and standard deviation values. Obviously, the AIU vowel triangle has reached the most uniform results over all formant planes and due to this reason it can be seen as the most proper vowel polygon to formant plane independent stress detection. The worst value has been achieved by AEI vowel triangle.

Formant plane	Cross-correlation [-]		$\bar{R}$
	$\bar{\mu}$	$\bar{\sigma}$	[-]
F1F2	0.3073	0.1902	0.6189
F1F3	0.3787	0.1606	0.4240
F1F4	0.4202	0.1699	0.4044
F1F5	0.3883	0.1258	0.3239
F2F3	0.3576	0.1604	0.4485
F2F4	0.3202	0.1659	0.5180
F2F5	0.3478	0.1524	0.4382
F3F4	0.4147	0.1120	0.2701
F3F5	0.4079	0.1379	0.3382
F4F5	0.3469	0.1642	0.4733
Average	0.3690	0.1539	0.4257

Tab. 1. Experimental results over all vowel polygons averaged in each formant plane for Method 5.

Vowel polygon	Cross-correlation [-]		$\bar{R}$
	$\bar{\mu}$	$\bar{\sigma}$	[-]
AEI	0.2215	0.1199	0.5414
AEU	0.4869	0.1248	0.2563
AEO	0.4123	0.1859	0.4504
AIO	0.5317	0.1657	0.3116
AIU	0.4806	0.1056	0.2198
AOU	0.4357	0.1425	0.3269
EIO	0.3740	0.1824	0.4876
EIU	0.3717	0.1254	0.3374
EOU	0.4709	0.1246	0.2645
IOU	0.4061	0.1398	0.3444
AEOU	0.2635	0.1103	0.4185
AEIO	0.3357	0.1025	0.3054
AEIU	0.2552	0.0666	0.2611
EIOU	0.2911	0.1218	0.4182
AIOU	0.2794	0.1107	0.3962
AEIOU	0.2864	0.0918	0.3205
Average	0.3690	0.1263	0.3538

Tab. 3. Experimental results over all formant planes averaged in each vowel polygon for Method 5.

Table 4 is an equivalent to Tab. 2 where average values of observed parameters are listed for all used methods in vowel polygon criterion (over all possible formant planes). In this case, generally higher values, i.e. less uniform, are reached for formant plane-independent stress detection, but significantly useful values are reached for both mean methods (Method 5 and Method 6). These methods are characterized by slightly significant positive dependency, small standard deviation value and very satisfactory R value. The worst values are reached by Method 2 and, obviously, the Method 4 does not work well as in the previous case (see Tab. 2).

Method number	Cross-correlation [-]		$\overline{R}$
	$\overline{\mu}$	$\overline{\sigma}$	[-]
1	0.2953	0.2490	3.4232
2	0.2830	0.2479	6.3388
3	-0.2434	0.2610	1.2928
4	-0.1490	0.3028	3.914
5	0.3690	0.1263	0.3538
6	0.3336	0.1477	0.4360

Tab. 4. Experimental results over all formant planes further averaged over all vowel polygons.

According to the achieved results, the basic usage of vector angle seemed useless for stress detection. Both mean methods reach much higher uniformity of mined results by cross-correlation. This fact can be caused by the event where each subject feels more or less the same stress level as the other caused by higher probability of final exam failure which leads to less self-confidence of each individual speaker and higher differences between normal and stressed speech.

The consistency of all mined R values is shown in Fig. 2, where the worst methods for stress detection are marked as light blue (Method 2) and orange (Method 4). By this observation, it can be set the statement of the unsuitability of using the area difference signum for high stress detection leading to high cross-correlation results variability and insignificant high stress detection. On the other hand, the most uniform cross-correlation results are received for both mean methods (Method 5 - purple and Method 6 - light brown) in plane and shape criteria. Method 1 is represented by green and Method 3 is marked
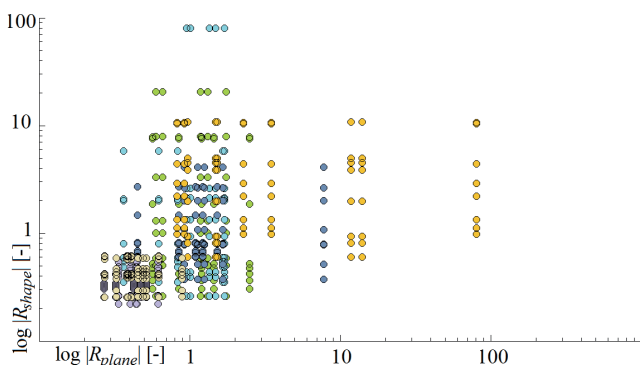

Fig. 2. Plane figuring out reached R for high stress influence. Both axes are in logarithmic scale due to better resolution.

by dark blue color. By this distribution illustrated in Fig. 2, both mean methods have been confirmed as the methods reaching the most consistent results in formant plane and selected shape criterion.

4.2 Efficiency of Vowel Polygons

In this sub-section, the suitability of stress detection will be observed for each possible vowel polygon separately because of not so significant results were achieved only in separated shape or plane criterion. The suitability, exactly the most significant and consist differences, are classified by their current efficiency which is based on results presented in the previous section. Generally, the efficiency of observed parameter x is defined by equation

$$\varepsilon = \frac{\sigma_x^2}{\overline{x}^2} \tag{2}$$

which can be modified for efficiency coefficient E_c as follows

$$E_c = \frac{CCV^2}{\left|R_{plane}\right| \cdot \left|R_{shape}\right|} \tag{3}$$

where CCV is previously calculated cross-correlation value for selected couple of observed parameters for current vowel polygon, R_{plane} is variation coefficient of relevant formant plane and R_{shape} is variation coefficient of relevant shape. Briefly, the value of efficiency coefficients signs the strength of observed couple of parameters for actual vowel polygon referred to statistical values over all relevant planes and shapes. The strength of observed vowel polygon is directly proportional to the E_c value - with increasing E_c the impact of current vowel polygon rises over others similar and relevant.

Experimentally achieved values of efficiency coefficient E_c for each vowel polygon and 6 different observation methods are presented in this section. Due to a big amount of achieved results, following tables list only the 5 top and 5 bottom values. Table 5 contains lists of the best and the worst E_c values. Obviously, significant difference between results of methods using vector angle and others exists. The worst results are achieved by cross-correlation methods of area difference value and its signum with vector angle; vice versa other methods reached more or less similar results, thus the best results are achieved for Method 5 followed by Method 6 and Method 1.

From mined results, the best shapes are AIU, AEU and AIO vowel triangles, supplemented also by formant planes F1F5, F2F5 and F3F5. Results on the bottom of the list are also interesting because, as it can be seen, the big amount of vowel polygons gives null results leading to non-suitability of their usage for stress detection. Generally, the usage of vowel triangles and formant planes containing the formant F5 can be finally evaluated as the best choices for stress detection as well as the usage of both mean methods (Method 5 and Method 6).

Efficiency coefficient E_c [-]						
Rank	Method					
	1	2	3	4	5	6
1.	AIU23 1.920	AEU12 1.661	AIO14 1.373	AEIO35 0.228	AEU23 4.320	IOU13 2.803
2.	AIU14 1.911	AIU12 1.491	AEU23 1.223	AEIU13 0.193	AIO35 4.141	AEU12 2.655
3.	AIU13 1.807	EOU24 1.450	AIO15 1.222	AEO35 0.186	EOU24 3.998	AIO14 2.619
4.	AOU23 1.754	AEU23 1.269	AIO35 1.008	AEIU25 0.184	AIU14 3.833	AIU12 1.910
5.	AOU24 1.735	IOU13 1.234	AIO23 0.982	AIOU35 0.157	AIU13 3.823	AEU25 1.827
…						
156.	AEIU35 0.000	AEIOU13 0.000	EIU23 0.002	AIU14 0.000	AEO24 0.031	EIU12 0.030
157.	AEIOU13 0.000	AIOU13 0.000	IOU23 0.002	AEO24 0.026	AEI13 0.026	AEIOU45 0.030
158.	AEIU12 0.000	AIOU34 0.000	AEO24 0.002	EIU25 0.000	AEI12 0.026	AIO25 0.028
159.	AIOU35 0.000	AIOU12 0.000	AIOU25 0.001	AIU24 0.000	EIO25 0.020	AEIOU34 0.010
160.	AIOU12 0.000	AIOU45 0.000	EOU23 0.000	AEOU13 0.000	AEO12 0.015	AEIU12 0.005

Tab. 5. Ranked list of the five best and five worst vowel polygons.

5. Conclusion

In this paper differences were presented within vowel polygon parameters and their mutual correlation between normal speech and stressed speech taken from the database Exam Stress. The relationships between observed parameter couples were observed by cross-correlation coefficient and statistical parameter called variation coefficient R for investigating the suitability of a reached result over formant planes and vowel shapes. These observations proved that means methods (Method 5 and 6) do not reach the highest cross-correlation values but are the most suitable over all vowel shapes and formant planes.

Furthermore, the appropriateness for possibly stress detection was classified by created efficiency coefficient based on classic efficiency equation for each individual vowel polygon separately. Several statements can be laid by this indicator. Methods 1, 5 and 6 reached the best results, and the worst results were achieved by Method 4 which is characterized by low values of the efficiency coefficient E_c (much lower than for other observed methods).

It was also proved that the lower formant planes contain foremost information about spoken phoneme while information of speaker's state and identity are attenuated. The best vowel shape for stress detection proves to be IOU, AIO, AIU, and AEU vowel triangles as well as AEIU and AEIO vowel tetragons. Obviously, the best formant planes for stress detection are F1F5, F2F5 and F3F5. In conclusion, stress can be possibly uncovered by usage of mentioned vowel shapes and formant planes (leaded to a various number of vowel polygons) by the fifth experimental method. In future, presented research will be applied on other language, e.g. English or German, speech under stress database to compare received results and to observe if presented findings are language-dependent or not.

Acknowledgments

References

[1] LE FEVRE, M., MATHENY, J., KOLT, G. S. Eustress, distress, and interpretation in occupational stress. *Journal on Managerial Psychology,* 2003, vol. 18, no. 7, p. 726–744. DOI: 10.1108/02683940310502412

[2] ALMEIDA, D. M., KESSLER, R. C. Everyday stressors and gender differences in daily distress. *Journal of Personality and Social Psychology,* 1998, vol. 75, no. 3, p. 670–680. DOI: 10.1037/0022-3514.75.3.670

[3] KURNIAWAN, H., MASLOV, A. V., PECHENIZKIY, M. Stress detection from speech and Galvanic Skin Response signals. In *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems.* Porto (Portugal), 2013, p. 209 - 214. DOI: 10.1109/CBMS.2013.6627790

[4] Bt JOHARI, N. A. a., HARIHARAN, M., SAIDATUL, A., YAACOB, S. Assimilate the auditory scale with wavelet packet filters for multistyle classification of speech under stress. In *Proceedings of International Conference on Biomedical Engineering.* Penang (Malaysia), 2012, p. 537–542. DOI: 10.1109/ICoBE.2012.6179076

[5] HANSEN, J. H. L., BOU-GHAZALE, S. Getting started with SUSAS: A Speech Under Simulated and Actual Stress database. In *Proceedings of EUROSPEECH 1997.* Rhodes (Greece), 1997, p. 1743–1746.

[6] GALANIS, D., KARABETSOS, S., KOUTSOMBOGERA, M., PAPAGEORGIU, H. Classification of emotional speech units call in centre interactions. In *Proceedings of the 4th IEEE International Conference on Cognitive Infocommunications.* Budapest (Hungary), 2013, p. 403–406. DOI: 10.1109/CogInfoCom.2013.6719279

[7] YAO, X., JITSUHIRO, T., MIYAJIMA, C., KITAOKA, N., TAKEDA, K. Psychical characteristics of vocal folds during speech under stress. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing.* Kyoto (Japan), 2012, p. 4609–4612. DOI: 10.1109/ICASSP.2012.6288945

[8] HE, L., LECH, M., MADDAGE, M. C., ALLEN, N. Stress detection using spectrograms and sigma-pi neuron units. In *Proceedings of the 5th International Conference on Natural Computation.* Tianjin (China), 2009, p. 260–264. DOI: 10.1109/ICNC.2009.59

[9] LECH, M., HE, L. Stress and emotion recognition using acoustic speech analysis. *Journal on Mental Health Informatics: Studies in Computational Intelligence,* 2014, vol. 491, p. 163–184. DOI: 10.1007/978-3-642-38550-6_9

[10] PRIBIL, J., PRIBILOVA, A. Determination of formant features in Czech and Slovak for GMM emotional speech classifier. *Radioengineering,* 2013, vol. 22, no. 1, p. 52–59.

[11] HASRUL, M. N., HARIHARAN, M., YAACOB, S. Human affective (emotion) behavior analysis using speech signals: A review. In *Proceedings of International Conference on Biomedical Engineering.* Penang (Malaysia), 2012, p. 217–222. DOI: 10.1109/ICoBE.2012.6179008

[12] GIDDENS, C. L., BARRON, K. W., BYRD-CRAVEN, J., CLARK, K. F., WINTER, S. A. Vocal indices of stress: A review. *Journal of Voice,* 2013, vol. 27, no. 3, p. 390.e21–390.e29. DOI: 10.1016/j.jvoice.2012.12.010

[13] PETTA, P., PELACHAUD, C., COWIE, R. *Emotion-Oriented Systems: The Humaine Handbook.* Berlin: Springer, 2011. DOI: 10.1007/978-3-642-15184-2

[14] SIGMUND, M. Statistical analysis of fundamental frequency based features in speech under stress. *Information Technology and Control,* 2013, vol. 42, no. 3, p. 286–291. DOI: 10.5755/j01.itc.42.3.3895

[15] TSE, A. C. Y., WONG, A. W.-K., WHITEHILL, T. L., MA, E. P.-M., MASTERS, R. S. W. Analogy instruction and speech performance under psychological stress. *Journal of Voice,* 2014, vol. 28, no. 2, p. 196–202. DOI: 10.1016/j.jvoice.2013.03.014

[16] DOMBROWSKI, M., McCLEERY, A., STANFORD, G. W., DOCHERTY, N. M. Stress reactivity of emotional and verbal speech content in schizophrenia. *Journal of Nervous & Mental Disease,* 2014, vol. 202, no. 8, p. 608–612. DOI: 10.1097/NMD.0000000000000169

[17] BUCHANAN, T. W., LAURES-GORE, J. S., DUFF, M. C. Acute stress reduces speech fluency. *Journal of Biological Psychology,* 2014, vol. 97, p. 60–66. DOI: 10.1016/j.biopsycho.2014.02.005

[18] BUTTERWORTH, B. Evidence from pauses in speech. *Journal of Language Production,* 1980, vol. 1, p. 155–176.

[19] STANEK, M., SIGMUND, M. Speaker dependent changes in formants based on normalization of vowel triangle. In *Proceedings of the 23th International Conference Radioelektronika.* Pardubice (Czech Republic), 2013, p. 337–341. DOI: 10.1109/RadioElek.2013.6530941

[20] SIGMUND, M. Influence of psychological stress on formant structure of vowels. *Elektronika ir Elektrotechnika,* 2012, vol. 18, no. 10, p. 45–48. DOI: 10.5755/j01.eee.18.10.3059

[21] SIGMUND, M. Introducing the database ExamStress for speech under stress. In *Proceedings of the 7th Nordic Signal Processing Symposium.* Rejkjavik (Iceland), 2006, p. 290–293. DOI:10.1109/NORSIG.2006.275258

[22] STANEK, M. Software for generation and analysis of vowel polygons. In *Proceedings of the 37th International Conference on Telecommunications and Signal Processing.* Berlin (Germany), 2014, p. 424–427.

[23] STANEK, M., POLAK, L. Algorithms for vowel recognition in fluent speech based on formant positions. In *Proceedings of the 36th International Conference on Telecommunications and Signal Processing.* Rome (Italy), 2013, p. 521–525. DOI: 10.1109/TSP.2013.6613987

[24] STANEK, M., SIGMUND, M. Comparison of speaker individuality in triangle area of plane formants spaces. In *Proceedings of the 24th International Conference Radioelektronika.* Bratislava (Slovakia), 2014, p. 1–4. DOI: 10.1109/Radioelek.2014.6828439

[25] RVACHEW, S., MATTOCK, K., POLKA, L., MENARD, L. Development and cross-linguistic variation in the infant vowel space: The case of Canadian English and Canadian French. *Journal of Acoustical Society of America,* 2006, vol. 120, no. 4, p. 1–10. DOI: 10.1121/1.2266460

[26] SHARIFZADEH, H. R., McLOUGHLIN, I. V., RUSSELL, M. J. A comprehensive vowel space for whispered speech. *Journal of Voice,* 2012, vol. 26, p. e49–e56. DOI: 10.1016/j.jvoice.2010.12.002

[27] WHITFIELD, J. A., GOBERMAN, A. M. Articulatory-acoustic vowel space: Application to clear speech in individuals with Parkinson disease. *Journal of Communication Disorders,* 2014, vol. 51, p. 19–28. DOI: 10.1016/j.jcomdis.2014.06.005

[28] MILIVOJEVIC, Z. N., BRODIC, D., MILIVOJEVIC, M. Z. The effects of the active hypoxia to the speech signal inharmonicity. *Radioengineering,* 2014, vol. 23, no. 2, p. 665–670.

[29] DELLWO, V. Rhythm and speech rate: A variation coefficient for deltaC. In *Proceedings of the 38th Linguistic Colloquium on Language and Language Processing.* Piliscaba (Hungary), 2003, p. 231–241.

[30] CHANDAKA, S., CHATTERJEE, A., MUNSHI, S. Support Vector Machines employing cross-correlation for emotional speech recognition. *Measurement,* 2009, vol. 42, no. 4, p. 611–618. DOI: 10.1016/j.measurement.2008.10.005

[31] ADAMI, A. G., HERMANSKY, H. Segmentation of speech for speaker and language recognition. In *EUROSPEECH 2003.* Geneva (Switzerland), 2003, p. 1–4.

[32] WRIGLEY, S. N., BROWN, G. J., WAN, V., RENALS, S. Speech and crosstalk detection in multichannel audio. *IEEE Transactions on Speech and Audio Processing,* 2005, vol. 13, no. 1, p. 84–91. DOI: 10.1109/TSA.2004.838531

[33] DIETRICH, M., ABBOTT, K. V. Psychobiological stress reactivity and personality in persons with high and low stressor-induced extralaryngeal reactivity. *Journal of Speech, Language and Hearing Research,* 2014, vol. 57, no. 6, p. 2076–2089. DOI: 10.1044/2014_JSLHR-S-12-0386

About the Authors ...

Miroslav STANEK received his M.Sc. from Brno University of Technology (BUT) in 2012. Currently, he is a master student at the Dept. of Forensic Engineering of BUT and a doctoral student at the Faculty of Electrical Engineering and Communication at the same university. His research interests are oriented on speech signal processing, exactly on speaker recognition and investigation of current speaker's state.

Milan SIGMUND received his M.Sc. degree in Biomedical Engineering and the Ph.D. degree in Speech Signal Processing, both from Brno University of Technology. Currently, he is a professor at the Dept. of Radio Electronics, Faculty of Electrical Engineering and Communication, Brno University of Technology.