

\$ sciendo

THE USE OF FORMANTS' CORRELATION IN ASSESSING THE SADNESS STATE OF THE SPEAKERS

ΒY

ŞTEFAN-ANDREI GHELŢU*

"Gheorghe Asachi" Technical University of Iaşi, Faculty of Electronics, Telecommunications and Information Technology, Iaşi, Romania

Received: March 31, 2023 Accepted for publication: October 7, 2023

Abstract. The purpose of this study is to further the comparative analysis of the variations in the correlation coefficients of the formants for the Romanian vowels during emotional speech. We created an annotated speech database composed of recordings of speakers pronouncing the following sentences: /vine mama/ (/mother is coming/), /aseară/ (/last night/), /cine a făcut asta/ (/who did this/), first on a neutral tone of voice, and then expressing sadness.

The analysis focuses on the influence that sadness has on the vocal signal. The formants and F0 (pitch) of each vowel were extracted. Statistical analysis techniques were applied in order to verify whether the variation of the correlation coefficients between F0 and F1-F4 presents significant variations or whether it tends to be homogenous under the following conditions: (1) same speaker – same vowel – different sentences; (2) same speaker – same vowel – emotional neutrality vs. sadness; (3) different speakers – same sentence – same vowel.

Keywords: correlation; formants; sadness; Romanian speech database.

© 2022 Ștefan-Andrei Ghelțu

^{*}Corresponding author; *e-mail*: stefan.gheltu@gmail.com

This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

1. Introduction

The aim of this study is to perform a comparative analysis of the correlation coefficient variations for the Romanian vowels during neutral and emotional speech.

According to the Merriam-Webster Dictionary (2002), emotions are "conscious mental reactions subjectively experienced as strong feelings, usually directed toward a specific object and typically accompanied by physiological and behavioural changes in the body". Klaus R. Scherer (2003) proposed a more comprehensive definition, stating that emotions are "relatively brief episodes of synchronized response of all or most organismic subsystems in response to the evaluation of an external or internal event as being of major significance".

Emotions result from the cognitive assessments of the event involved in triggering the emotional reaction. Regarding their effect, emotions have a strong behavioural impact, generating specific changes adapted to context. Therefore, emotions: a) highlight the significance or relevance of a stimulus; b) prepare the body for action by mobilizing the available physiological resources; c) indicate the individual's emotional state, as well as his/her intentions towards the others, thus facilitating communication (Scherer, 2003).

Summarizing, emotion processuality involves: (1) the existence of a triggering stimulus/event; (2) its cognitive assessment; (3) the occurrence of physiological changes (pulse, blood pressure, tremor, changes in the respiratory rate etc.); (4) the manifestation of associated behavioural changes (the fight, attack or freeze reaction) and (5) the evaluation of the perceived state or subjective experience.

Previous literature presented two main directions of research integrating most studies on emotions, providing valid scientific data and allowing the establishment of two explanatory theoretical models: 1) *discrete theories* and 2) *dimensional theories* of emotions. In the former, the first contributions were Darwin's empirical observations on the evolutionary role of emotions and their universal character. Paul Ekman (1999) proposed the model of the six basic, innate and universal emotions: *joy, sadness, anger, fear, surprise* and *disgust*. R. Plutchik (1997) added acceptance and anticipation in the category of fundamental emotions, emphasizing their primary role in the process of mobilizing adaptation resources. On the other hand, dimensional theories explain emotions analysing their modulation on the following three dimensions: (1) valence; (2) dominance; (3) arousal. The authors who support this paradigm analysed a relatively low number of emotions characterized by specific patterns of responses at the physiological level and in the speaker's voice signal.

Regarding the vocal expression of emotion, in order to explain this process, Klaus R. Scherer (2003) adapted a Brunswikian lens model. The model synthesizes the process following three steps: encoding, transmission and decoding. During the first stage, emotion is embedded into the vocal signal, in

the following stage, the transmission of the semantic and emotional content occurs, and the third and final stage consists of decoding and evaluating the information. When the message reaches the receiver, he undertakes a double evaluation process: on the one hand, the informational content is analyzed, and on the other, the recognition of the emotional state is performed.

Within this process, a factor of major relevance is represented by the indicator cues, which allow for the modelling of the vocal signal, taking into account specific patterns of the associated emotion. According to the measurement's degree of objectivity, Klaus R. Scherer (2003) classifies the vocal parameters into: *acoustic-phonetic parameters* (speech rate, F0, intensity, spectral slope, harmonic-to-noise-ratio, formants and formant bandwidth) and *psychoacoustic parameters* (perceived loudness/pitch/rhythm).

In accordance to these parameters, sadness is characterized by: a lower speech rate (fewer syllables per second), lower values compared to other emotions for F0 mean, F0 deviation and F0 range, a lower level of amplitude and harmonics, higher formant bandwidth and a longer pronunciation time for syllables (Scherer, 2003).

For Romanian language, the first studies regarding the influence of emotions at the level of correlation of the vowel formants during dynamic pronunciation have been carried out by (Teodorescu and Feraru, 2007; Teodorescu *et al.*, 2012; Teodorescu *et al.*, 2014; Ghelțu and Teodorescu, 2018; Teodorescu, 2019a, b, c; Ghelțu, 2020).

Previous results showed that pitch tends to be slightly decreased when speakers express sadness and, for male speakers, it is more difficult to differentiate between sadness and neutral tone for the following vowels: /e/, /i/, /a/ and /ă/ (Zbancioc *et al.*, 2011). Furthermore, sadness is more easily distinguishable from joy and anger compared to neutral tone (Teodorescu and Feraru, 2007). Both sadness and neutral tone are low-arousal emotions; in consequence, recognition problems may arise in differentiating the two in terms of vocal signal, despite their valence dissimilarity.

Continuing the analysis of sadness, the correlation coefficients between F0 and the formants presented lower values compared to those of the formant pairs F1-F4 (Ghelţu and Teodorescu, 2018; Ghelţu, 2018; Ghelţu, 2019). Moreover, previous data indicates that for negative valence emotions (sadness and anger), higher order formants are significantly stronger correlated compared to lower order formants. With regards to the neutral tone, a tendency towards homogeneous variation of data was observed: there have been less oscillations between extremely low values and extremely high values as opposed to emotional speech.

The current analysis focuses on the influence of sadness on the vocal signal. We investigated the effect of sadness on the correlation level of the formants in the case of Romanian language vowels. Then, we performed a comparative analysis of the level of correlation between the formants of the vowels when speakers used a neutral tone versus when they expressed sadness. Moreover, we analysed the influence attributed to the stressing of vowels regarding the correlation degree of the formants. Also, we are interested in analysing the effect that the position of the vowel within the word structure has on the degree of correlation. At the same time, we aim at analysing the differences between the correlation coefficients calculated for the fundamental frequency (F0) and the formants F1-F4.

The article is structured in four sections: the first summarizes the contributions of several recent studies, the second presents the method of analysis, the third section illustrates the results obtained and the fourth includes the conclusions of the study, a few limitations and future research directions.

2. Method

As experimentation methodology, for the current study an annotated speech database was created composed of recordings of the speakers pronouncing the short sentences: */mother is coming/* (/vine mama/), */last night/* (/aseară/) and */who did this/* (/cine a făcut asta/). The files had a 24 KHz sampling frequency. The average pronunciation time was 6 ms (0.006 s). The sentences were first pronounced on a neutral tone of voice and then expressing sadness. Twenty speakers (10 male, 10 female) aged 24 to 35, coming from the north-eastern part of Romania, participated in this study; all of them had higher education degrees and did not show any sign of speech pathology. The formants and F0 (pitch) of each vowel were extracted. Segmentation was done manually using the software Praat (Boersma, 2001; Will, 2013). A total number of 120 files were analysed: 10 male + 10 female pronunciation x 3 sentences x 2 emotions.

Statistical analysis techniques were applied in order to verify whether the variation of the correlation coefficients between F0 and F1-F4 presents significant oscillations or whether it tends to be homogenous under the following conditions:

(1) same speaker – same vowel – different sentences;

(2) same speaker – same vowel – emotional neutrality vs. sadness;

(3) different speakers – same sentence – same vowel.

The Pearson correlation coefficient was calculated for couples of formants of the vowels by applying the following formula:

$$C_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 (y_i - \bar{y})^2}}$$
(1)

where:

x represents the value of the first formant in the pair;

y represents the value of the second formant;

n represents the total number of values.

Moreover, we applied multiple linear regression analysis in order to verify whether F1, F2, F3 are effectively predicted by combinations of superior order formants. The data obtained after the correlation analysis and multiple linear regression analysis are reported in the following sections.

3. Results

The preliminary results are in line with the previous studies that found emotion-dependent variations in the local correlation between formants. Tables 1 and 2 present the correlation coefficients computed for F0 and F1-F4 formant pairs, first for the neutral tone and then for sadness. Both for male and female speakers, superior order formants (F3 and F4) present higher correlation values than lower order formants (F1 and F2).

VC	Values of the correlation coefficients for neutral tone - /Aseara/									
Neutral Tone (Female Speakers)						al Tone (Male Spe	eakers)		
	A1	A2	Ε	Ă	A1	A2	Е	Ă		
correl F0,F1	0.126	-0.234	-0.126	0.167	0.002	-0.175	-0.299	-0.210		
correl F0,F2	-0.024	-0.127	0.160	-0.184	0.019	-0.103	-0.158	-0.051		
correl F0,F3	-0.057	-0.268	-0.170	-0.120	-0.119	-0.300	-0.644	-0.287		
correl F0,F4	0.054	-0.427	-0.003	-0.206	0.070	-0.286	-0.540	-0.251		
correl F1,F2	0.568	0.390	0.411	0.448	0.605	0.516	0.068	0.770		
correl F1,F3	0.469	0.285	0.169	0.418	0.581	0.294	0.190	0.715		
correl F1,F4	0.334	0.274	-0.014	0.362	0.330	-0.017	0.083	0.596		
correl F2,F3	0.564	0.526	0.516	0.584	0.721	0.498	0.419	0.755		
correl F2,F4	0.491	0.470	0.488	0.575	0.468	0.288	0.505	0.692		
correl F3,F4	0.660	0.646	0.558	0.615	0.622	0.549	0.505	0.753		

 Table 1

 Values of the correlation coefficients for neutral tone - /Aseară/

Regardless of the speaker's emotional state (neutral tone vs. sadness), correlations between F0 and formants tend to have lower values compared to those calculated for the formant pairs (F1-F4).

For the vowel $|\check{a}|$, when male speakers use neutral tone, the correlation level between the formant pairs (F1-F4) is higher than in the case of female speakers.

Regarding the /i/ vowel, when female speakers express sadness, the correlation values between F1-F4 are lower compared to those associated to the neutral tone. This result is in line with previous data on the phonetic and acoustic particularities of the /i/ vowel, indicating that during its sustained pronunciation, the correlation degree between formant pairs is significantly lower compared to the vowels /a/, /e/, /o/, /u/, /ă/ and /â/.

As for the vowel $/\check{a}/$, both for neutral tone and sadness, the correlation values of F1-F4 were significantly higher for male speakers compared to female speakers. Moreover, there have been no significant differences between the

correlation values of the formants when speakers used neutral tone versus when they expressed sadness.

As shown in Tables 5 and 6, for male speakers, the vowel /u/ has the highest correlation values of the formant pairs. We observed very strong correlation values higher than 0.70, not only for neutral tone, but also for sadness. In the case of female speakers, the same vowel (/u/) in the same sentence and occupying the same position within word structure has significantly lower correlation values. As an example, for sadness, the correlation between the formats of the vowel /u/ is 0.77 (F1-F3; male speakers) and is significantly higher than 0.20 (F1-F3; the corresponding value for female speakers).

Sadness (Female Speakers)					Sac	lness (Ma	ale Speak	ers)
	A1	A2	Е	Ă	A1	A2	Е	Ă
correl F0,F1	-0.081	0.076	0.159	-0.011	0.045	-0.324	-0.181	-0.169
correl F0,F2	-0.235	-0.016	0.085	-0.020	-0.118	0.059	0.283	-0.055
correl F0,F3	-0.222	0.120	0.066	0.012	-0.193	-0.177	-0.273	-0.158
correl F0,F4	-0.076	0.145	-0.109	0.047	-0.112	-0.116	-0.546	-0.274
correl F1,F2	0.314	0.249	0.123	0.706	0.527	0.205	0.171	0.652
correl F1,F3	0.314	0.375	0.012	0.622	0.570	0.239	0.284	0.681
correl F1,F4	0.171	0.257	0.069	0.437	0.399	-0.035	0.168	0.546
correl F2,F3	0.512	0.430	0.654	0.543	0.761	0.665	0.564	0.692
correl F2,F4	0.678	0.306	0.573	0.543	0.574	0.462	-0.057	0.460
correl F3,F4	0.678	0.673	0.682	0.619	0.644	0.610	0.405	0.763

Table 2
Values of the correlation coefficients for sadness - /Aseară/

 Table 3

 Values of the correlation coefficients for neutral tone - /Cine a făcut asta/

Neutral (Female Speakers)									
	Ι	Е	A1	Ă	U	A2	A3		
correl F0,F1	-0.043	0.037	0.058	0.007	0.023	0.030	-0.003		
correl F0,F2	-0.129	-0.034	-0.176	0.101	-0.106	0.088	-0.007		
correl F0,F3	-0.086	0.180	-0.146	0.054	-0.306	-0.029	0.109		
correl F0,F4	0.029	0.376	-0.286	-0.075	-0.037	-0.094	0.169		
correl F1,F2	0.251	0.427	0.422	0.515	0.389	0.207	0.560		
correl F1,F3	0.635	0.624	0.378	0.297	0.309	0.345	0.501		
correl F1,F4	0.625	0.386	0.376	0.253	0.297	0.216	0.429		
correl F2,F3	0.625	0.624	0.476	0.538	0.505	0.508	0.566		
correl F2,F4	0.625	0.386	0.352	0.313	0.406	0.359	0.485		
correl F3,F4	0.509	0.671	0.744	0.499	0.551	0.592	0.629		

Table 4 Values of the correlation coefficients for sadness - /Cine a făcut asta/									
Sadness (Female Speakers)									
	Ι	Е	A1	Ă	U	A2	A3		
correl F0,F1	0.087	0.121	0.191	-0.145	0.121	0.078	0.066		
correl F0,F2	-0.071	-0.035	0.079	0.076	0.227	0.172	-0.034		
correl F0,F3	-0.124	-0.157	0.057	0.318	0.297	0.245	0.008		
correl F0,F4	0.038	-0.123	0.197	0.182	-0.047	0.232	-0.145		
correl F1,F2	0.123	-0.059	0.523	0.536	0.306	0.355	0.585		
correl F1,F3	0.211	0.026	0.317	0.384	0.207	0.348	0.540		
correl F1,F4	0.137	0.068	0.364	0.250	0.252	0.301	0.518		
correl F2,F3	0.632	0.643	0.683	0.654	0.567	0.754	0.675		
correl F2,F4	0.455	0.621	0.551	0.240	0.292	0.614	0.559		
correl F3,F4	0.515	0.656	0.729	0.437	0.475	0.634	0.635		

 Table 5

 Values of the correlation coefficients for neutral tone - /Cine a facut asta/

Neutral (Male Speakers)									
	Ι	Е	A1	Ă	U	A2	A3		
correl F0,F1	0.152	0.172	-0.251	-0.022	-0.061	-0.048	-0.001		
correl F0,F2	0.202	0.177	-0.143	-0.033	-0.048	0.122	0.076		
correl F0,F3	0.124	-0.009	-0.176	-0.068	-0.101	0.135	0.023		
correl F0,F4	0.094	-0.012	-0.191	-0.195	-0.203	-0.027	-0.039		
correl F1,F2	0.397	-0.376	0.670	0.809	0.690	0.694	0.758		
correl F1,F3	0.514	0.033	0.720	0.742	0.729	0.563	0.702		
correl F1,F4	0.561	0.378	0.784	<i>0.783</i>	0. 777	0.569	0.641		
correl F2,F3	0.535	0.599	0.730	0.864	0.812	0.651	0.760		
correl F2,F4	0.459	0.104	0.614	0.767	0.564	0.496	0.714		
correl F3,F4	0.701	0.526	0.771	0.785	0.732	0.720	0.772		

Table 6

Values of the correlation coefficients for sadness - /Cine a facut asta/

Sadness (Male Speakers)									
	I E A1 Ă U A2 A3								
correl F0,F1	0.074	0.146	0.137	0.017	0.411	-0.048	0.242		
correl F0,F2	-0.141	-0.046	-0.025	-0.024	-0.010	-0.077	0.139		
correl F0,F3	-0.078	-0.029	-0.130	-0.127	-0.206	-0.101	0.167		
correl F0,F4	-0.229	-0.021	0.072	0.002	-0.113	-0.053	0.088		
correl F1,F2	0.309	0.066	0.486	0.730	0.772	0.457	0.70 7		
correl F1,F3	0.302	0.253	0.587	0.699	0.779	0.558	0.626		
correl F1,F4	0.265	0.477	0.657	0.694	0.707	0.496	0.610		
correl F2,F3	0.678	0.645	0.692	0.815	0.831	0.683	0.727		
correl F2,F4	0.405	0.309	0.657	0.749	0.734	0.557	0.671		
correl F3,F4	0.661	0.617	0.755	0.809	0.785	0.619	0.770		

Stefan-Andrei Gheltu

The gender differences identified for the /u/ vowel also remain significant when the speakers voice is neutral. This result partially disagrees with our previous data (Ghelţu and Teodorescu, 2018; Ghelţu, 2019; Ghelţu, 2020), which suggested that there were no significant statistical differences between the male and female speakers in terms of the formants level of correlation. But regarding the generalization of our previous conclusion, it must be considered that the analysed data mainly consisted of the dynamic pronunciation of the vowels /i/, /e/, /a/ or their sustained pronunciation, including the vowel /u/.

railes of the correlation coefficients for neutral tone "rithe mama										
Neutral Tone (Female Speakers)					Neut	ral Tone (ral Tone (Male Speakers)			
	Ι	Е	A1	A2	Ι	Е	A1	A2		
correl F0,F1	0.041	0.119	0.216	-0.137	-0.154	-0.498	-0.398	-0.011		
correl F0,F2	0.003	0.147	0.429	-0.305	-0.220	-0.112	0.205	-0.160		
correl F0,F3	0.125	0.035	-0.288	-0.095	-0.266	-0.416	-0.247	-0.181		
correl F0,F4	0.062	0.254	-0.001	-0.225	-0.146	-0.385	-0.482	-0.193		
correl F1,F2	-0.186	0.148	0.508	0.653	0.359	0.154	0.621	0.677		
correl F1,F3	0.154	0.390	0.504	0.664	0.532	0.402	0.596	0.743		
correl F1,F4	0.064	0.340	0.495	0.652	0.342	0.382	0.388	0.609		
correl F2,F3	0.640	0.698	0.526	0.665	0.532	0.402	0.669	0.743		
correl F2,F4	0.421	0.530	0.504	0.536	0.342	0.382	0.514	0.609		
correl F3,F4	0.485	0.772	0.706	0.699	0.396	0.800	0.742	0.793		

 Table 7

 Values of the correlation coefficients for neutral tone - /Vine mama/

Regardless of the speakers' gender, we observed again lower correlation values between F0 and the formants compared to the values calculated for the formant pairs (F1-F4). For the vowel /i/, the correlation level of F1-F4 formant pairs remains low in both cases (neutral tone vs. sadness).

Summarizing, in the sentence /vine mama/, for vowels /e/, /a1/ and /a2/:

(a) the correlation between higher order formants is significantly stronger when speakers used neutral tone compared to when they expressed sadness (male speakers);

(b) the correlation between higher order formants is only marginally significant when speakers used neutral tone compared to when they expressed sadness. (female speakers).

These results sustain as preliminary conclusions that (1) the correlations between formants are emotion-dependent and (2) individuals differ in the correlation level of the formants.

Moreover, regarding the stressed vowel /a/ in the three sentences analyzed, we extracted the formants and applied multiple linear regression analysis considering F1, F2... as dependent variables and superior order formants as independent variables. Thus, we tested to what extent combinations between superior order formants are effective predictors for estimating changes in the data series of the dependent variable. Obtained results are summarized in Tables 9-11.

28

When female speakers express both sadness and neutral tone, we observed a tendency towards an increased power of prediction for the regression model, due to the exclusion of the previous formants (F1, F2, F3).

In the case of male speakers expressing sadness, the highest value for the coefficient of determination (R^2) is the one corresponding to F2. Thus, the higher order formants (F3-F4) predicted 44% of the dependent variable variance, which, in this case, was F2.

When male speakers use neutral tone, there are significantly higher values for the coefficient of determination (R^2) compared to the values obtained for the same sentence, same vowel, neutral tone, but female speakers. Also, for male speakers, neutral tone, predictors F2, F3 and F4 explained 54% of the dependent variable variance, which, in this case, was F1.

When male speakers use neutral tone, there are significantly higher values for the coefficient of determination (\mathbb{R}^2) compared to the values obtained for the same sentence, same vowel, neutral tone, but female speakers. Also, for male speakers, neutral tone, predictors F2, F3 and F4 explained 54% of the dependent variable variance, which, in this case, was F1.

On the other hand, for female speakers, neutral tone, the same predictors explained only 12% of the criterion variable variance, F2. Moreover, when female speakers expressed sadness, the higher order formants (F3-F4) predicted 60% of the criterion variable variance, F2.

Similarly to the previous case (the second /a/ vowel in the sentence /*cine a facut asta*/), for neutral tone, it has been observed that R^2 values tend to be higher for male speakers compared to those calculated for female speakers. Also, F4 significantly increased the power of predicting the variance of the dependent variable F3, except in the case of sadness, for male speakers.

<i>V</i> .	aiues of i	ne correi	or saaness - /v ine mama/					
Sa	Sa	dness (Male Speakers)						
	Ι	Е	A1	A2	Ι	Е	A1	A2
correl F0,F1	0.435	-0.153	-0.153	0.033	-0.017	0.015	0.331	0.027
correl F0,F2	0.060	0.234	0.000	-0.088	0.251	0.117	0.218	0.027
correl F0,F3	0.229	0.237	-0.088	-0.048	0.525	-0.046	0.134	0.121
correl F0,F4	0.173	0.230	-0.090	-0.037	-0.134	0.002	-0.244	0.143
correl F1,F2	0.231	0.249	0.467	0.650	-0.280	0.069	0.672	0.688
correl F1,F3	0.118	0.296	0.563	0.652	-0.095	0.371	0.220	0.545
correl F1,F4	-0.001	0.332	0.484	0.431	0.317	0.377	-0.016	0.533
correl F2,F3	0.597	0.604	0.545	0.788	0.496	0.664	0.408	0.702
correl F2,F4	0.328	0.592	0.401	0.580	0.248	0.169	0.176	0.628
correl F3,F4	0.270	0.490	0.688	0.680	0.402	0.504	0.385	0.697

 Table 8

 Values of the second state of the

		<u> </u>	1/2/							
	Stressed vowel /a2/									
V	М	ale	Female							
D			1 cmute							
	Sadness	Neutral	Sadness	Neutral						
F	Y=619.06+0.05F2+0.1	Y=287.20+0.36F2+0.1	Y=277.24+0.09F2+0.1	Y=286.09+0.15F2+0.02F3+0.						
1	0F3-0.06F4	0F3-0.10F4	3F3-0.00061F4	001F4						
	$R^2 = 0.11$	$R^2 = 0.31$	$R^2 = 0.15$	$R^2 = 0.16$						
	F(3, 546)=23.96	F(3 449) = 6970	F(3 1451)=91.06	F(3,557) = 36,79						
	n<0.001	n<0.001	n<0.001	n<0.001						
	P 01001	P 0.001	P 01001	P 01001						
F	Y=382.25+0.37F3+0.0	Y=551.30+0.35F3+0.0	Y=916.04+0.20F3+0.0	Y=463.15+0.26F3+0.09F4						
2	4F4	01F4	01F4	$R^2 = 0.30$						
	$R^2 = 0.44$	$R^2 = 0.24$	$R^2 = 0.18$	F(2,558)=121.80						
	F(2,547)=221.56	F(2,450)=74.14	F(2,1452)=165.07	p<0.001						
	p<0.001	p<0.001	p<0.001							
F	Y=630,90+0.46F4	Y=973.22+0.40F4	Y=-543.06+0.79F4	Y=908.43+0.43F4						
3	$R^2 = 0.37$	$R^2 = 0.30$	$R^2 = 0.44$	$R^2 = 0.41$						
	F(1,548)=324.79	F(1,451)=194.32	F(1,1453)=1159.24	F(1,559)=291.37						
	P<0.001	p<0.001	p<0.001	p<0.001						

 Table 9

 Linear regression parameters for stressed vowel /a2/ - /Aseară/

Table 10

Linear regression parameters for stressed vowel /a2/ - /Cine a făcut asta/

	Stressed vowel /a2/										
VD		Male	Female								
	Sadness	Neutral	Sadness	Neutral							
<i>F1</i>	Y=-	Y=-324.61+0.32F2-	Y=66.82+0.39F2+0.05	Y=434.50+0.025F2+0.09F3+0.00							
	130.58+0.03F2+	0.005F3+0.14F4	F3-0.03F4	5F4							
	0.13F3+0.08F4	R ² =0.54	R ² =0.28	R ² =0.12							
	R ² =0.35	F(3,610)=245.81	F(3,1084)=146.40	F(3,758)=34.69							
	F(3.564)=102.26	p<0.001	p<0.001	p<0.001							
	p<0.001										
<i>F2</i>	Y=-	Y=179.03+0.42F3+0.03	Y=-	Y=606.722+0.21F3+0.05F4							
	528.44+0.44F3+	F4	117.66+0.43F3+0.09F4	R ² =0.26							
	0.18F4	R ² =0.42	R ² =0.60	F(2,759)=135.79							
	R ² =0.49	F(2,611)=225.86	F(2,1082)=825.61	p<0.001							
	F(2,565)=280.28	p<0.001	p<0.001	-							
	p<0.001										
<i>F3</i>	Y=258.17+0.61	Y=-681.67+0.85F4	Y=-812.90+0.84F4	Y=-219.91+0.71F4							
	F4	R ² =0.50	R ² =0.46	R ² =0.34							
	R ² =0.38	F(1,612)=623.17	F(1,1083)=947.46	F(1,760)=399.91							
	F(1,566)=348.97	p<0.001	p<0.001	p<0.001							
	p<0.001										

	Table 11										
	Linear regression parameters for stressed vowel /a1/ - /Vine mama/										
	Stressed vowel /a1/										
VD	М	lale	Fem	ale							
	Sadness	Neutral	Sadness	Neutral							
<i>F1</i>	Y=451.78+0.35F2-	Y=296.17+0.14F2+0.15	Y=-245.01+0.18F2+	Y=-							
	0.004F3-0.04F4	F3-0.05F4	0.14F3+0.10F4	59.12+0.25F2+0.06F3+							
	R ² =0.47	R ² =0.45	R ² =0.37	0.08+F4							
	F(3,544)=160.81	F(3,585)=161.32	F(3,742)=147.32	R ² =0.35							
	p<0.001	p<0.001	p<0.001	F(3,601)=112.45							
		_	_	p<0.001							
<i>F2</i>	Y=615.62+0.22F3+0.00	Y=-523.02+0.66F3+	Y=497.42+0.28F3+0.001	Y=488.92+0.14F3+0.11							
	1F4	0.03F4	F4	F4							
	R ² =0.16	R ² =0.44	R ² =0.29	R ² =0.31							
	F(2,545)=54.48	F(2,586)=238.62	F(2,743)=157.33	F(2,602)=135.25							
	p<0.001	p<0.001	p<0.001	p<0.001							
F3	Y=588.34+0.51F4	Y=-328.48+0.76F4	Y=-335.27+0.78F4	Y=-247.21+075F4							
	R ² =0.14	R ² =0.55	R ² =0.45	R ² =0.48							
	F(1,546)=92.96	F(1,587)=717.96	F(1,744)=627.33	F(1,603)=573.12							
	p<0.001	p<0.001	p<0.001	p<0.001							

T 11 44

4. Conclusions

Emotions have a significant influence on the level of correlation of vowel formants during individual dynamic speech. Moreover, individuals differ in the correlation level of the formants and some of the vowels exhibit stronger variations of the correlations between formants in emotional speech. For the current participant sample, the correlation values between the formant pairs were only marginally different when speakers used neutral tone compared to when they expressed sadness. This could be a consequence of the fact that the emotion of sadness has a pattern of low arousal and negative valence and is similar to a greater degree to the state of emotional neutrality, as opposed to happiness, anger or fear.

The statistical results allowed the formulation of the following conclusions:

(1) Both for male and female speakers there have been no significant differences between the correlation values of the formants when speakers used neutral tone versus when they expressed sadness;

(2) For sadness, the higher order formants have a significantly higher correlation level compared to the lower order formants (F1 and F2) and compared to the correlation degree between F0 and F1-F4.

(3) For female speakers, both for sadness and for neutral tone, we observed a tendency towards an increased power of prediction for the regression model, due to the exclusion of the previous formants.

(4) For male speakers, the vowel /u/ has the highest correlation values of the formant pairs, both for sadness and neutral tone. On the other hand, for female speakers, the same vowel (/u/) in the same sentence and occupying the same

position within word structure has significantly lower correlation values, even if the voice expressed sadness or neutral tone.

(5) For the vowel/ \check{a} /, when male speakers use neutral tone, the correlation level between the formant pairs (F1-F4) is higher than in the case of female speakers.

(6) For the vowel /i/, the correlation degree between formant pairs is significantly lower compared to vowels /a/, /e/, /o/, /u/, /ă/ and /â/.

We considered the previous major limitations of our contribution on this research direction and in the present study we included a higher number of participants. Also, the gender equality ratio (10 male, 10 female) was respected.

Compared to previous statistical analysis (Ghelțu and Teodorescu, 2018; Ghelțu, 2019; Ghelțu, 2020), the current study showed more significant results (higher power of prediction). The results complement the statistical analysis of the formants correlation of the vowels and contribute to characterizing the acoustic-phonetic pattern of the Romanian vowels. Also, these results show strong correlations and power of predictions in some cases, which may help improve speech coding and compression.

One of the major limitations of the present study remains the insufficient number of speakers and the absence of longitudinal approaches of the changes in the vocal signal of the same speakers. Therefore, the results of the current study must be interpreted within the context of the particular sample used. Larger scale studies, with a more diverse range of participants could point out the extent to which our results could be generalized. The approach of the study targets the area of inferential statistics (data prediction and comprehension), explaining to a greater degree the empirical aspects investigated. Multiple regression analysis allowed the identification of the formants combinations that would best predict the changes of the dependent variable (represented by the formant ranking lower than the last formant considered a predictor).

Regarding the future research directions, we intend to extend the number of subjects and include representative samples of speakers in database, so that the results can be generalised to a larger number of individuals in the target population. Moreover, we plan to do a comparative study on the variations of the correlation level of the vowels of the Romanian language, taking into account the comparison between normal speech and pathological speech.

Acknowledgements. I acknowledge the help of Prof. Horia-Nicolai Teodorescu, who suggested the topic, the research plan and the method, helped derive the analysis results and revised several versions of the article.

REFERENCES

- Boersma P., *Praat, a system for doing phonetics by computer*, Glot International., 2001, vol. 5, no. 9/10, pp. 341-345.
- Dictionary, Merriam-Webster, Merriam-Webster, 2002, https://www.merriamwebster.com/dictionary/emotion?utm_campaign=sd&utm __medium=serp&utm_source=jsonld (accessed 29 Jun 2021).
- Ekman P., *Basic emotions. Handbook of cognition and emotion*, vol. 98, no. 16, 1999, pp. 45-60.
- Feraru S.M., Teodorescu H.N., Zbancioc M.D., SRoL Web-based resources for languages and language technology e-learning, International Journal Computers Communications & Control, Vol. 5, no. 3, pp. 301-313, Sep 2010.
- Ghelţu S., Teodorescu H.N., Correlations and Dependencies of the Formants During Speech in Romanian, 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Iasi, Romania, 2018, pp. 1-6, doi: 10.1109/ECAI.2018.8678995.
- Ghelţu S., *Correlation of the Formants of the Romanian Vowels*, 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Iaşi, Romania, 2018, pp. 1-5, doi: 10.1109/ECAI.2018.8679075.
- Ghelțu S., Correlations of the Vowel Formants in Romanian and Emotional Effects on Formants, 2019 International Conference on Applied Electronics (AE), Pilsen, Czech Republic, 2019, pp. 1-4, doi: 10.23919/AE.2019.8867028.
- Ghelţu S., A Correlational and Predictive Analysis of the Relationship between F0 and the Formants of Romanian Vowels, 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Bucharest, Romania, 2020, pp. 1-6, doi: 10.1109/ECAI50035.2020.9223249.
- Plutchik R.E., Conte H.R. (Eds.), *Circumplex models of personality and emotions*, American Psychological Association, 1997.
- Sherer K.R., Johnstone T., Vocal expression of emotion, In R.J. Davidson, K.R. Sherer, & H.H. Goldsmith, Handbook of Affective Sciences, 2003, Oxford University Press, pp. 433-456.
- Teodorescu H.N., Feraru S.M., A study on Speech with Manifest Emotions, 10th International Conference on Text, Speech and Dialogue (TSD), Pilsen, Czech Republic, September 3-7, 2007, Lecture Notes in Computer Science, Springer Verlag, vol. 4629/2007, ISBN 978-3-540-74627-0, pp. 254-262.
- Teodorescu H.N., Feraru S.M., Zbancioc M.D., Pistol L., Trandabăț D., Sounds of the Romanian Language Corpus, International Journal of Computers Communications & Control vol. 5, no. 3, pp. 301-313, 2012.
- Teodorescu H.N., Feraru S.M., Zbancioc M.D., Trandabăț D., Romanian Speech Database SRoL, 2014.
- Teodorescu H.N., Using local variance, Allan- and Hadamard variances in speech analysis – Pitch analysis, 2019 International Symposium on Signals, Circuits and Systems (ISSCS), Iași, Romania, 2019a, pp. 1-4, doi: 10.1109/ISSCS.2019.8801757.
- Teodorescu H.N., Allan and Hadamard Variance Analysis of Formants during Sustained Pronunciations with Applications, 2019 International Conference on Speech

Technology and Human-Computer Dialogue (SpeD), Timişoara, Romania, 2019b, pp. 1-6, doi: 10.1109/SPED.2019.8906554.

- Teodorescu H.N., *How much are formants correlated and what that means?* International Symposium on Signals, Circuits and Systems (ISSCS), Iași, Romania, 2019c, pp.1-4, doi: 10.1109/ISSCS.2019.8801758.
- Will S., Using Praat for linguistic research, University of Colorado at Boulder Phonetics Lab.2013,

http://savethevowels.org/praat/UsingPraatforLinguisticResearchLatest.pdf (accessed 21 Feb 2018).

Zbancioc M.D., Teodorescu H.N., Feraru S.M., *Statistical characteristics of the formants* of the Romanian vowels in emotional states, 6th Conf. on Speech Technology and Human-Computer Dialogue (SpeD), IEEE, Brasov, Romania, May 18-21, 2011.

UTILIZAREA CORELAȚIEI FORMANȚILOR ÎN EVALUAREA STĂRII DE TRISTEȚE A VORBITORILOR

(Rezumat)

Scopul acestui studiu propune continuarea analizei comparative privind variațiile coeficienților de corelație ai formanților vocalelor limbii române în timpul vorbirii dinamice. Metodologia experimentală a presupus alcătuirea unei baze de date (annotated speech database) compusă din înregistrări audio în care vorbitorii pronunță următoarele trei propoziții: (1) /vine mama/, (2) /aseară/ și (3) /cine a făcut asta/, mai întâi pe un ton neutru, apoi exprimând ca stare emoțională, tristețea.

Analiza urmărește investigarea modificărilor pe care emoția tristețe le produce la nivelul semnalului vocal. După prelucrarea înregistrărilor, pentru fiecare vocală în parte s-au extras frecvența fundamentală (F0) și formanții (F1-F4). Ulterior, s-au aplicat tehnici de analiză statistică pentru a verifica în ce măsură variațiile coeficienților de corelație calculați între F0 și F1-F4 prezintă variații semnificative sau din contra, tind să fie omogene în următoarele condiții:

(1) același vorbitor - aceeași vocală - propoziții diferite;

- (2) același vorbitor aceeași vocală neutralitate emoțională vs. tristețe;
- (3) vorbitori diferiți aceeași propoziție aceeași vocală.