Classification in Gnathophonics – Preliminary Results

H.N. Teodorescu[1,2], M. Feraru[1,2]
[1] Institute for Computer Science of the Romanian Academy
[2]Technical University of Iaşi
B-dul Carol I nr. 8, 6600 Iaşi
ROMÂNIA
hteodor@etc.tuiasi.ro

Abstract—**We present several preliminary results related to the automatic classification of gnatho phonic sounds. The classification is based on the pseudo-formant concept, and on the pseudo-distance concept. A specific pseudo-distance has been introduced by the first author in relation to this task and to sensorial mechanisms in general. The classification is performed using tools provided by the Mathematica™ package.**

***Index Terms*—Euclidian distance, features space, gnathophonic sounds, pseudo-formants**

I. INTRODUCTION

GNATHOSONICS is a small discipline at the border between dentistry and sound analysis, consisting in a set of investigation methods and techniques aimed to determine the state of the teeth, jaw, and temporo-mandibular articulation, in view of diagnostic. The basic method has been pioneered by Watt [1] and others [2] in the 1970s and 1980s, with the first automated apparatus determining the time and number of occlusal sounds reported in [3]. Gnathophonics is a term proposed by the first author for a set of methods and techniques proposed and experimented first in the 1980s by the first author, in collaboration with V. Burlui and C. Morarasu [4]. It should be emphasized that other authors have also investigated the influence of the dentition state on the voice, for example [5]. Gnathophonics has been developed as a paired discipline, at the border of dentistry, language, and voice analysis, in view of determining, based on the voice analysis, of the state of the gnathic system, moreover of analyzing the influence of the state of the gnathic system on voice production. Gnathophonics involves methods and tools pertaining to acoustics, phonetics, technical voice analysis, and dentistry. Details on gnathosonics and gnathophonics, as well as related results are exemplified in [6]-[12].

In this paper, we report on further improvements of the methods, specifically on the attempt to extract features in the space domain in view of classification of imperfectly spelled fricative consonants in case of dental pathologies.

This paper is structured on four section; fist we recall what gnathophonics is, what are the parameters characteristics of gnathophonics and, we exemplify by images the gnathophonic sounds. In the second section we discuss the features space and the Euclidian distances. In the third part we present the results obtained from our analysis. The last section of the paper contains the discussions and conclusions regarding our analysis.

II. METHOD

We recall for the readers unfamiliar with signal processing concepts that the spectrum of a sound represents the “image” in the frequency domain of the vocal signal. The spectrum of vowels presents maxima of the envelope in the frequency domain, around specific frequencies, maxima that are named formants. The fundamental frequency, named pitch, is the resonant frequency of the vocal tract.

It is known that deficiencies in frontal dentition produce alterations of the spelled labial and frontal fricatives, among others. We recall that the consonants are classified by several criteria, such as:

- the articulation point: labial (*b, v, p, f*), dental (*d, z, t, l*), palatal (*n, j, g*);
- the articulation manner: plosive (*b, p, d, t*), fricative (*v, f, z, s, j, ş*), nasal (*m, n*);
- the vibration of the vocal tracts: resonant (*b, d, g*) and mute (*p, t, c, k*).

The first author has developed a simple test, based on a selection of a set of words that are similar, but include fricatives that are often confused. These words have been spelled by several subjects with normal and deficient dentitions and recorded as explained in [8]-[10], [12]. The words are: vata / fata; var / far; vuiet (spelling vvvvvuiet) / vuiet (normal spelling, vuiet) / fui /vaiet (vvvvvaiet) / vaiet / faieton / vecin / fecior / vânt (spelling vvvvvânt) / vânt / fân / vvvvvvvine, vvvine, vvvine / vine / fine / vehement / ferment / vierme / fierbe / vâjâit / vvvvvvvâjjjjjâit / vvvvâjjjjjâie / ffffâşşşşşşâie / ffffâşşşşşâit / fâşâit / sâsâit / sssssâsssssâie / gâjâit / zâzâie / bââzzzzââââie / bâzâie.

The deficiency of one or several teeth causes important modification in the consonants spectrum for the consonants *s, ş, f, v, j*, and *z*.

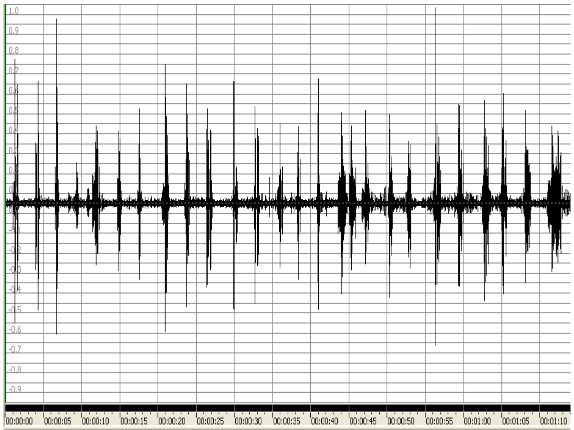

Fig. 1. Example of gnathophonic recordings

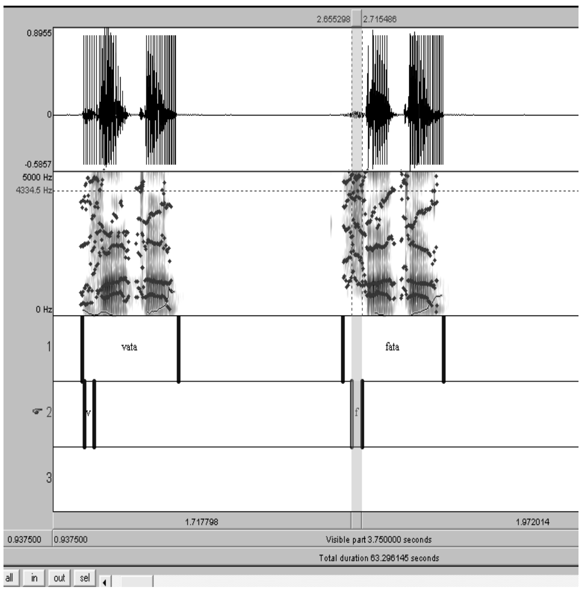

Fig. 2. Example of annotation, spellings: vata – fata, subject # 280269m

The recordings were manually annotated by the second author and we computed the values of the pseudo-formants. Here, we use the term "pseudo" because a sounds like "*v*" or "*f*" do not have true formants; yet, there are emphasized components in the spectra of such sounds, component that look similar to formants on the sonogram. The pseudo-formants were detected using the Praat^{TM} software. In figures 1 and 2, we exemplify gnathophonic recordings and corresponding annotations.

III. THE FEATURES SPACE AND THE EUCLIDIAN DISTANCE

The preliminary step of classification is to choose the features space of the elements for classification. There are no rules for this. We can establish some semi-empiric rules, some on experience, and some on deduction, which can guide us to choose the characteristics. Thus rules are:

1. The features must be specific; a parameter which has the same value no mater if is the "object" of classification is uncharacteristic.
2. The parameters must be independent, otherwise their number can decrease.
3. The number of parameters must be adequate for distinguish all the classes.
4. The number of the parameters must be small to make easier the classifier.
5. The parameters must have a clear semantic content for the experts; in other words to be easy to interpret.

In general, we can choose many parameter sets for a classification problem, all of these sets satisfying the previously mentioned requirements. For example, for the gnathophonic classification, it may be preferable to use mixed parameters, namely in the frequency domain, moreover in the time domain. Features that can be used in gnathophonic classification may include values of the formants, vowel and consonant energy compared to average energy in the specified word, vowel and consonant duration etc.

We used the discrimination rely on the fricative spelling (*s, f, ş*) compared with the close sounds (*s/f, s/ş, ş/j, f/v*) in similar words, in which the paired consonants are in the same position. For example, the first consonant, in an accentuated syllable, followed by the same vowel or sonorous group (fata /vata, fui / vuiet, etc). The chosen parameters are:

- the presence / the absence of pseudo-formants, detected by Praat^{TM};
- the values of pseudo-formants;
- the average energy, on the central part of the respectively sound. In the next figure, we observe the pseudo-formants, the fundamental frequency, and the boundaries between words.

The distance between two pronunciations v1 and v2, for the same vowel, with the formants vector, $\bar{f}_1$ respective $\bar{f}_2$ is computed following the equation:

$$\bar{f}_i = \left(f_{0i}, f_{1i}, f_{2i}, f_{3i}, f_{4i}\right), \ i = 1,2,$$

as

$$\Delta\left(\bar{f}_1, \bar{f}_2\right) = \sqrt{\sum_{k=0}^{4}\left(f_{k1} - f_{k2}\right)^2} \qquad (1)$$

The algorithm automatically selects the classes so that the distance from the vowel to the class must be minimum. The number of classes can be automatically determined or can be pre-established. We analyzed the pairs of pronunciations of the same vowel (in the same word and sentence) for four emotional states. Because the speakers are both women and men, we suppose that the best classification is based on four classes, $(F, E_1), (M, E_1), (F, E_2), (M, E_2)$, where F - women, M – men, E_1, E_2 two emotional states.

The score of classification for pair (sadness, happiness), using four classes for the vowels *i, e,* a_1, a_2, from the sentence "Vine mama". The results are summarized in Table 1.

The pronouncing quality is due to the discrimination between the spellings of some different fricative consonants, but which are closer like characteristics, for example "*f*" and "*v*".

We define the spectral distance between two sounds, one with the strictly consonants characteristics (without fundamental frequency and formants) being the distance between the values of the spectral amplitude to a selected frequency set. For example, using Praat™, the selected frequencies are the focus frequency of the formants (false or true).

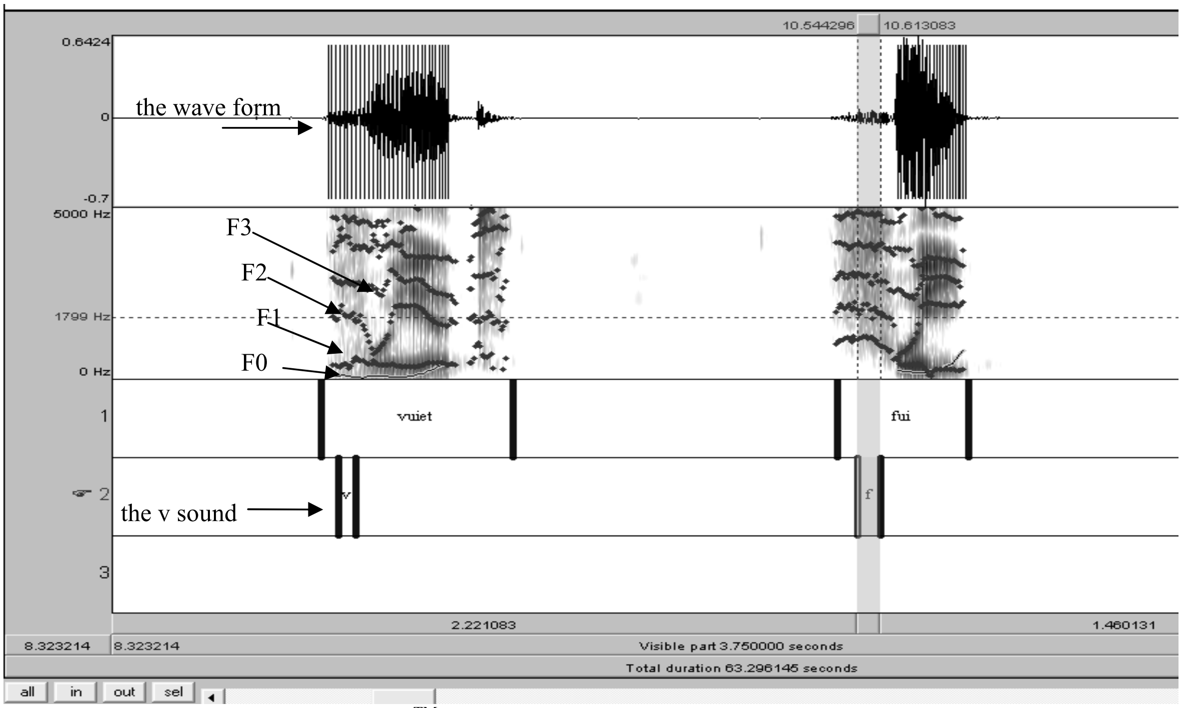

Fig. 3. The processing results using Praat™, spelling vuiet-fui, Subject # 280269m. The acoustic (phoneme), syllable, and word annotations are shown.

Because the spectral separation for sounds is performed by the auditory organ on a logarithmic frequency scale, we use the distance proposed in [11], definition that is based on logarithms of the frequencies of the formants:

$$d(a,b)=\sum_{k=0}^{4} d_k = \sum_{k=0}^{4} \frac{\ln|F_{ka}-F_{kb}|}{\ln(\frac{F_{ka}+F_{kb}}{2})}. \qquad (2)$$

Here, F_{ka} represent the *k*-th formant for sound denoted by "a". In the case when one sound does not have a detected formant, we assign for that formant the elementary distance (the logarithm ratio), being one.

There is a technical justification for using the logarithms, because the formants values are large, using the effective values leads to a majority weight related with higher formants, dominate in frequency and variability.

The graphics of the distance function, according with our definition and with the definition of the Euclidian norm are presented in the next figure, for a space with double variable.

The specific values for the elementary distance are under 1, typical between 0.7 and 0.9. The values under 0.5 for the elementary distance show a small discrimination on the frequency values of the pseudo-formant.

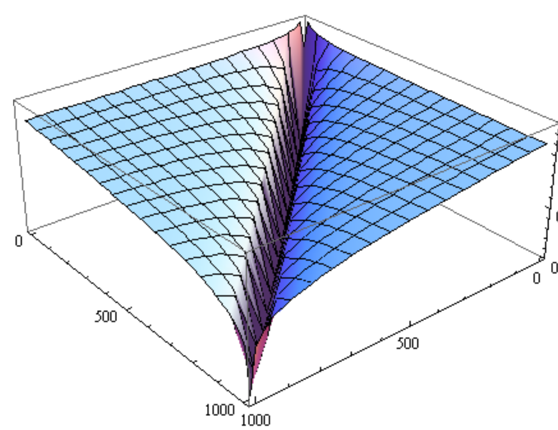

Fig. 4. The graphic of the distance function (2), based on the normalized logarithm of the absolute difference between the pseudo-formants

TABLE II
EXAMPLE OF PSEUDO-FORMANT CARACTHERISTICS OF THE "F" AND "V" CONSONANTS, IN SIMILAR WORDS/ SYLLABE

Gnathophonic recording; subject # 7777m; eccentric mal-relation		
Vată	**Fată**	
147	Undefined	1

841	1433	0.907244
2164	2439	0.725558
3351	3204	0.616496
		3.249298
Gnathophonic recording; subject # 17543m; young person		
Vată	**Fată**	
undefined	undefined	1
895	865	0.501657
2407	1831	0.82992
3437	2928	0.772737
		2.104314

TABLE III
EXAMPLE OF PSEUDO-FORMANT CARACTHERISTICS OF THE "F" AND "V" CONSONANTS, IN SIMILAR WORDS/ SYLLABE

Gnathophonic recording; subject # 7777m; eccentric mal-relation		
Vuiet	**Fui**	
101	143	0.778029
298	1345	1.036142
1629	2225	0.844854
3077	3134	0.502809
		3.161834
Gnathophonic recording; subject # 17543m; young person		
Vuiet	**Fui**	
104	Undefined	1
325	1289	1.026559
1150	1941	0.908784
2899	3056	0.632122
		3.567466

A good pronunciation supposes an adequate distance between the closer sounds, like "*f*" and "*v*". The small distance represents the deficiency of the pronunciation. We exemplify for three people, two without particular problems of the dentition state, and one with multiple problems.

The situation of the person [8] is as following, according to the doctor-made diagnostic:

- Subject # 7777m – 63 years old, multiple cements, multiple crown, advance usage dentition, sub-dimensional maxillary floor, eccentric mal-relation;
- Subject # 17543 – 30 years old, three fillings (obturations) (maxillary), four obturations (on the mandible);
- Subject # 12457m - 30 years old, one filling (maxillary), one filling (mandible).

IV. RESULTS AND DISCUSSIONS

The exemplified results and the analysis of other results reveal that the discrimination between "*f*" and "*v*", measured through the distance between the consonants is significantly reduced, in different words which include the respective consonants, to the older person (see Tables above) and to persons with dentition deficiencies, mainly in the frontal dentition. The use of the distance (2) was found to improve the discrimination between correct spellings and imperfect spellings of the consonants, in the formants / frequency domain, as well as the related automatic detection of the pathological signs for the stomato-gnathic system.

We suggest that further improvements of the diagnostic might be obtained using sets of rules and fuzzy logic. In figure 5, we suggest a "high-pass" membership function for the degree of the discrimination.

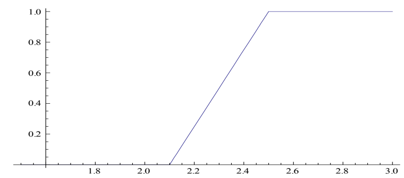

Fig .5. The degree of discrimination – sketch of the membership function

The degree of discrimination is (d-2.1)/(2.5-2.1), if $d \in (2.1, 2.5)$, and 1 for d>2.5 (fuzzification), where *d* is the distance between similar consonants, like *f* and *v*. Notice that this is just an initial proposal, not validated by tests. It is provided for exemplification purposes only.

In Table IV, we sketch some rules that we empirically and tentatively derived for automated diagnostic in gnathophony; these rules are also tentative only and not validated by extensive tests.

TABLE IV
RULES BASED ON THE PSEUDO-FORMANTS CARACTHERISTICS AND ON THE DISTANCE VALUES BETWEEN CONSONANTS WITH SIMILAR SPECTRUM

Words	Characteristics	The conclusion regarding the degree of discrimination	The conclusion regarding the pronunciation
vată / fată	Exist pseudo-fundamental frequency for *v*, but is not to *f*	discriminative consonants	correct
	no pseudo-fundamental frequency to none AND the determined distance on pseudo-formants >2.1	discriminative consonants degree of discrimination = (d-2.1)/(2.5-2.1), if d is between 2.1 and 2.5, respectively 1 for d>2.5 (fuzzification)	correct, with respectively degree of discrimination
	other cases	non-discriminative consonants	pathological pronunciation
vuiet / fui	Exist pseudo-fundamental frequency for *v*, but not for *f*	discriminative consonants	correct
	exist pseudo-fundamental frequency for both *AND* the determined distance on pseudo-formants >2.1	discriminative consonants degree of discrimination = (d-2.1)/(2.5-2.1), if d is between 2.1 and 2.5, respectively 1 for d>2.5 (fuzzification)	critical pronunciation *f*

	other cases	non-discriminative consonants	pathological pronunciation
vecin/ fecior	Exist pseudo-fundamental frequency for *v*, but not for *f*	discriminative consonants	
	There is pseudo-fundamental frequency for both *AND* the determined distance on pseudo-formants >2.1	discriminative consonants degree of discrimination = (d-2.1)/(2.5-2.1), if d is between 2.1 and 2.5, respectively 1 for d>2.5 (fuzzification)	
Etc.			

We choose the characteristics vectors (the total distance for pseudo-formants, without fundamental frequency, for the four words), for five people [5], and we obtains:

Subject 7777m
2.2492, 2.3838, 2.1962, 2.0628

Subject 280269m
2.0816, 2.2258, 2.2025, 2.5925

Subject 17543m
2.1043, 2.5674, 2.7782, 2.7640

Subject 26653m
2.4487, 2.3943, 2.2565, 2.3950

Subject 12457
2.7227, 2.4900, 2.7004, 2.4080

This vector was used in Euclidian classifier with the Mathematica™ software. We obtain the next results:

a. Classification with two classes
FindClusters {2.2492,2.3838,2.1962,2.0628},
{2.0816,2.2258,2.2025,2.5925},
{2.1043,2.5674,2.7782,2.7640},
{2.4487,2.3943,2.2565,2.3950},
{2.7227,2.4900,2.7004,2.4080}

Results:
{{{2.2493,2.3838,2.1962,2.0628},{2.0816,2.2258,
2.2025,2.5925},{2.4487,2.3943,2.2565,2.3950},
{2.7228,2.4900,2.7004,2.4080}},
{{2.1043,2.5674,2.7782,2.7640}}}

Class A:
{2.2493,2.3838,2.1962,2.0628},{2.0816,2.2258,
2.2025,2.5925},
{2.4487,2.3943,2.2565,2.3950},{2.7228,2.4900,
2.7004,2.4080}

Class B:
{{2.1043,2.5674,2.7782,2.7640}

b. Classification with three classes
FindClusters
{2.2492,2.3838,2.19622,2.0628},
{2.0816,2.2258,2.2025,2.5925},
{2.1043,2.5674,2.7782,2.7640},
{2.4487,2.39432,2.2565,2.3950},
{2.7227,2.4900,2.7004,2.4080}

Result:
{{{2.2493,2.3838,2.1962,2.0628},{2.4487,2.3943,
2.2565,2.3950}},{{2.0816,2.2258,2.2025,2.5925},
{2.1043,2.5674,2.7782,2.7640}},{{2.7228,2.4900,
2.7004,2.4080}}}

Class A:
{2.2493,2.3838,2.1962,2.0628},{2.4487, 2.3943,
2.2565,2.3950}

Class B:
{2.0816, 2.2258, 2.2025, 2.5925},{2.1043,2.5674,
2.7782, 2.7640}

Class C:
{2.7228,2.4900,2.7004,2.4080}

Above, the numerical values are truncated to four decimals; the numerical results produced by the software have nine decimals.

V. Conclusions

We have investigated the possibility to automatically classify vowels and especially fricative consonants uttered by healthy subjects and by subjects with imperfect and pathologic dentitions, in view of automatic diagnostic in dentistry.

In the case of gnatho-phonic sounds, the spectral representation allows to choose the specific characteristics (features) related to formants and pseudo-formants. Based on these features and using a specific distance function defined by means of normalized ratio of logarithms, as well as degrees of discrimination, some semi-empirical rules have been tentatively proposed in view of achieving an automated diagnostic, at least for a few pathologies.

While the results reported are preliminary, we believe the paper lays down an appropriate set of theoretical and methodological tools for classification and improved diagnostic in gnathophonics.

Acknowledgment

Research partly performed in relation to the Romanian Academy “priority research” theme “Cognitive Systems” and to a CEEX grant nr 46/2005. The theoretical analysis regarding various distances and pseudo-distances has been performed by the first author mainly with the support of the Grant CEEX Viasan /2005.

REFERENCES

[1] Watt, DM, Gnathosonic Diagnosis and Occlusal Dynamics (Accompanied by Cassette Entitled Occlusal and Temporomandibular Joint Sounds"). Praeger Publishers: East Sussex, 1981

[2] Y. Wakabayashi, D.M. Watt, "Gnathosonic monitoring of patients with gnathic dysfunctions". J Oral Rehabil. 1978, pp. 29-34.

[3] Teodorescu HN, Burlui V, Leca PD., Gnathosonic analyser. Med Biol Eng Comput. 1988 Jul; 26(4):428-31.

[4] Burlui V., Teodorescu H.N., Moraraşu C.S.: La fonction phonatoire chez l'edenté total. Analyse en frequence. *Les Cahiers de Prothese* (France), No. 88, Decembre 1994, pp. 63-68 1994

[5] W. Hedzelek, T. Hornowski, "The analysis of frequency of occlusal sounds in patients with periodontal diseases and gnathic dysfunction". J Oral Rehabil. 1998 pp. 139-45.

[6] H.N. Teodorescu, Gnathophonetics – A New Discipline Analyzing Relations between Speech and the Stomato-Gnathic System. Simpozionul naţional "Bazele performanţei şi inventică" organizat în cadrul "Zilelor Academice Ieşene". 9 Septembrie 2006

[7] H.N. Teodorescu, Occlusal Sound Analysis Revisited. Proc. 3rd Int. Conf. Advances in Medical, Signal and Information Processing, MEDSIP 2006, IET, Glasgow, UK

[8] HN Teodorescu et al., Sunetele Limbii Române, Annotated and Documented Language Corpus for the Romanian Language, http://www.etc.tuiasi.ro/sibm/romanian_spoken_language/ro/sunete_gnatofonice_gnatosonice.htm

[9] H.N. Teodorescu, M. Feraru, Micro-corpus de sunete gnatosonice şi gnatofonice, In Pistol, Cristea, Tufiş (eds.) Resurse lingvistice şi instrumente pentru prelucrarea limbii române. Editura Universităţii "Al. I. Cuza" Iaşi, 2007, pp. 21-30, ISBN 978-973-703-297-3

[10] H.N. Teodorescu, M. Feraru, Corpus de Sunete Gnatosonice, CD-Volume, Ed. Performantica, ISBN 9789737304483, Iaşi, România 2007

[11] Horia-Nicolai L. Teodorescu, Hybrid- and Pseudo-Distances in Pattern Recognition – Medical Applications. Submitted, WSEAS Conf. on Computers, Heraklion, Greece, July 23-25, 2008

[12] HN Teodorescu, Monica Feraru, Diana Trandabăţ, Studies on the Prosody of the Romanian Language: The Emotional Prosody and the Prosody of Double-Subject Sentences. SPED 2007 Iaşi. C. Burileanu, HN Teodorescu (Eds.), Advances In Spoken Language Technology, Ed. Academiei Române, Buc. 2007, pp. 169-180

Note on contributors. The authors' contributions to the research are as follows: The first author has proposed the topic, planned the research, contributed the concepts, developed the methodology, developed the new distances concepts, performed the classification using Mathematica tools, proposed rules, and wrote the paper. The second author has gathered data, annotated voice files, and used PRAAT to extract pseudo-formants, moreover contributed some pictures and some (about 30%) of the translation and editorial matters.