

# A Study on Speech with Manifest Emotions

Horia-Nicolai TEODORESCU \*, \*\*

Monica FERARU \*\*

\* Institute for Computer Science, Romanian Academy

\*\* Technical University of Iasi, Romania

TSD 2007

# Overview

- Introduction
  - Existing approaches
  - Methodology, database, and analysis tools
  - Results of the analysis of speech with manifest emotions
  - Discussions and conclusions
- 

# Introduction

What is the prosody?

(Merriam - Webster, H.N. Teodorescu)

- the rhythm and intonation aspects of a language
- a communication manner which includes the attitude, the emotions

What is the emotion?

- A motivation-related answer adapted to the social environment.

# What researchers said about emotions?

- Emotions characterized a progressive smooth transitions (Schlosberg, 1954).
- For understanding emotional vocal communication, it is necessary a combination of production and perception studies of emotions (Scherer, 1991).
- Emotions are: psychology state, a physiology state, a cognitive process? (V. Auberger, Prosodies et emotion)

# Speculative hypotheses (H.N. Teodorescu)

- The same emotion is represented differently in presence of different interlocutors, depending on the relationship the speaker has with the interlocutor.
- The set of characteristics includes duration of phonemes, duration of pauses, pitch trajectory, including the first formants, moreover the higher formants, the subtle mixture of linear and nonlinear processes of speech generation.

# Emotional speech databases

- Recordings with emotions produced by the “normal” speakers or by actors.
- The sentence are selected from news, papers or depending of the goal of the emotional database.
- The recordings required technical and documentation protocols.
- The database are validated and after the analysis can be start.

# The Greek emotional database

- The study was oriented towards the evaluation of the simulated emotional states by free answers (86.9%) and false answers (89.6%).
- The goal of this research was to improve the naturalness of synthesized voice.
- The recordings were made in three different contexts:
  - in order to reflect the reaction of the speaker to a concrete stimulus (authentic emotion);
  - preparing the environment in order to help psychologically the speaker to simulate the indicated emotion;
  - simulating the emotions only by imagining a context.

# The German emotional database

- The emotions are: anger, happiness, fear, sadness, disgust, boredom and neutral tone. The validation commission recognized 80% of the simulated emotional states.
- It contains:
  - results of the perception tests
  - results of measuring the fundamental frequency, the energy, the duration, the intensity and the rhythm.

# The Danish emotional database

- The emotions are happiness, surprise, sadness, anger and neutral tone. They were correctly recognized in a proportion of 67%.
- The database contains information about the profile of the speaker and the questionnaire.
- The happiness state was mostly confused with surprise and the sadness state was confused with the neutral tone.

# The Spanish emotional database

- The emotions are: happiness, desire, fear, fury, surprise, sadness and disgust.
- The goal was to describe a useful methodology in the validation of the simulated emotional states. The results are useful in generating synthesized speech with emotion.
- The analyzed parameters were the fundamental frequency trajectory, the time, the rhythm and the energy. They represented graphically the wave form and the pitch contour.

# The methodology

- **Recording protocol**
- **The emotional speech database**
- **Processing tools used in the analysis**
  - Praat
  - Klatt analyzer
  - GoldWave
  - Wasp

# Recording protocol

- The speaker signed an informed consent in accordance with to the Protection of Human Subjects Protocol to the U.S. Food and Drug Administration and with Ethical Principles of the Acoustical Society of America for Research Involving Human Subjects.
- The database contains two types of protocols, namely the recording technical protocol and the recording documentation protocol.

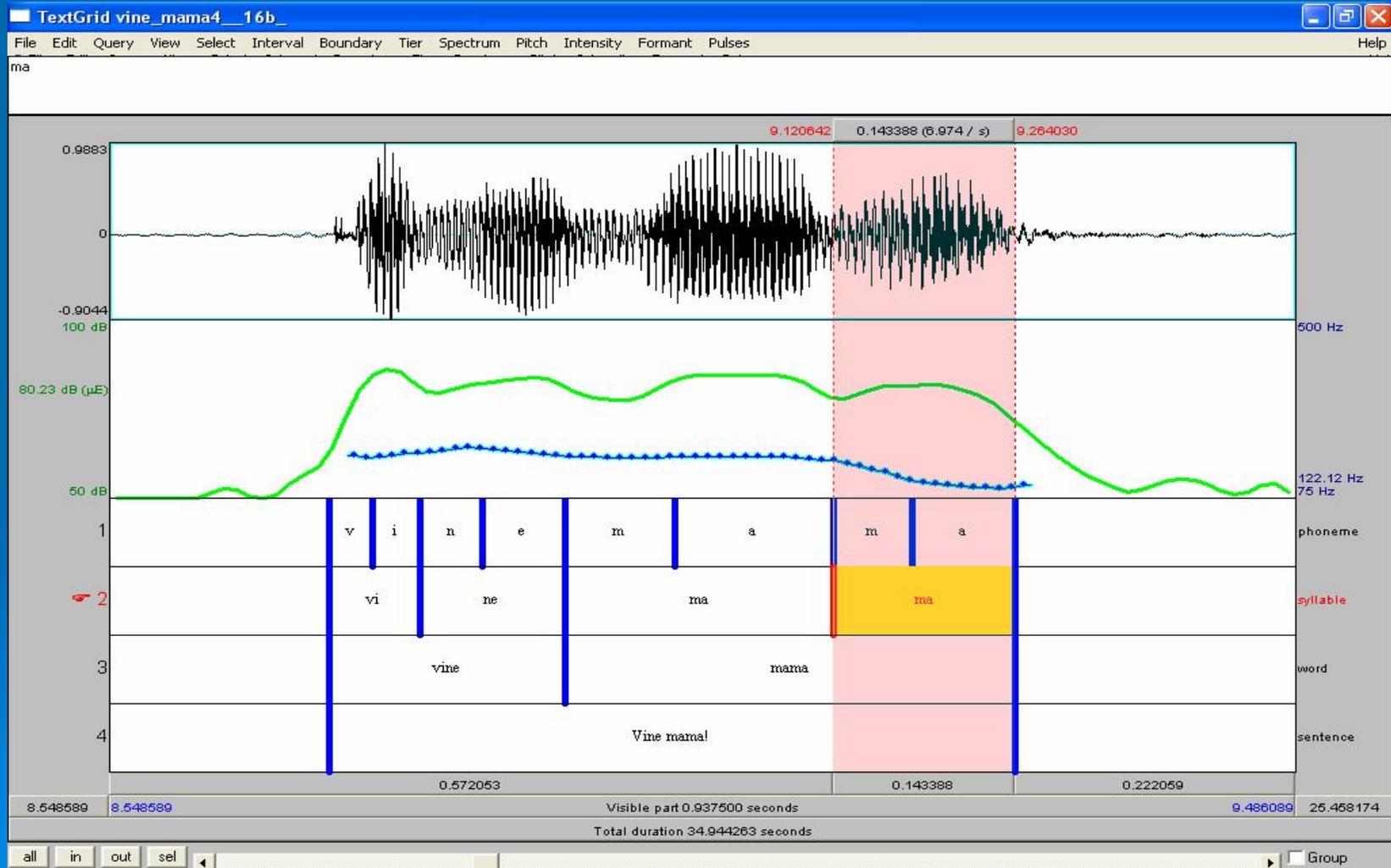
# The emotional speech database

- short sentences ; sadness, happiness, anger and neutral tone.

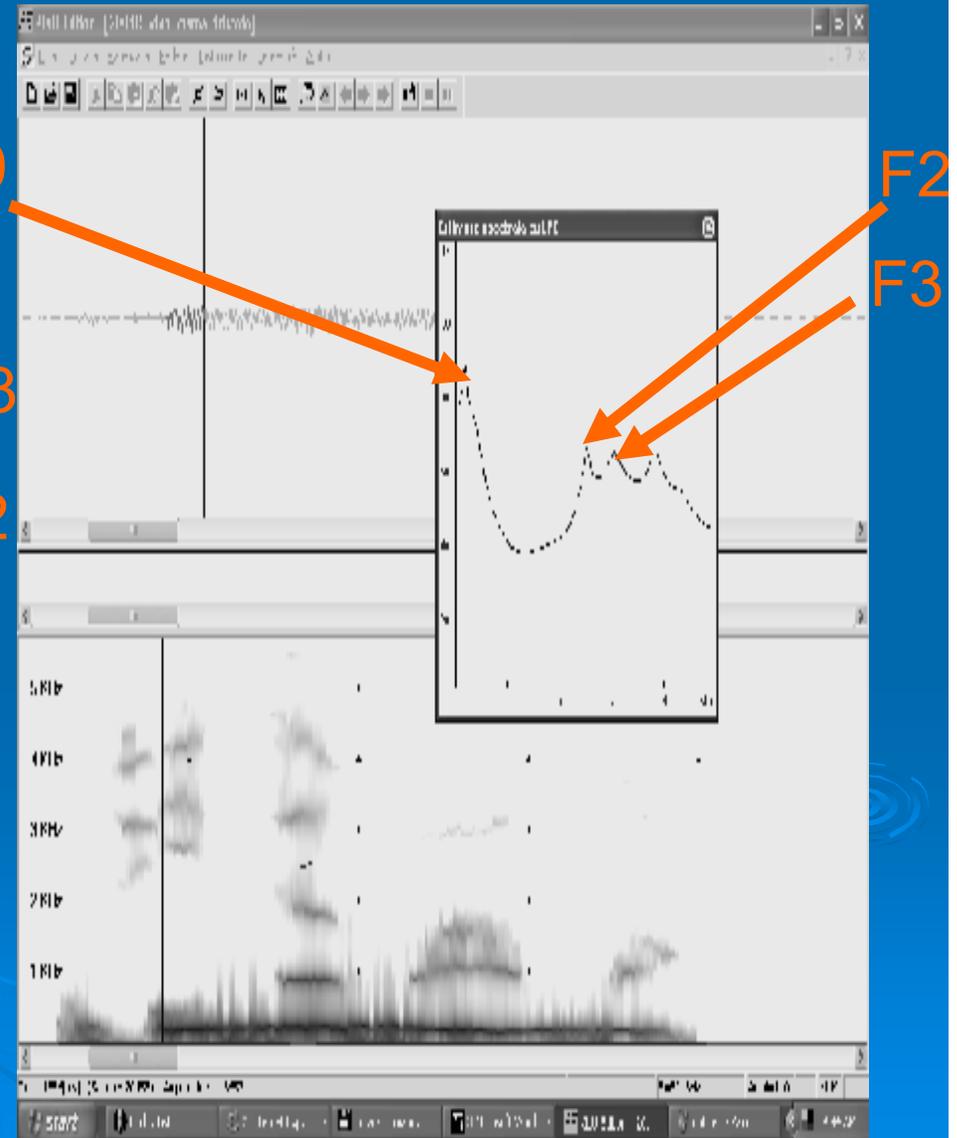
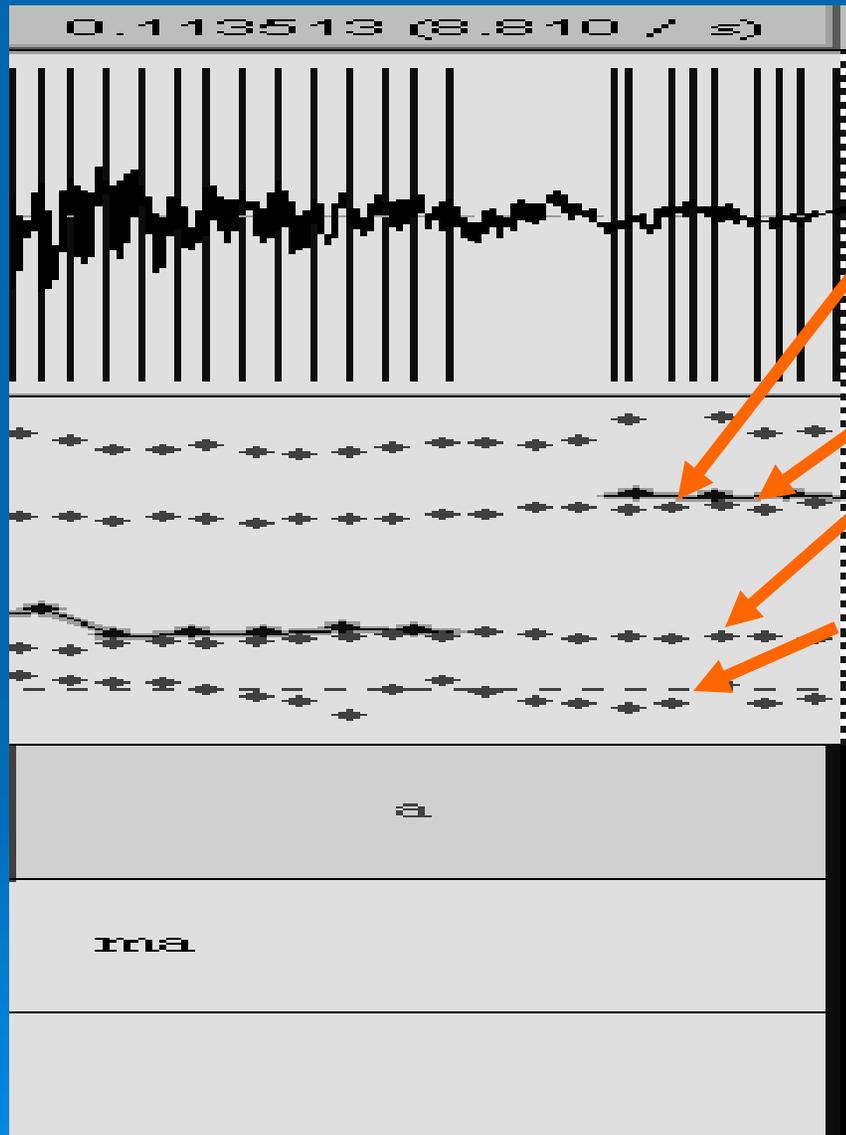
The sentences are:

1. *Vine mama.* (Mother is coming)
  2. *Cine a facut asta.* (Who did that?)
  3. *Ai venit iar la mine.* (You came back to me)
  4. *Aseara.* (Yesterday evening).
- The emotion confusion matrix identified by more than 67% listeners
  - The sentences were annotated using the Praat at several levels: phoneme, syllable, word and sentence.

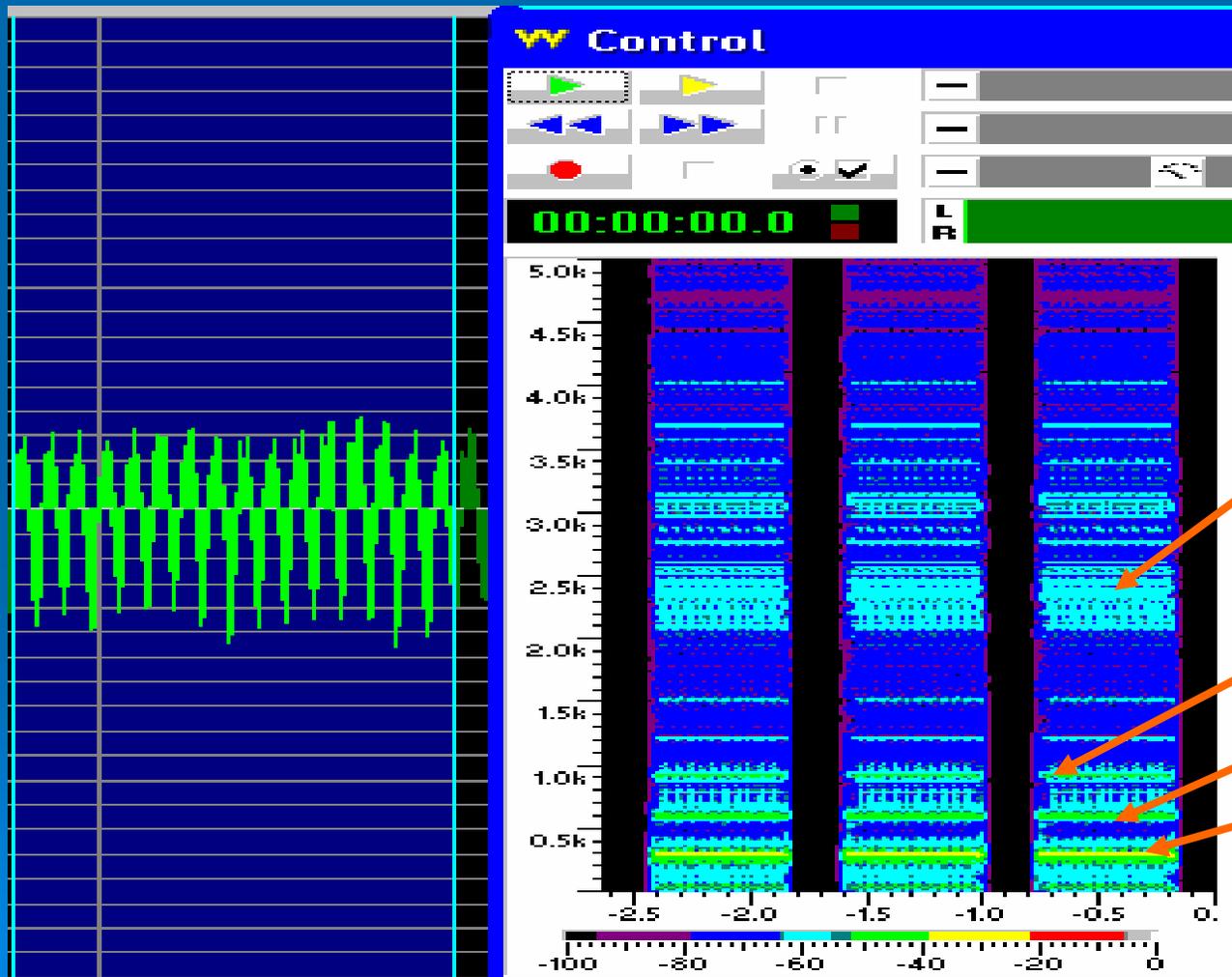
# Annotation for the sentence Vine mama



# Processing tools (I)



# Processing tools (II)



F3  
F2  
F1  
F0



# Results for joy compared with furry states

Subject	F0			F1			F2		
	a	ea	ă	a	ea	ă	a	ea	ă
20048f	±	+	±	±	+	-	+	+	+
01312f	+	+	±	±	±	±	-	+	+
55555f	-	±	-	-	-	-	±	±	±
123456f	-	+	+	±	+	+	+	+	+
77777m	±	±	+	+	-	+	+	+	+
263315m	±	±	±	-	±	-	+	-	±
14411f	±	-	±	-	-	-	-	-	-
26653	±	±	±	±	±	-	-	-	-

# Results (III)

- The accentuated vowels (like the vowel "i" from the word "vine" and the vowel "a", first "a" from the word "mama") don't offer important information compared with non - accentuated vowels (vowel "e" from the word "vine" and vowel "a", the last from the word "mama").
- The vowel "i" from the word "vine" has random values of the formants; therefore, it does not help in the emotion recognition.

# Results (IV)

- The obtained values for the F0 formant for all the persons decrease in happiness compared with the sadness state.
- The obtained values of the formants F1 and F2, for the vowel “a” (the first “a” from the word “mama”) have the tendency to decrease in happiness compared with sadness state.

# Results (V)

- The happiness from sadness states and fury from sadness states can be distinguished in all cases.
- The happiness from fury states is difficult to distinguished in all cases.
- The emotional intra-speaker states can be clearly distinguished but we cannot specify the emotional inter-speaker states.

# Information site

[http://www.etc.tuiasi.ro/sibm/romanian\\_spoken\\_language](http://www.etc.tuiasi.ro/sibm/romanian_spoken_language)

Version 1.0



Alegeți limba  
ROMÂNĂ

Version 1.0



Select  
ENGLISH

Version 1.0



Choisissez  
LE FRANCAIS

Member of CLARIN

Prof. Horia-Nicolai Teodorescu , Inițiator, coordonator de proiect

Team\*:

Monica Feraru, Doctorand

Marius Zbancioc, Asistent Universitar

Adrian Verbuță, Doctorand

Raluca Ganea, Doctorand

Laura Pistol, Asistent Cercetare

Diana Trandabăț, Asistent Cercetare

Ramona Luca, Doctorand

Mihaela Hnatiuc, Asistent Cercetare

Oana Voroneanu, Asociat

\* Listă deschisă. Viitorii colaboratori vor fi trecuți pe pagină.

Web design: Diana Trandabăț, Doctorand și Horia-Nicolai Teodorescu, Profesor

Concept: Prof. Horia-Nicolai Teodorescu

Administrator sit Laura Pistol, Asistent Cercetare

# Sentences simulating emotional states of mind

🔗 Sounds archive

🔗 Vowels

🔗 Consonants

🔗 Diphthongs

🔗 Sentences

🔗 Standard prosody

🔗 Emotional tones

🔗 Linguistics particularities

🔗 Protocols

🔗 Analysis Instruments

🔗 References

🔗 Other links



## Short sentences in emotional tones - files comparing the states of mind

Short sentences simulating the following emotional states of mind: happiness, sadness, joy, hate, optimism, pessimism, sorrow  
Every file contains seven utterances of the selected sentence, one for every emotional state of mind in the order above mentioned. The recordings comply with the Recording Protocol and were saved in 16 and 24 bits mono in .wav format.

Speaker profile	.wav 16 bits	.wav 24 bits	.ogg 16 bits	.ogg 24 bits	.txt	Annotation*
Speaker profile #1	Aseară Cine a făcut asta Vine mama					
Speaker profile #2	Aseară Cine a făcut asta Vine mama					
Speaker profile #3	Aseară Cine a făcut asta Vine mama					
Speaker profile #4	Aseară Cine a făcut asta Vine mama					
Speaker profile #5	Aseară Cine a făcut asta Vine mama					

\* The annotations were made in Praat. To view them, load in Praat the desired sound and the related TextGrid file.

The analysis of emotions in determining the formants and the results obtained can be found [here](#).

[Back to sentences.](#)

# Sentences simulating happiness state

## Voiced Sounds of the Romanian Language

- Home
- Romanian Language Phonetics
- Sounds archive
  - Vowels
  - Consonants
  - Diphthongs
  - Sentences
    - Standard prosody
    - Emotional tones
    - Linguistics particularities
- Protocols
- Analysis Instruments
- References
- Other links

If you have accidentally arrived at this page without reading and fully agreeing with the [Legal Disclaimer and Copyright Notice](#), please [click here](#).



### Emotional tones - joy

The recordings comply with the [Recording Protocol](#) and were saved in 16 and 24 bits mono in .wav format.

Speaker profile	.wav 16 bits	.wav 24 bits	.ogg 16 bits	.ogg 24 bits
Speaker profile #1	Aseară	Aseară	Aseară	Aseară
	Cine a făcut asta			
	Vine mama	Vine mama	Vine mama	Vine mama
	Ai venit iar la mine			
Speaker profile #2	Aseară	Aseară	Aseară	Aseară
	Cine a făcut asta			
	Vine mama	Vine mama	Vine mama	Vine mama
	Ai venit iar la mine			
Speaker profile #3	Aseară	Aseară	Aseară	Aseară
	Cine a făcut asta			
	Vine mama	Vine mama	Vine mama	Vine mama
	Ai venit iar la mine			
Speaker profile #4	Aseară	Aseară	Aseară	Aseară
	Cine a făcut asta			
	Vine mama	Vine mama	Vine mama	Vine mama
	Ai venit iar la mine			

# Examples of sounds

➤ 123456f – happiness state



➤ 01312f – sadness state



➤ 263315m – fury state



➤ 12312m – neutral tone



# Discussions and conclusions

- We have indicated a methodology to choose a stable section of the vowels for the analysis and to compare different analyzers results in order to improve consistency in measurements.
- No tool provides irrefutable results.
- The most informative prosodic feature is the pitch for sentences uttered with manifested emotions.
- This conclusion is similar with the results reported for other languages.

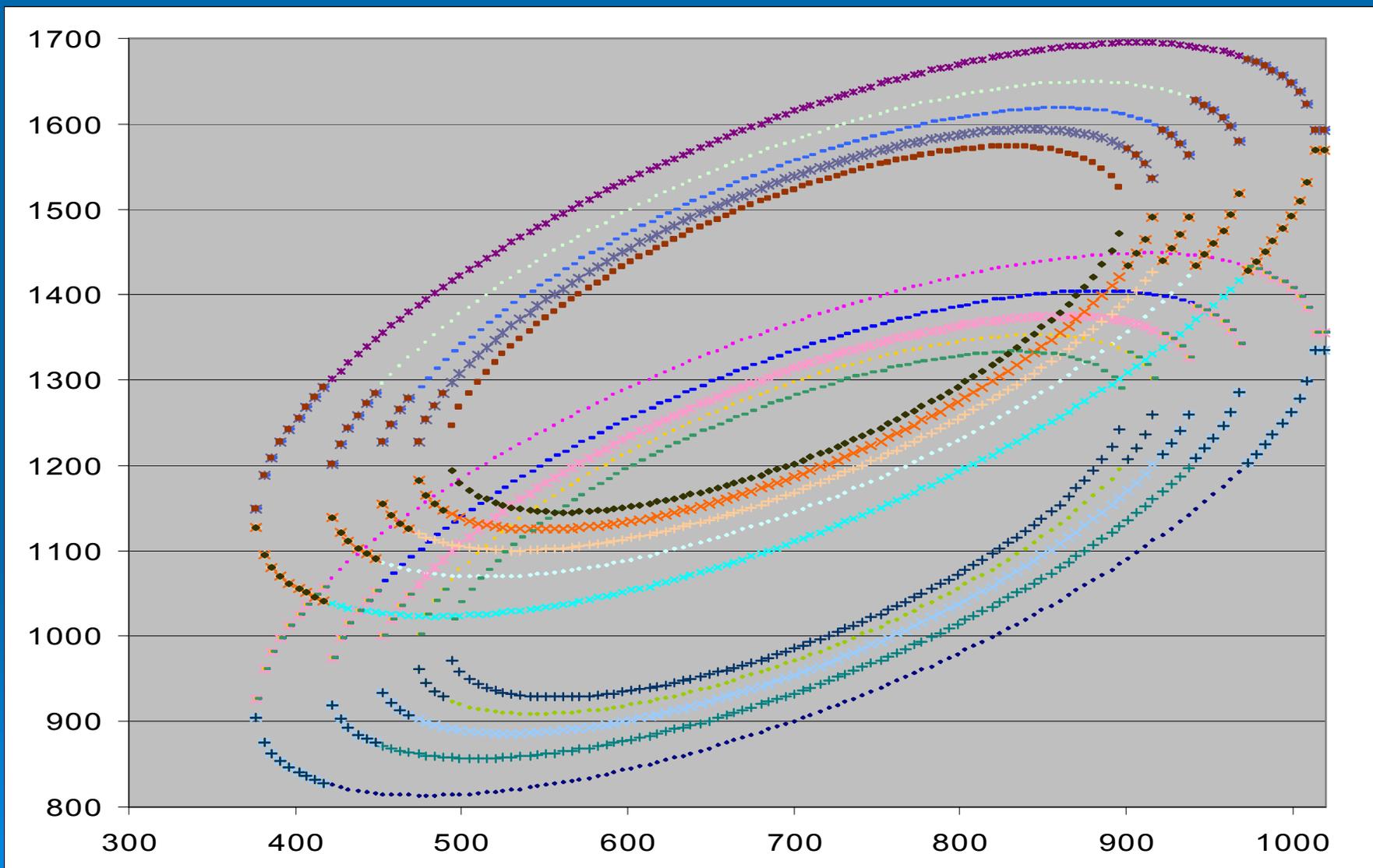
# Further work

- More statistical data (more subjects);
- Automatic classifier;

# References

- <https://nats-www.informatik.uni-hamburg.de/intern/proceedings/2004/LREC/C/pdf/41.pdf>
- <http://pascal.kgw.tu-berlin.de/emodb/>
- <http://kom.aau.dk/~tb/speech/Emotions/defs.pdf>.
- <http://serpens.salleurl.edu/intranet/pdf/239.pdf>.

# The statistical results



Thank you!

