

# Acoustic, Semantic and Personality Dimensions in the Speech of Traditional Puppeteers

M. Rusko\*, J. Hamar\*\* and Š. Beňuš\*\*\*,\*

\* Institute of Informatics of the Slovak Academy of Sciences, Bratislava, Slovakia

\*\* Department of Aesthetics, Faculty of Philosophy of the Comenius University, Bratislava, Slovakia

\*\*\* Constantine the Philosopher University, Nitra, Slovakia

Milan.Rusko@savba.sk, juraj.hamar@sluk.sk, sbenus@ukf.sk

**Abstract**— Puppeteering represents a traditional but slowly dying folk art in which impersonation based on modifying the puppeteer’s voice belongs among the core expressive and stylistic devices. A study of the formal and functional aspects of voice quality and other prosodic features in puppeteers’ production as well as their acoustic correlates is thus an integral part of studies into communicative social signals especially in emotional and expressive speech. This analysis becomes even more relevant in cases when puppeteers’ performance represents a very salient communication code in a particular culture. Better understanding of such traditional code opens possibilities of applications into several speech processing domains. Our data consist of recordings of plays and excerpts by two puppeteers (a father and a son) and verbal descriptions of personality features of the characters that the actor tries to express through their voices. We propose an approach to the taxonomy of archetypal characters based on analyzing the psychological, aesthetic and acoustic-phonetic aspects of their personalities. More specifically, we identify salient links between acoustic-prosodic features on the one hand and semantic and personality features on the other hand.

## I. THE ART OF TRADITIONAL PUPPETEERS AS A SOURCE OF KNOWLEDGE ABOUT SPEECH AND COMMUNICATION

Characters in a traditional puppet play are diverse and their personalities are expressed not only through visual mode as costumes and movement patterns but primarily through semantic, prosodic and voice quality characteristics. The puppeteer thus has to adjust his/her own features to suit for each character as well as to differentiate among the characters. Better understanding of formal and functional aspects of such traditional code linking personality with speech characteristics opens many possibilities for applications into several speech processing domains. This paper presents an extension of an earlier study [1] that aimed at uncovering mutual relations between the personality of the character and its semantic features, as well as between the personality and the acoustic features of the character’s voice.

Understanding the art of puppeteers is intimately linked to the study of communicative social signals for several reasons. First, many plays are meant for children and thus the relationship between the social signals the puppets send during a play represent the core social signals thoroughly ingrained in the culture and easily decoded by small children. Second, a puppet play is multi-dimensional communication par-excellence since the communication

channels exist between the puppeteer and the audience, or between individual characters present in the play. Third, this type of acted speech highlights the most archetypal characteristics of human behavioural interaction, allowing thus identification of the most salient features in these social signals. Fourth, this type of communication is multi-modal and extremely dynamic, which are prototypical features of any spontaneous human-human communication. Finally, in addition to improving our understanding in the above mentioned areas, knowledge gained from studying highly expressive and emotional communication like this has a great potential to be utilized in advanced – that is, more natural, expressive, or human-like – interactive dialogue information systems and intelligent avatars.

Our corpus is rather limited both in terms of quality of recordings and coverage of puppeteers, but it makes up for these limitations by representing unique material of Slovak cultural heritage as a fruitful field for explorations into the most salient acoustically communicated personal traits. The core of the corpus, that we plan to extend in the future, represents old records of low acoustic quality of eight pieces played by Bohuslav Anderle (1913-1976) recorded by his son Anton, newer records of parts of the same plays played by his son Anton Anderle (1944-2008) demonstrating the typology of voices (21 voices of different characters together with the actor’s natural voice have been recorded by the first two authors), and finally, verbal descriptions of personality features of the characters by A. Anderle. The main goal of the paper is to present a classification of voice displays of characters, techniques of voice changes, and their acoustic correlates. Additionally, we propose an approach to the taxonomy of archetypal characters based on analyzing the psychological, and aesthetic aspects of their personalities.

## II. A BRIEF HISTORICAL DESCRIPTION

Despite the records of a relatively large number of puppeteers in the first half of the 20<sup>th</sup> century (roughly 20), only a single actively performing puppeteer in the 21<sup>st</sup> century was A. Anderle. Such a sharp decline cannot be entirely due to general diminishing folk traditions and is probably linked to systematic efforts of the communist regime to suppress this tradition. A. Anderle represented the third generation of puppet theatre performers and his father Bohuslav was a respected puppeteer with up to 40 scripts in his repertoire.

In 1970’s, A. Anderle tape-recorded some of the performances of his father and learned the craft under his

supervision. In early 1990s Anton re-started the old tradition and started performing the plays from his father’s repertoire again. Until his unexpected death in 2008, Anton Anderle’s enjoyed respect and admiration both from the viewers as well as other puppeteers in Europe.

### III. INDIVIDUALITY AND ITS COMPONENTS, THE NOTION OF PERSONALITY

In this work, individuality is understood as a psychological entity, a unit consisting of three components [2]. First, personality is a rather stable component that remains practically unchanged during one’s life. Second, mood may change in time, but its changes are rather slow. Third, emotions are the most dynamical and change rather rapidly. In this paper we focus mainly on the first component personality and leave the other two components for subsequent studies.

Personality is most commonly described using the formal psychological model called the Five Factor Model [3], [4] with the factors representing the basis of the personality space. Assuming, that the personalities of the characters from puppet plays represent archetypes, we have adopted a very simple annotation convention. Each of the personality dimensions (Neuroticism, Extraversion, Openness to experience, Agreeableness and Conscientiousness) will be assigned only one of three values: 1, 0 or -1. For instance, N1 denotes that the character is neurotic, N0 means that this dimension is not applicable, i.e. not important for expressing the character’s personality, and N-1 denotes the absence of neuroticism, i.e. that the character comes across as secure and confident, which is opposite to neurotic.

### IV. PARAMETERS DESCRIBING VOICE QUALITY AND SETTINGS OF THE ARTICULATORS

A common way of describing voice settings uses the notion of a reference or neutral setting. This reference corresponds to a normal position relative to possible adjustments [5]. Discrete or continuous variation in voice settings is then depicted as deviations from the reference/neutral setting.

Following the basic pattern outlined in [5], one can then attempt to classify voice qualities primarily in terms of articulatory descriptions. Although our features primarily consist of prosodic characteristics extractable from the acoustic signal, and there exist multiple non-linear relationships between articulation and acoustics in speech [11], we consider Laver’s articulatory-based classification as currently the most suitable for labeling highly expressive speech. For annotation we used a simple set of labels derived from Laver’s terminology e.g. Labial protrusion = LP, Laryngopharyngealized = LPH, Denasal = DN, Harsh whispery creaky falsetto = HWCF, etc.

Laver’s classification scheme is considered to be carefully worked-out, and it is being used widely. Despite this, however, some speech qualities are not covered, e.g. smiling or weepy speech. In producing these types of speech, complex positioning of the articulators (wide high/top and wide and low/bottom mouth corner positioning) along with special phonation modes (vibrato etc.) are used, and these are not included in the scheme. Pathological phenomena occurring in spoken utterances, whether acted or natural, such as lisping, stammering,

muttering are not included in the scheme either; we have added the PAT (Pathological) annotation mark for them.

Considering prosodic features, we denote slow speech as SRL (Speech rate low), fast speech as SRH (Speech rate high), large pitch range as PRH (Pitch range high), small pitch range as PRL (Pitch range low), and low voice pitch is denoted as LOW.

A complex feature covering both voice quality and prosody is vocal effort. We denote high vocal effort as VEH (Vocal effort high) and low vocal effort as VEL (Vocal effort low).

### V. PROSODY-EMOTIONS RELATIONSHIP IN EXPRESSIVE SPEECH RESEARCH

There has been extensive research into the systematic links between prosody and emotions. The studies mainly investigate three prosodic dimensions: melodic contour of the fundamental frequency (F0), metrical and rhythmical structuring of prosodic units, and voice quality and their relation to discrete (e.g. happiness, or binary valence/arousal space) or continuous (arousal, power and valence labeled on a scale between -1 and 1) representations of affective state of the speaker; see for example [6], [7] or [8] for an overview. In these domains, the finding that seems most salient across various corpora is that pitch correlates positively with the degree of arousal [9]. Additionally, arousal also correlates well with mean of F0, mean intensity, speech rate, and pitch range [10].

### VI. SEMANTIC DIMENSION

The description of personalities encoded in the speech and acting of puppet characters requires at least a three-dimensional space consisting of semantic, visual and auditory dimensions. In the semantic domain the character is best described following its functions in the play. Table I. gives classification of functions of the characters (the concept) on the basis of elementary semantic oppositions, proposed by us for aesthetic and semantic description of characters for the purposes of this study.

TABLE I.  
TAXONOMY OF CHARACTERS BASED ON ELEMENTARY SEMANTIC OPPOSITIONS

Criteria	One pole	code	Opposite pole	code
Sex	Male	<b>XM</b>	Female	<b>XF</b>
Anthropological view	human	<b>HH</b>	Non-human	<b>HN</b>
Age	Old	<b>AO</b>	Young	<b>AY</b>
Morality	Positive	<b>MP</b>	Negative	<b>MN</b>
Aesthetics	Tragical	<b>ET</b>	Comical	<b>EC</b>
Ethnicity	Domestic	<b>RO</b>	Foreign	<b>RF</b>
Social status	Noble	<b>SN</b>	Low	<b>SL</b>
Intelligence	Clever	<b>IH</b>	Stupid	<b>IL</b>

The classification includes eight dimensions. Some of them are binary, e.g. Sex or Anthropological view, some are coarsely continuous, e.g. Social Status or Ethnicity,

and some represent a fine-grained continuum, e.g. Age, Morality, or Intelligence. We will initially code all dimensions as binary since even Age, Morality, or Intelligence are considered archetypal and thus extremely polar for the purposes of the plays.

## VII. VISUAL AND ACOUSTIC DIMENSIONS

Simultaneously with the semantic dimension, certain archetypes of traditional characters participate in the creation of personalities of individual characters. These archetypes are mostly expressed through the visual and auditory layers. Hence, the three layers cannot be considered separately but as a system of interleaving and interdependent relationships. The visual domain includes all elements of the visual realization of a puppet (its face, costume, material and animation). The auditory domain corresponds to emotional, social and linguistic markers of a character through mediating the plot and the text strategy or bearing poetics and aesthetics of the whole play. The auditory layer "vitalizes" and underlines the characteristics and contexts introduced in the semantic and visual layers. It includes primarily speech, but also musical accompaniment as well as non-verbal sounds produced either by the actor or generated by the motion of the puppets.

As our field research also has an ethno-musicologist aspect, we elicited from A. Anderle the descriptions of characters' personalities and explanations of the means for changing his voice that he uses to bring these personality features to life. The actor-puppeteer thus based his descriptions on both socio-psychological characteristics of the characters as well as the acoustic and articulatory means for expressing them.

TABLE II.  
TYPOLOGY OF CHARACTERS AS GIVEN BY ANTON ANDERLE

Character typology	Properties	Voice
I. Negative male type	Intriguer, bad knight	High volume, hyper-articulation
II. Positive male type	Leading man	-
III. Royal type	dignified, deliberate, wise	Low pitch, monotonous
IV. Bad man	-	hoarse, low pitch
V. Swaggerer	Convivial, bold farmer, folk type, straight man, unshuffling, not cunning, frank	Pharyngeal resonance, great pitch range
VI. Lead woman	-	young, soft modal
VII. Old woman	-	-
VIII. Bad old woman	Cunning, sarcastic	Increased hoarseness, articulator setting as for smile
IX. Good old woman	-	Low falsetto, medium pitch range

In Table II we present an initial partial taxonomic model for the classification of basic types of voices. The first column follows the actor verbal descriptions and the other two columns our additions: the second column gives our judgment of some features related to personality, and the third column to speech quality. The missing cells represent cases for which the description

was not provided by the actor or The complete lists of characters belonging to the particular groups can be found in [12]. This model is further extended in Table III and described in the following section.

## VIII. RELATIONSHIP BETWEEN PERSONALITIES AND SEMANTIC AND ACOUSTIC CHARACTERISTICS

We have two standpoints from Anton Anderle to the issue of character types and voices he renders. We first recorded his demonstration and brief description of individual character types. During our next meeting he summarized the typology of characters for us by characterizing a character archetype (in some cases the voice type used as well) and by presenting several examples of this type including the title of the play where each of them acts.

The first recording aimed at better understanding actor's ideas regarding the personalities of his characters. We visited Anton Anderle in his house and asked him to play the basic voice types he uses, and to give a brief description of each of them. Professional quality audio record was made, and the presentation was also recorded by digital video camera. Fig. 1 shows Anton Anderle recording sample voices in his house in Banská Bystrica.

Although this was no easy task, the actor was able to present classification of his characters off-hand; his classification was based on several basic characteristics including both psychological features and acoustic and phonetic means for expressing them.



Figure 1. Anton Anderle recording sample voices in his house in Banská Bystrica

Table III presents a short list of several selected characters, their names, the dramatic piece in which they perform, and our remarks. We had intended to record larger parts of plays, or possibly whole plays later. Unfortunately, due to unexpected death of the actor this intention could not be realized.

In the same table we also present the application of our classification scheme - a partial overview of the taxonomy of basic types of voices given by the actor supplemented with our judgments of their personality dimensions (in code), speech quality description and its code, as well as some additional remarks.

The complete list of characters with coding and information illustrated in Table III includes 24 voices (the

TABLE III. A FRACTION (FIRST 8 CHARACTERS) OF THE ACTOR'S CLASSIFICATION AND OUR ASSESSMENT OF PERSONALITY DIMENSIONS, AND VOICE QUALITY; OUR REMARKS (THE COMPLETE TABLE CAN BE FOUND IN [12])

Name of the character (title of relevant play in brackets)	The actor's concept	Personality parameters values	Speech quality, acoustic parameters, other remarks	Semantic features
Actor	Neutral, natural	E1, O1, A1	MV	XM, HH;
Persian shah's daughter (Faust)	Young, gentle judicious, relaxed, womanly, kind, well-mannered	N-1, O1, A1	F, VEL	XF, HH, AY, MP, SH, IH,
Dona Karolinka ( <i>Don Šajn</i> )	Young, gentle, judicious, relaxed, kind, womanly, polite	N-1, O1, A1	F, VEL	XF, HH, AY, MP, SH, IH,
Žabinka – Gašparko's wife (several plays)	Hysterical, funny	N1, O-1,	RL, F, SRH sounding, sharp	XF, HH, AY, MP, EC, RO, SL;
Countess (Belengardo)	Angry, assertive, provocative, shrill	N1, E1, A-1	F, VEH, SRH	XF, HH, AO, MN, SH,
Angel (Faust)	Insisting	A1, C1	TV, F, VEH, PRH specific (falling) intonation outlines	XF, HN, MP
Faust (Faust)	Scholar, philosopher, serious	O1, C1	MV pathetic	XM, HH, AO, ET, SH, IH
Gašparko (all plays)	Funny, cheeky, liar,	N1, E1, O1	RL, TV, PRH, SRH, PAT raised larynx, lisping (s-š, c-č)	XM, MP, EC, RO, SL;
Belengardo (Belengardo)	Intrigant, self-assured, vigorous, loud, self-assured, convivial	E1, A-1, C1	LS, HV, higher throat resonance	XM, HH, AO, MN, SH;

TABLE IV. COUNTS AND MUTUAL COMMON OCCURRENCES OF PERSONALITY DIMENSIONS AND SEMANTIC CHARACTERISTICS

***	XM	XF	HH	HN	AO	AY	MP	MN	ET	EC	RO	RF	SH	SL	IH	IL	SUM	%
N1	3	4	4	2	2	1	2	4		4	3		1	4		3	37	13,4
N-1	4	2	4	1	2	1	3	1	1	1		1	3	1	3		28	10,1
E1	7	2	7	1	3		2	4		3	2	2	2	5		2	42	15,2
E-1	2	1	3		1		1	1		2	1	1		3		1	17	6,16
O1	3	2	3		1	1	2		1	1	1		2	1	3		21	7,61
O-1	2	2	4		1	1	3	1		2	1	1		3			21	7,61
A1	5	4	5	3	2	1	5	2		2	1		2	3	3	2	40	14,5
A-1	2	3	3	2	2			5	1				2		1		21	7,61
C1	8	3	6	4	2		4	3	1	1		1	4	1	2		40	14,5
C-1	1		2		1					1		2		2			9	3,26
SUM	37	23	41	13	17	5	22	21	4	17	9	8	16	23	11	9	276	100
%	13,41	8,33	14,86	4,71	6,16	1,81	7,97	7,61	1,45	6,16	3,26	2,90	5,80	8,33	3,99	3,26	100	***

TABLE V. COUNTS AND MUTUAL COMMON OCCURRENCES OF PERSONALITY DIMENSIONS AND VOICE CHARACTERISTICS

***	PAT	SRH	SRL	PRH	PRL	VEH	VEL	LOW	WV	F	CF	HV	RL	TV	MV	LV	LS	CR	LP	LL	N	BV	DN	SUM	%	
N1	3	4		2		3				3	1		2	1			2					2			23	12,6
N-1			3		2		2	5	1	2		1				2		2	1	2		1			24	13,1
E1	2	4		2		3		2		2		1	1	1			2					1	2	2	26	14,2
E-1	1		2		1		1					1										1			7	3,83
O1	1			1		2		2		2			1	1	2										10	5,46
O-1		1	2		1	1	2	1		1	1		1				1				1				13	7,1
A1	2	1		1		2	3	2		5			1	1	1	1	1	1	1		2				25	13,7
A-1		3		1		3		1		2	1	1					2				1				15	8,2
C1			3	1	2	2	1	5	1	3	1	2		1	1	2	2	2	1	3		1			34	18,6
C-1							2																2	2	6	3,28
SUM	9	13	10	8	6	13	10	19	2	20	3	7	5	5	5	6	10	6	3	7	6	6	4	183	100	
%	4,92	7,10	5,46	4,37	3,28	7,10	5,46	10,38	1,09	10,93	1,64	3,83	2,73	2,73	2,73	3,28	5,46	3,28	1,64	3,83	3,28	3,28	2,19	100	***	

actor's own voice and 23 characters). These data can be used for a first analysis of mutual links among personality factors, semantic and articulatory-acoustic features. Tables IV and V (at the end of the paper) provide such a first descriptive 2D analysis of mutual common occurrence of personality dimensions and semantic and acoustic characteristic. In the following, we evaluate and discuss the raw frequencies of marks assigned during annotation and relative ratios in percentages shown in these two tables.

As expected, the 2D analysis performed on a relatively limited number of data – does not provide clear answers to the queries related to coding of personality characters by aesthetic-semantic and acoustic speech means. However, the results in Table IV still suggest some dependencies. For example, negative moral features (MN) can be observed with neurotic (N1), extrovert (E1) and competitive (A-1) characters. Comical characters (EC) are often neurotic (N1). High social position (SH) is connected with calmness (N-1), extroversion (E), openness to new impressions (O) and strong-mindedness (C). Similar personality characteristics also tend to correlate with wisdom (IH).

Results in Table V suggest that actors use mostly pitch changes in their voice (LOW+F+CF=22.95%) to express diversity of characters. While female voices (F+CF=12.57% of the total of assigned acoustic marks) are naturally expressed by falsetto, low voices (LOW=10.38% acoustic marks) correlate robustly with the N-1 factor, i.e. with calm and self-assured nature, and obviously with orderliness and resolution (C1). Additionally, most often used acoustic means include speech rate (SRH+SRL=12.57%) and voice effort intensity (VEH+VEL=12.57%). High speech rate is usually related to neuroticism (N1), extroversion (E1), but also to competitiveness and assertiveness (A-1). On the other hand, slow speech (SRL) tends to be linked to reliability (C1).

Considerable range of frequencies of the basic tone in melodic structures (PRH) and high voice effort (VEH), have also been used several times to express neurotic and extrovert nature. More data would be necessary for us to be able to evaluate the function of additional voice properties.

## IX. CONCLUSION

The goal of this paper was to improve our understanding of the acoustic and auditory correlates of personality dimensions. We introduced a novel approach to the analysis of functional variation, i.e. the need to express personalities of particular characters, in the speech and vocal features of a puppeteer. We argued that the system of stylized personality expressions by a puppeteer provides an excellent source of information both for understanding cognitive aspects of social communicative signals in human-human interactions as well as for utilization of observed patterns of human behavior in applications based on interactive voice systems in human-machine interactions.

The paper describes the collection of speech material and suggestions for annotation schemes. Specifically, Five Factor Model of personality dimensions [2], [3] and a model of voice quality features [5] with more acoustic characteristics added were adapted and used for creating a database from this speech material. We also introduced a semantic typology of characters utilizing elementary semantic oppositions. With the help of a simple code for the three dimensions, it was then possible to describe all characters that appear in our recorded database of puppet plays.

Results from two-dimensional analyses are tabulated, and suggest correlation among certain semantic and acoustic characteristics and the personality factors.

We hope this work improves our understanding of the relationship between the acoustic and semantic aspects of personality which, along with emotions, affection, attitude and mood, belong to the fundamental areas of expressive speech research.

## ACKNOWLEDGMENT

This research was supported by the grant 2/0202/11 awarded by the Scientific Grant Agency of the Slovak Ministry of Education.

## REFERENCES

- [1] M. Rusko, J. Hamar, "Character Identity Expression in Vocal Performance of Traditional Puppeteers," Text, Speech and Dialogue, 9th International Conference, TSD 2006, LNAI 4188, pp. 509-516.
- [2] A. Egges, S. Kshirsagar, N. Magnenat-Thalmann, "Imparting Individuality to Virtual Humans. First International Workshop on Virtual Reality Rehabilitation, pp. 201-108, 2002.
- [3] J. M. Digman, "Personality structure: Emergence of the five factor model," *Annual Review of Psychology*, 41, pp. 417-440, 1990.
- [4] R.R. McCrae,; O.P. John, "An introduction to the five-factor model and its applications," *Journal of Personality* 60, pp.175-215, 1992.
- [5] J. Laver, *The gift of speech*, Edinburgh, UK: Edinburgh University Press, 1991.
- [6] M. Schroder, *Speech and Emotion Research: An Overview of Frameworks and a Dimensional Approach to Emotional Speech Synthesis*, Ph.D. dissertation, Univ. of Saarland, Germany, 2003.
- [7] H. Gunes, B. Schuller, M. Pantic, and R. Cowie, "Emotion Representation, Analysis and Synthesis in Continuous Space: A Survey", Proc. of the 1st International Workshop on Emotion Synthesis, rePresentation, and Analysis in Continuous spacE, IEEE Press, pp. 827-834, 2011.
- [8] B. Schuller et al., "Acoustic emotion recognition: A benchmark comparison of performances," in Proc. IEEE ASRU, 2009.
- [9] R.A. Calvo & S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Trans. On Affective Computing*, 1(1), pp. 18–37, 2010.
- [10] G. L. Huttar, "Relations between prosodic variables and emotions in normal american english utterances," *J. Speech Hearing Res.*, 11, pp. 481–487, 1968.
- [11] K. N. Stevens, "On the quantal nature of speech," *Journal of Phonetics*, 17, pp. 3-45, 1989.
- [12] M. Rusko, J. Hamar, "Akustické a sémantické aspekty vyjadrenia personality postáv v tradičnom bábkovom divadle," In: HAMAR, J., et. al., *Ludové bábkové divadlo v Európe*, Bratislava: Slovenské centrum UNIMA, pp. 45 – 67, 2010.