ON THE STUDY OF THE RELATIONSHIP BETWEEN THE PHONEMIC CONTENT AND THE EMOTIONAL VALENCE OF WRITTEN LANGUAGE

Velina Slavova, Filip Andonov

Abstract: *This study investigates the relationship between the sublexical content of texts in English and the emotional valence they evoke. The statistical analysis of emotionally annotated sentences revealed a strong correlation between this sublexical presentation and the evaluations of valence provided by the readers. The weights of the sublexical units representing the emotional valence, obtained from the corpus, were applied on data outside of it, proving the validity of the finding.*

Keywords: *Natural language processing, Sound-symbolism, Corpus linguistics, Text emotion recognition*

ITHEA Keywords: *I.2.7 Natural Language Processing*

DOI: https://doi.org/10.54521/ijita30-03-p01

Introduction

The assumption researched in this paper is based on the fact that languages are used to convey information.

In compliance with the distinction between information and data accepted in the domain of data mining, where the process of data mining is defined as the extraction of implicit, previously unknown, and potentially useful information from data, the term information, used hereafter, is based on the following definition: “Data are sensory stimuli that we perceive through our senses. Information is data that has been processed into a form that is meaningful to the recipient.”[Davis & Olson, 1985].

Information is encoded by the speaker, who is the source of the language message, and is decoded by the recipient, who is assumed to use the same conceptual frame when decoding it (Fig. 1). The general idea of the research presented here is based on the hypothesis that inside the phonetic characteristics of speech both meaning and emotion are encoded and transmitted.

Figure 1. General Outline

Specialized research leads to the conclusion that emotions are fundamental characteristics of several animal species and the communication in their societies is used to convey emotional information. As for the human emotions, their study and categorization dates back to antiquity. Two models of emotion are used in the contemporary scientific fields – a discrete model and VAD. The discrete model [Ekman, P., 1999] is based on the way we label emotions in our daily interactions. There is however no consensus on the labels and as a result several sets of labels exist.

Recently in science a different approach is preferred. The most popular model [Russell & Mehrabian, 1977] is called VAD. It is an abbreviation of the words Valence, Arousal and Dominance. Those are the three dimensions that all emotions are placed in. Valence indicates the positive or negative (pleasant or unpleasant) aspect of the emotion. Arousal indicates the degree of excitement. Dominance shows one’s status regarding the emotion-causing event.

Some results from other domains

The studies, presented here, investigate the hypothesis that emotions can be found in the phonemic content of speech and can therefore be found in written language.

Although the exact “biological” mechanisms by which language transmits emotion are not fully understood, recent results have shown a clear connection between language constituents and emotion. Numerous pieces of evidence from functional magnetic resonance imaging (fMRI) also showed the existence of specific brain regions activated only by unpleasant words and regions activated only by pleasant words. Such results lead to the idea that some indicators of emotional valence appear in the “meaning-communication protocol” – that is – they are detectable in speaking and writing.

The analysis presented next concerns emotional processing in reading. It is known from fMRI and event-related potential (ERP) studies that after visual recognition of the word-input, the orthographic information is transmitted for „analysis” to the auditory cortex where it automatically gains its corresponding phonetic content. Recognition of lexical meaning proceeds in accord with the processing of presented spoken word-sequences – that is, as perceived speech. For that reason in our study work with written language, assuming that the transmission of emotion is detectable in the phonemic content of the text itself.

Several recent studies confirm this assumption. For example Ariani, Ullrich, and colleagues [Aryani, Kraxenberger, Ullrich et al, 2016] performed several text-emotion studies of German poetry and found that the sublexical representation contains features that can predict emotional effects. In these studies, the authors applied the reasoning of Gestalt psychology and examined the language phonological content into layers. The sublexical level investigated by Ariani and colleagues showed that both single phonemes and sub-syllabic segments (onset, nucleus, and coda) provide information that is relevant for the investigation of emotion-related phonological iconicity. In the approach presented here we work with the assumption that a text is a whole-body composed of gradual phonemic layers – from phones to syllables, words, sentences, etc. and all of these layers influence the reader.

Another recent study of Adelman and colleagues [Adelman, Estes & Cossu 2018] conducted a data-analysis emotion evaluated corpora of English, German, Dutch, Spanish, and Polish words. They discovered that for all these languages there is a correlation between the phonological properties of individual phonemes and the emotional valence of these words. All of these findings, the authors reasoned, would be likely to benefit a person's chances of survival in cases of suddenly announced danger.

The important result that led us to investigate a more specific phonemic layer is the exploration of the emotion-effects of syllables, composed of two phonemes. Kawahara and Shinohara [Kawahara & Shinohara 2012] investigated phoneme pairs of the type consonant-vowel (c-v) in pseudo-words, organized in two types of phonological stimuli of non-words: “abrupt” (“Stop condition” – e.g. [/tadi/]) and “sonorant” (“Sonorant condition” – e.g. [/maji/]). The study demonstrated a tripartite trans-modal sound-symbolic relationship between auditory speech-sounds, visual shapes, and emotions.

Having those previous research results in mind, we decided to investigate further the possibility of emotional information encoded in the sublexical content of pairs of sounds, containing one vowel and one consonant, further called biphones.

Performed investigations

We conducted several studies of these phenomena on texts in English, gradually increasing our understanding of it.

In our first attempt to study the phenomena we investigated the relationship between the sublexical level of texts in natural language (English) and the valence of the texts (Fig. 2), based on phoneme pairs of the type – vowel-consonant (v-c), that are the inverse of what Kawahara, S., & Shinohara investigated. Two phonological types of v-c were investigated – vowel-abrupt (e.g., /ʊt/) and vowel-sonorant (e.g.,/iːm/). The statistical result showed that subsets of both types of v-c phoneme pairings were correlated with valence [Slavova, 2019].

The result from this experiment was promising and we continued our research. Our primary goal was to accumulate a statistically significant amount of reliably evaluated texts.

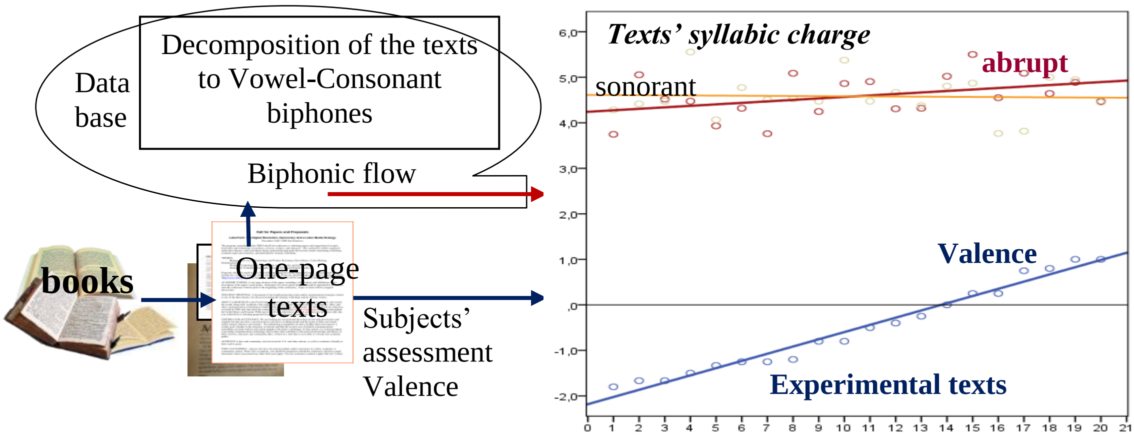

Figure 2. Settings of the Initial experimental and part of the result

We found a solution by using the EmoBank corpus [EmoBank] - a large-scale text corpus manually annotated with emotion according to the psychological

Valence-Arousal-Dominance scheme [Buechel, S., & Hahn, U. 2022]. EmoBank contains 10,549 sentences taken from 136 different documents, classified into 7 genres – blog, essays, fiction, letters, newspaper, “SemEval”, and travel-guides. (“SemEval” is a dataset of news headlines from major newspapers such as the New York Times, CNN, and BBC News, as well as from the Google News search engine).

The EmoBank corpus was locally stored and treated to extract phonemic metadata from emotionally evaluated texts. The text of the sentences was parsed and the phonetic transcription of each word was found using Oxford Online Dictionary and stored in the local database (for details see [Andonov & Slavova 2019]

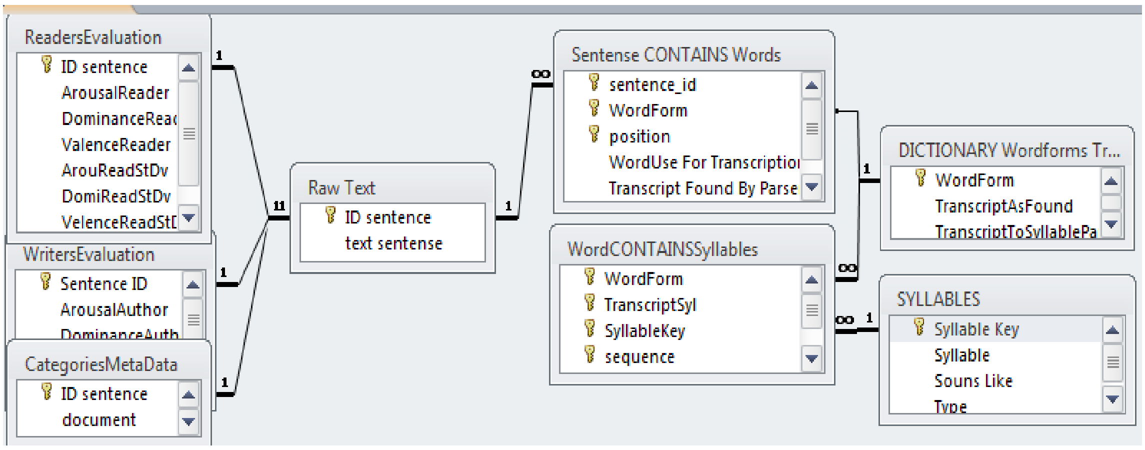

Figure 3. Database for the investigation of the sentences in EmoBank

The get to the sublexical level we investigated biphones consisting of one vowel and one consonant in either order, because in speech such biphones make the word pronunciation possible and are the building blocks of word formation.

As a first step, models for all the possible biphones were obtained using the full set of phonemes as defined in the English language – the set V of vowels and the set C of consonants (Fig. 4). Next, the phonemic transcriptions of the words as found in Oxford dictionary were parsed using these models and the biphonic decomposition of the words was obtained.

biph	vowels	ɪ	e	æ	ɒ	ʌ	ʊ	iː	eɪ	aɪ	ɔɪ	uː	əʊ	aʊ	ɪə	eə	ɑː	ɔː	ʊə	ɜː	ə	i	u
nsonants																							
m		mɪ	me	mæ	mɒ	mʌ	mʊ	miː	meɪ	maɪ	mɔɪ	muː	məʊ	maʊ	mɪə	meə	mɑː	mɔː	mʊə	mɜː	mə	mi	mu
n		nɪ	ne	næ	nɒ	nʌ	nʊ	niː	neɪ	naɪ	nɔɪ	nuː	nəʊ	naʊ	nɪə	neə	nɑː	nɔː	nʊə	nɜː	nə	ni	nu
ŋ		ŋɪ	ŋe	ŋæ	ŋɒ	ŋʌ	ŋʊ	ŋiː	ŋeɪ	ŋaɪ	ŋɔɪ	ŋuː	ŋəʊ	ŋaʊ	ŋɪə	ŋeə	ŋɑː	ŋɔː	ŋʊə	ŋɜː	ŋə	ŋi	ŋu
r		rɪ	re	ræ	rɒ	rʌ	rʊ	riː	reɪ	raɪ	rɔɪ	ruː	rəʊ	raʊ	rɪə	reə	rɑː	rɔː	rʊə	rɜː	rə	ri	ru
l		lɪ	le	læ	lɒ	lʌ	lʊ	liː	leɪ	laɪ	lɔɪ	luː	ləʊ	laʊ	lɪə	leə	lɑː	lɔː	lʊə	lɜː	lə	li	lu
w		wɪ	we	wæ	wɒ	wʌ	wʊ	wiː	weɪ	waɪ	wɔɪ	wuː	wəʊ	waʊ	wɪə	weə	wɑː	wɔː	wʊə	wɜː	wə	wi	wu
j		jɪ	je	jæ	jɒ	jʌ	jʊ	jiː	jeɪ	jaɪ	jɔɪ	juː	jəʊ	jaʊ	jɪə	jeə	jɑː	jɔː	jʊə	jɜː	jə	ji	ju
p		pɪ	pe	pæ	pɒ	pʌ	pʊ	piː	peɪ	paɪ	pɔɪ	puː	pəʊ	paʊ	pɪə	peə	pɑː	pɔː	pʊə	pɜː	pə	pi	pu
b		bɪ	be	bæ	bɒ	bʌ	bʊ	biː	beɪ	baɪ	bɔɪ	buː	bəʊ	baʊ	bɪə	beə	bɑː	bɔː	bʊə	bɜː	bə	bi	bu
t		tɪ	te	tæ	tɒ	tʌ	tʊ	tiː	teɪ	taɪ	tɔɪ	tuː	təʊ	taʊ	tɪə	teə	tɑː	tɔː	tʊə	tɜː	tə	ti	tu
d		dɪ	de	dæ	dɒ	dʌ	dʊ	diː	deɪ	daɪ	dɔɪ	duː	dəʊ	daʊ	dɪə	deə	dɑː	dɔː	dʊə	dɜː	də	di	du
k		kɪ	ke	kæ	kɒ	kʌ	kʊ	kiː	keɪ	kaɪ	kɔɪ	kuː	kəʊ	kaʊ	kɪə	keə	kɑː	kɔː	kʊə	kɜː	kə	ki	ku
g		gɪ	ge	gæ	gɒ	gʌ	gʊ	giː	geɪ	gaɪ	gɔɪ	guː	gəʊ	gaʊ	gɪə	geə	gɑː	gɔː	gʊə	gɜː	gə	gi	gu
tʃ		tʃɪ	tʃe	tʃæ	tʃɒ	tʃʌ	tʃʊ	tʃiː	tʃeɪ	tʃaɪ	tʃɔɪ	tʃuː	tʃəʊ	tʃaʊ	tʃɪə	tʃeə	tʃɑː	tʃɔː	tʃʊə	tʃɜː	tʃə	tʃi	tʃu
dʒ		dʒɪ	dʒe	dʒæ	dʒɒ	dʒʌ	dʒʊ	dʒiː	dʒeɪ	dʒaɪ	dʒɔɪ	dʒuː	dʒəʊ	dʒaʊ	dʒɪə	dʒeə	dʒɑː	dʒɔː	dʒʊə	dʒɜː	dʒə	dʒi	dʒu
f		fɪ	fe	fæ	fɒ	fʌ	fʊ	fiː	feɪ	faɪ	fɔɪ	fuː	fəʊ	faʊ	fɪə	feə	fɑː	fɔː	fʊə	fɜː	fə	fi	fu
v		vɪ	ve	væ	vɒ	vʌ	vʊ	viː	veɪ	vaɪ	vɔɪ	vuː	vəʊ	vaʊ	vɪə	veə	vɑː	vɔː	vʊə	vɜː	və	vi	vu
θ		θɪ	θe	θæ	θɒ	θʌ	θʊ	θiː	θeɪ	θaɪ	θɔɪ	θuː	θəʊ	θaʊ	θɪə	θeə	θɑː	θɔː	θʊə	θɜː	θə	θi	θu
ð		ðɪ	ðe	ðæ	ðɒ	ðʌ	ðʊ	ðiː	ðeɪ	ðaɪ	ðɔɪ	ðuː	ðəʊ	ðaʊ	ðɪə	ðeə	ðɑː	ðɔː	ðʊə	ðɜː	ðə	ði	ðu
s		sɪ	se	sæ	sɒ	sʌ	sʊ	siː	seɪ	saɪ	sɔɪ	suː	səʊ	saʊ	sɪə	seə	sɑː	sɔː	sʊə	sɜː	sə	si	su
z		zɪ	ze	zæ	zɒ	zʌ	zʊ	ziː	zeɪ	zaɪ	zɔɪ	zuː	zəʊ	zaʊ	zɪə	zeə	zɑː	zɔː	zʊə	zɜː	zə	zi	zu
ʃ		ʃɪ	ʃe	ʃæ	ʃɒ	ʃʌ	ʃʊ	ʃiː	ʃeɪ	ʃaɪ	ʃɔɪ	ʃuː	ʃəʊ	ʃaʊ	ʃɪə	ʃeə	ʃɑː	ʃɔː	ʃʊə	ʃɜː	ʃə	ʃi	ʃu
ʒ		ʒɪ	ʒe	ʒæ	ʒɒ	ʒʌ	ʒʊ	ʒiː	ʒeɪ	ʒaɪ	ʒɔɪ	ʒuː	ʒəʊ	ʒaʊ	ʒɪə	ʒeə	ʒɑː	ʒɔː	ʒʊə	ʒɜː	ʒə	ʒi	ʒu
h		hɪ	he	hæ	hɒ	hʌ	hʊ	hiː	heɪ	haɪ	hɔɪ	huː	həʊ	haʊ	hɪə	heə	hɑː	hɔː	hʊə	hɜː	hə	hi	hu

Figure 4. The biphones

Thus having the biphonic decompositions of the words, the biphonic content of each sentence in EmoBank was obtained.

The frequencies of the sublexical units - the biphones - defined in such way turned out to have important differences in positive and negative texts from EmoBank. The plot in Fig. 5 shows the number of occurrences of some biphones in 500 positive and 500 negative sentences from EmoBank. Certain biphones participate with quite a different frequency in positive and negative sentences.

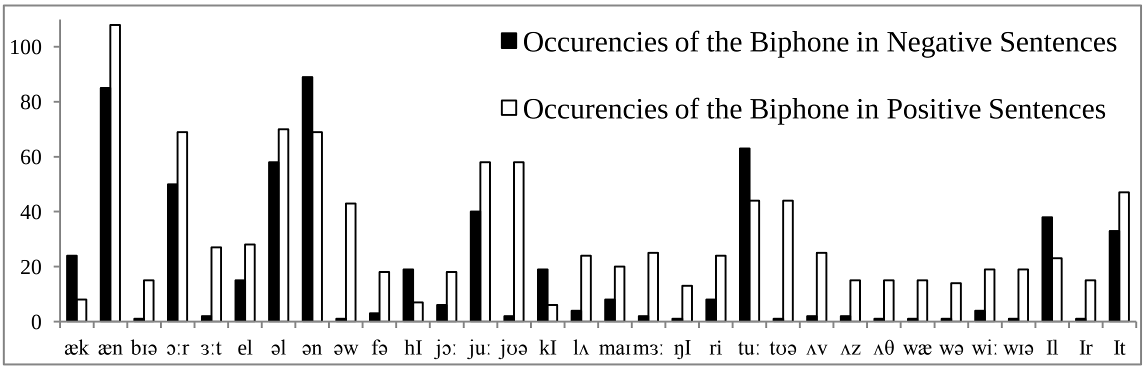

Figure 5. Frequencies of biphones in positive and negative sentences from EmoBank.

The main principle underlying the used statistical approach is based on applying frequency measures for the participation of each biphone in the valence-evaluated sentences of EmoBank.

To treat statistically the corpus' sentences, they were first filtered. Only sentences with valence-score different than 0 and with high between evaluators agreement (SD of the evaluations < 1) were taken. This filtering provided 6,466 sentences, covering all 7 genres. The filtered sub-sample comprises 88,676 occurrences of transcribed words containing 212,244 occurrences of biphones with 746 distinct biphones, where the occurrences of proper nouns and numerals were intentionally discarded.

Statistical analysis of the relationship between emotional valence and the phonemic content

Further, the sentences from EmoBank were grouped into bigger textual portions to obtain statistically informative sources.

The data observation showed that a textual material should contain around 3300 words to have a reasonable chance of containing a representative subset of the biphones found in the corpus. For composing textual data-sources of that required length, we sorted the sentences according to their valence and then split the sorted list of sentences into fractions [Slavova, 2020]. That resulted in 22 valence layers of approximately 3300 words each, where all the filtered sentences were included (each sentence belonging to only one layer). Further, the sentences from each of these layers are considered as a distinct textual source called an artificial document. The valence of the artificial documents was calculated as the mean of the valence of the sentences from the corresponding valence layer. We discovered that the valence of the artificial documents is well fitted by a linear function (see [Slavova, 2021]).

To compare the biphonic content of the artificial documents with their valence, a frequency measure for the biphones - a Biphonic Ratio per Document (RatBiph),

$$RatBiph_{ij} = k.\frac{NBiph_{ij}}{NWord_{j.}.NBiph_{j.}} \quad (1)$$

Where:

$NBiph_{ij}$ is the number of occurrences of biphone i (i=1 to 746) in the artificial document j,

$NWord_j$ is the number of transcribed words in the document j (j = 1 to 22) and $NBiph_j$ is the number of biphones in the syllabic flow of the artificial document j.

As the Biphonic Ratios have very small values, to facilitate the observation of the obtained metadata, they were multiplied by $k=10^6$.

The Biphonic Ratio per Document shows the extent of involvement, per word, of a given biphone in a particular artificial document. Using this frequency measure for the biphones, the 22 artificial documents were presented in a 746-dimensional statistical space – one axis for each biphonic ratio. Further this statistical space is called a biphonic space, in which each document is depicted by its corresponding biphonic ratios.

Many biphonic ratios are highly correlated between them. For example, the ratio of /ɑːd/ is correlated with this of /æf/ at r = 0,486. This phenomenon was expected, as it is a normal phonemic property of any language.

The regression procedure, applied to deal with the multicollinearity in the language data, was Principal Component Regression (PCR) - a technique often used for regressing data with numerous correlated predictors.

Following the procedure of PCR, the first step consisted of dimension reduction using Principal Component Analysis (PCA). The initial 746-dimensional space was reduced (having eigenvalues > 1, where 100% of the variance was extracted) to 21 principal components (PCs).

This reduced space, expressing independent features of the phonological metadata, is further called a PC Biphonic space. The 22 documents' coordinates were recalculated and the documents – “placed” in the obtained PC Biphonic space. Following the standard procedure of Principal Component Regression, the 21 PCs were used as predictors of valence instead of the 764 individual biphonic ratios. The linear regression of valence on the PCs showed that 20 of the 21 PCs have practically no influence on valence. The result showed that the documents' valence is not correlated with 20 of the 21 PCs. Documents' valence displayed a very strong and reliable correlation with one PC2. [Slavova, V. 2021].

The result of this analysis shows a statistically reliable pattern. It makes evident the existence of a relation between the biphonic content of a long text and the valence of the emotion that the text evokes in the reader.

Further, it was tested whether the used sublexical approach gives results that are in agreement with other results in the domain of emotional sound-symbolism.

VALENCE	Biphone	ɑ:b	ɑ:d	ɑ:ð	ɑ:dʒ	ɑ:ʃ	ɑ:f	ɑ:g	ɑ:h	ɑ:k	ɑ:l	ɑ:m	ɑ:n	ɑ:p	ɑ:r	ɑ:s	ɑ:t	ɑ:tʃ	ɑ:v	ɑ:z	ɑ:ʒ	ɑ:θ	ɒb	ɒd	ɒð	ɒdʒ	ɒʃ	æb	æd	æð	ædʒ	æʃ	æf	æg	æk
-1,1734127	ArtDoc01						4,7	0,3		0,3		2,7	2	0,3	15	5,4	1		0,7	1			1,7	1,7				1,7	2,7		0,7	3,3		0,7	19
-0,8732426	ArtDoc02		0,3	1	0,7		2,8			0,3		2,8	1,7		16	3,5	2,1		0,3		0,3	0,3	3,8	2,4		0,3		2,1	2,1		0,7	3,5		1	13
-0,7098639	ArtDoc03		0,3	0,6	0,3		3	0,3		0,9	0,6	1,8	2,4	0,3	16	3,6	0,9			0,3			3,3	0,9		0,9		0,9	3,9	0,3	0,3	3	0,9	0,6	11
-0,5999999	ArtDoc04	0,3		0,6		0,3	3,6			0,6	0,3	1	2,9	0,3	15	3,9	0,6			0,6	0,3		1,6	1	0,3		0,3	1	3,2		0,6	2,6	0,3	0,3	8,1
-0,4526240	ArtDoc05		1,8	0,5	0,3		2,3	1,3			0,5	0,3	1,5		15	4,6							2	1		0,3	0,3	1	1,3		1	1	0,5	1	11
-0,4000001	ArtDoc06		0,2	0,2	0,5		2,4			0,3	0,2	1,4	2,4	0,2	12	3,1	0,9		0,2			0,2	2,1	2,1	0,2	0,2		1	1,9	0,2		1,9	0,3	0,3	9,1
-0,3843538	ArtDoc07		0,3				1,7	0,3		1	0,7	0,7	2,7	0,3	16	4,1	1,4			0,3		0,3	2,4	1,7		0,3		1,7	3,1		0,7	3,8		0,3	8,2
-0,2265954	ArtDoc08		0,2	0,2	0,6		2,3			0,6	0,2	0,5	1,3		11	2,6	0,5	0,2		0,6			1,9	0,3	0,3	0,3		0,8	1,1		0,2	1,9	0,2	0,5	7,5
-0,2000000	ArtDoc09		0,4	0,4	0,2		0,4	0,2				1,3	1,9		15	3,6	0,8						2,5	2,3		0,2		1	2,1		0,2	1	0,2	0,6	8,6
-0,2000000	ArtDoc10					0,3	4,2	0,3		0,3	0,8	0,6	3,9		15	6,7	0,6			0,3		0,3	0,6	1,7		0,6		1,4	1,1	0,6	0,6	1,4	0,3	1,7	11
-0,1997733	ArtDoc11		0,4		0,4		4,8	1,1		0,7		2,6	3,3	0,4	14	4,8	1,1	0,4		0,4	0,4	0,4	2,2	2,6		0,4	0,4	0,7	1,8	0,4	0,7	2,2	0,7	0,7	16
0,1704406	ArtDoc12		0,2	0,5	0,6		1,9	0,2		0,2	0,5	1,1	2,2	0,2	13	2,4	1,6					0,2	1,6	2,1		0,3	0,2	0,6	1	0,6	0,3	0,3		1,1	7,9
0,2000000	ArtDoc13		0,4	0,2	0,4		2,4	0,2	0,2	0,2		0,7	3,4		13	4,5	1,3		0,2				1,5	1,5		0,7		2,1	1,3	0,4	0,7	0,7	0,2	0,4	6,7
0,2000000	ArtDoc14			1,7	0,3		1,7	0,3		0,3		0,7	2,8	0,3	17	5,5	1						2,1	2,1				1,4	3,1	0,3	0,3	1	0,3		11
0,2000000	ArtDoc15		0,8		0,8		2,3	0,3		1,3	0,3	1	2	0,3	16	3,6	2,3			0,8	0,3	0,3	1,8	2,3		0,8		1,8	1,8			4,1	0,5		7,1
0,2020409	ArtDoc16		1,5	0,9	0,6		1,5			0,9		2,6	2,4	0,3	23	3,5	2,6	0,6	0,3	1,2			0,6	2,1		0,3		1,2	2,1		0,6	1,2	0,3	1,2	11
0,3529479	ArtDoc17				0,5		2			0,7		0,7	2,5		15	3	1,5	0,2	0,2				1,5	1,2		0,2		1	3,2		0,5	2,7	0,5	1,5	8,5
0,4000001	ArtDoc18			0,2	0,7		2,2			0,7	0,2	1,3	2,6		15	1,7	0,7			0,7	0,2		0,9	1,3		0,4		1,1	2,6	0,4	0,4	0,9	0,2	0,7	8,1
0,4289116	ArtDoc19		0,9		0,5		2,7			1,4		0,5	4,1	0,2	16	3	1,1						1,4	1,6		0,2		1,1	1,1	0,5	0,5	1,8	0,5	0,7	11
0,5901360	ArtDoc20		1	0,2	1		4,5	0,5				0,5	4,7		16	3,2	0,5		0,2	0,2			2	2		0,5		2,5	2,7		0,2	2,2	0,2	0,5	6,7
0,7316893	ArtDoc21		2,2		0,3		2,7			0,8		0,8	4,1		15	2,7	2,5			0,5		0,3	1,1	1,1		0,8		1,1	2,5		1,4	1,9	0,5	1,4	10
1,0296928	ArtDoc22				1,2		3,7	0,4		0,4		2,1	2,1		23	5,8	1,7			0,4			1,2	2,5		0,4	0,4	2,9	1,7	0,8	1,2	2,5	0,4	2,1	6,6

Figure 6. Part of the statistical space of the biphonic ratios RatBiph

Adelman and colleagues, based on analysis of emotion-evaluated vocabularies of English, German, Dutch, Spanish, and Polish, reported that the first phoneme in the words has particular importance for the word’s emotional charge. Using the approach and procedure described here, we checked whether this positional effect concerns the first biphone too. The biphones placed on the first position of the words (a subset of 532 biphones out of the 746 biphones in the complete words) were taken and the obtained first-biphones’ statistical space was analyzed. The dimension reduction of the “first-biphone” space (21 PCs, 100% variance extracted) revealed that the First PC of this space is strongly correlated (Person’s $r = -0{,}918$, $p < 0.001$) with the valence assessed by EmoBank’s readers.

This result is in agreement with the effect of the first phoneme, reported by Adelman and colleagues. It should be noted that the PC1 based on the first biphones is very strongly correlated (Person’s $r = 0.938$, $p < 0.001$) with the PC2 of the complete biphonic space. The explanation of this property of the first phonemic position requires further attentive analysis.

The result of the applied principle component regression makes evident the existence of a strong relationship between the phonemic content of a valence-homogeneous text (about 1.5 pages long) and the emotional valence that it inspires in the reader.

Thus, the phonemic content of the communication can be used as a relevant characteristic of the emotional valence, embedded in the language communication.

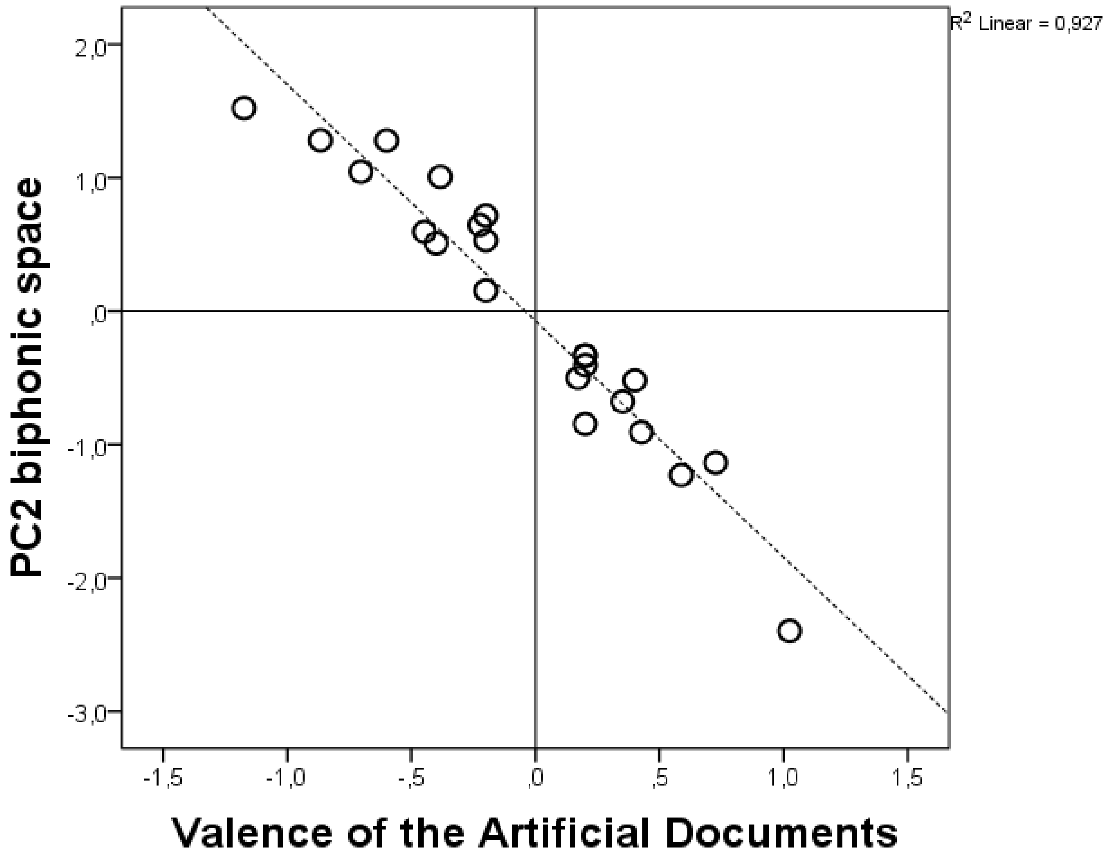

Figure 7. The Pearson’s correlation between Valence and PC2 is r = - 0.961 (p < 0.001).

Exporting the dependencies out of the corpus

Our effort was focused on searching dependencies between the biphonic content and the emotional valence in our corpus of emotionally annotated texts. We found that there is a strong correlation within the corpus. A required further step is to confirm the validity of findings outside the corpus. We did this by applying the calculated dependencies to a big variety of texts outside the corpus

[Slavova & Andonov 2022]. There are several differences between the real world texts and the artificially created documents from the corpus. The sentences in EmoBank in every artificial document are homogeneous in valence, but do not represent a coherent text. The real world texts on the other hand are rarely homogenous in valence on the scale of several pages, as the narrative changes.

To apply the relatedness between biphonic content and emotional valence parameters found in EmoBank outside of it, a “weight” WBiph was attributed to each biphone, that we called a “biphonic weight”. The biphonic weight WBiph was obtained as a correlation coefficient (Pearson) of the interdependence between the participation scores of the biphone and the valence of portions of evaluated sentences (see [Slavova 2020]). The biphonic weight indicates whether the biphone is related to positive or negative valence and shows the strength of this relation.

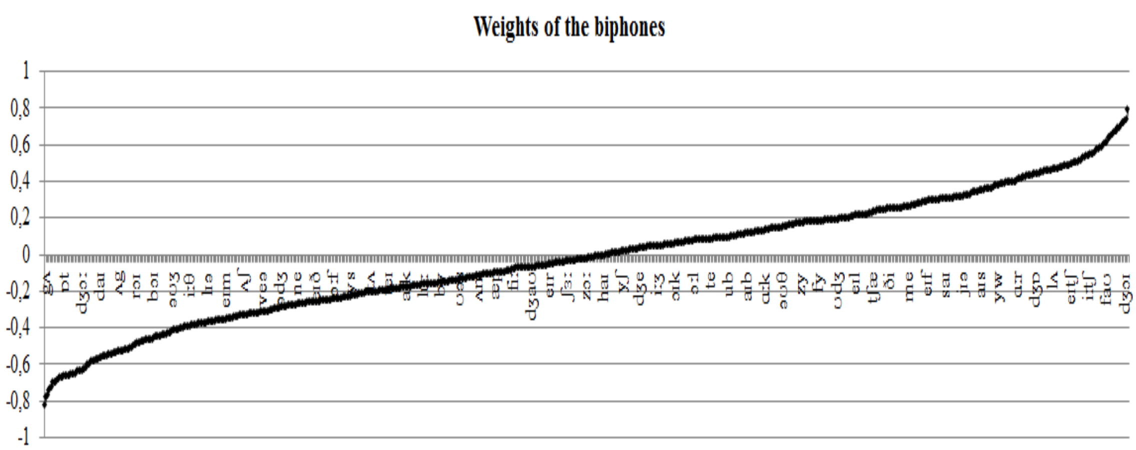

Figure 8. Weights WBiphi of the biphones

Assembling test samples

We performed a statistical analysis based on internet news texts (Fig. 2) in two separate samples. In the first sample (A) the articles were manually selected for inclusion based on the presence of keywords with known emotional valence. We used the 100 most positively (e.g. love, peace, happiness) and 100 most negatively (e.g. horrible, kidnapped, error) charged words from Glasgow Norms [Scott et al, 2018] and NoVAD [Warriner, Kuperman & Brysbaert, 2013] dictionaries. A 38% of the texts were evaluated by 7 participants proficient in English in order to check if the presence of a single keyword in the headline is a reliable predictor of the overall emotional valence of the text. We found 96.8% fit between the emotional keyword in the headline and the readers' assessments and thus we concluded that the presence of an emotional keyword in the headline is reliable and this list of emotional keywords can be used for automatic retrieval of positive and negative news items.

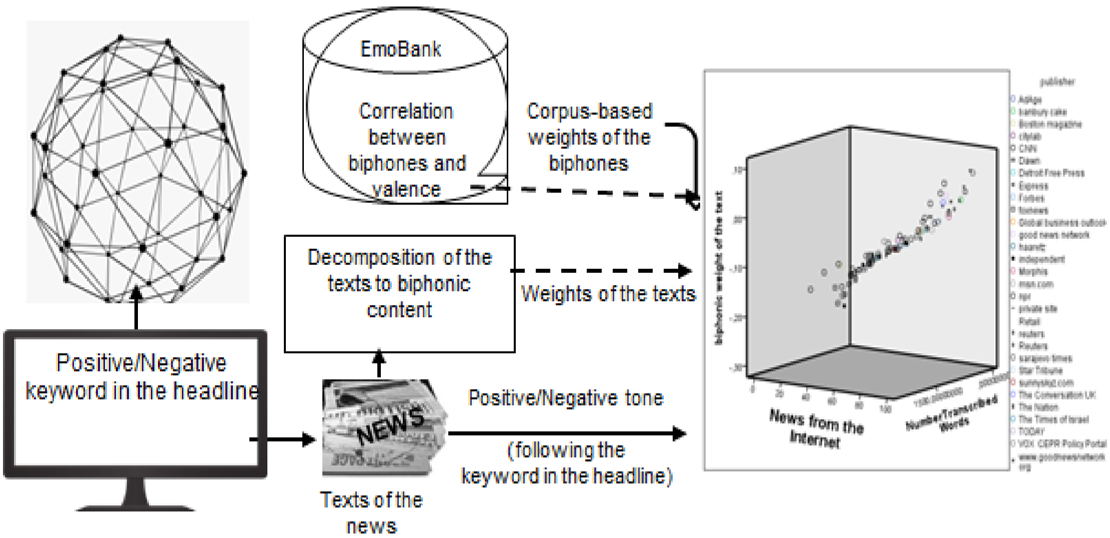

Figure 9. General scheme of the test of the dependencies outside the corpus

Sample B contains automatically downloaded news items using the same list of keywords. A hundred news items published on the Internet were downloaded during three distinct one-week periods in 2021. Covered topics include crime and justice, environment, business, economy, sports, politics and society, science, etc. The texts were assigned to either positive or negative class depending on the keyword in the headline.

Next the texts of the news items from both samples were decomposed to obtain their biphonic content (the occurrences of proper nouns and numerals were intentionally discarded). We had 50,259 occurrences of biphones in Sample A and 175,312 occurrences of biphones in Sample B. To evaluate the indicated by the biphonic content Valence, we used the following measure:

$$WENews_j = \frac{\sum_{i=1}^{746} \quad .n_i.WBiph_i}{NWord_j} \tag{2}$$

Where:

WNews is the biphonic weight of the news item j;

$WBiph_i$ is the weight of the biphone i ;

$NWord_j$ is the number of words in news item j.

Here is shown the interdependency between the phonemic characteristics of Sample B and the emotional tone of the texts selected by keyword and evaluated by subjects. The median weight derived from the corpus splits the series well in correspondence with the assessed positive and negative emotional tones (Fig. 10).

In Sample A the correlation between the phonemic weight of the series of texts and their positive-negative valence (dummy coded) is $r = 0.565$ ($p<0.001$). For Sample B we applied a regression procedure to the weights $WNews_j$ of the texts, where the dependent variable was “tone of the headline” (dummy coded). The phonemic weights of the texts in Sample B correlate with the emotional tone at $r = 0.625$, $p< 0.0001$.

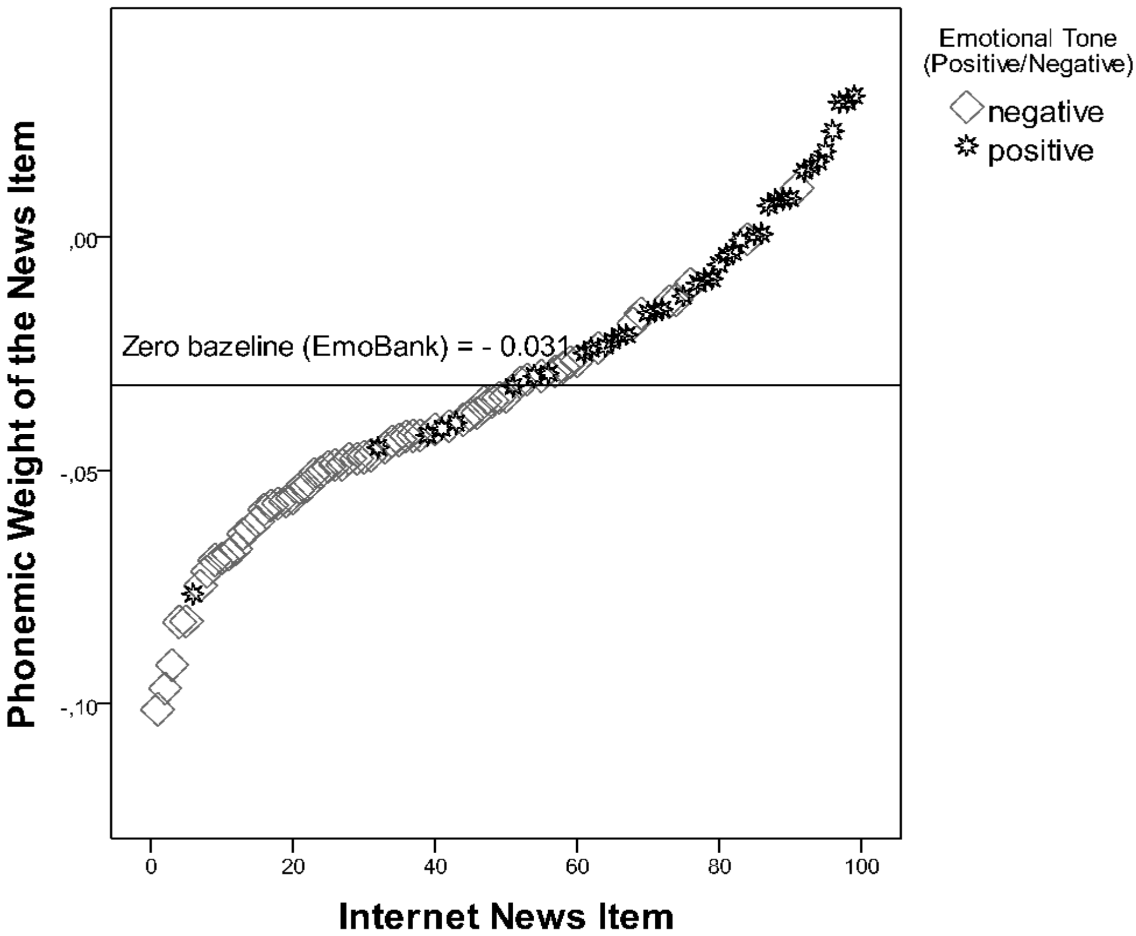

Figure 10. Phonemic weights of the texts from Sample B and their emotional tone as announced in the headline.

We found that the histogram of the statistical variable WNews is bimodal. We tested whether the factor causing this bimodality is the positive or negative tone. The weights WNews showed a clear split in two statistical groups with different Means for the two sub-samples (Mean positive = –0.09 and Mean negative = –0.46). The performed ANOVA (with Levene test for homogeneity of variances), gave definitive result: Mean squares was 0.001 within the groups, 0.032 between the groups, with $F = 62.01$, $p < 0.0001$.

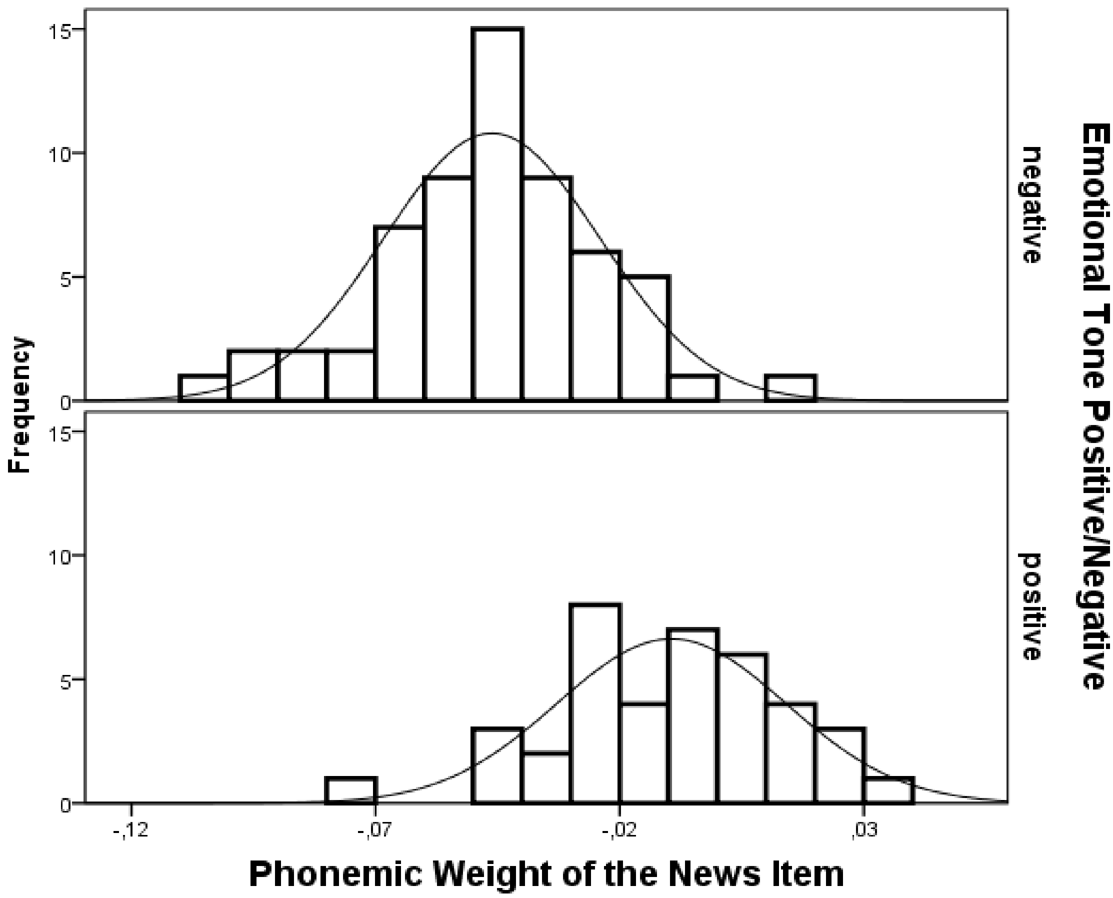

Figure 11. Histograms - the weights of the positive and the negative subsemples in Sample B

Conclusion

The statistical results show that the phonemic content of real-world texts displays clear association with emotional valence. We are not aware of an explanation of the observed results. If we consider that there exist two sources of information concerning emotional valence - the meaning of the message and the entire phonemic content structured in some chosen frame of discourse our finding shows that these two channels transmit similar emotional information. Thus the phonemic content of the words evolved reflecting the emotional component of the meaning that they possess.

From this result, it is not possible to explain according to what principle the observed phonetic type of emotion transfer has entered the language content. In all cases, texts' biphonic decomposition shows such a property.

As the texts in the artificial documents are composed of sentences, one may expect that the emotional valence perceived by the readers is stimulated by the sentences' meanings.

That suggests that the set of all the words used in the sentences of lengthy valence-homogeneous texts is not with arbitrary phonemic content, but with a phonemic content that reflects the involved emotional valence.

Researchers in the domain of robotics started to propose and implement systems based on phoneme occurrence rates and to use them as the input features for classifiers.

Recent results confirm the significance of the phonemic features - the level of accuracy when using such an approach is comparable to that of current stateof-the-art methods and, importantly, the required training time is significantly reduced from hours to minutes.

Bibliography

[Adelman et al, 2018] Adelman, J. S., Estes, Z., & Cossu, M. Emotional sound symbolism: Languages rapidly signal valence via phonemes. Cognition, 175, 122-130. https://doi.org/10.1016/j.cognition.2018.02.007

[Andonov and Slavova, 2019] Andonov F. & Slavova V. Technical aspects of the extraction of phonological features for emotion recognition in texts, International Journal Information Content & Processing, Volume 6, Number 2, 2019, p-p 21-31.

[Aryani et al, 2016] Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A. M., & Conrad, M. Measuring the basic affective tone of poems via phonological saliency and iconicity. Psychology of Aesthetics, Creativity, and the Arts, 10(2), 191. https://doi.org/10.1037/aca0000033

[Buechel and Hahn, 2022] Buechel, S., & Hahn, U. Emobank: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. arXiv preprint arXiv:2205.01996.

[Davis and Olson, 1985] Davis, G.B., and M. Olson. Management information systems: Conceptual foundations, methods and development. McGraw-Hill, New York. 1985.

[Ekman, 1999] Ekman, P. Basic emotions. Handbook of Cognition and Emotion, 1999, 45-60. https://doi.org/10.1002/0470013494.ch3

[Kawahara and Shinohara, 2012] Kawahara, S., & Shinohara, K. A tripartite trans-modal relationship among sounds, shapes and emotions: A case of abrupt modulation. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 34, No. 34). https://escholarship.org/uc/item/47b452vw

[Russel and Mehrabian, 1977] Russell, J. A., & Mehrabian, A. Evidence for a three-factor theory of emotions. Journal of research in Personality, 1977, 11(3), 273-294. https://doi.org/10.1016/0092-6566(77)90037-X

[Scott et al, 2018] Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. The Glasgow Norms: Ratings of 5,500 words on nine scales. Behavior research methods, 2018, 1-13. https://doi.org/10.31234/osf.io/akzyx

[Slavova and Andonov, 2021] Slavova, V., & Andonov, F. How deeply are emotions encoded in language communication and is this detectable in text. International Journal “Information Theories and Applications, 1(3), 271-299. https://doi.org/10.54521/ijita28-03-p04

[Slavova and Andonov, 2022] Slavova, V., & Andonov, F. Bad news or good news when recognizing emotional valence using phonemic content. In 2022 21st International Symposium INFOTEH-JAHORINA (INFOTEH) (pp. 1-6). IEEE. https://doi.org/10.1109/INFOTEH53737.2022.9751339

[Slavova, 2019] Slavova, V. Towards emotion recognition in texts–a sound-symbolic experiment. International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 7(2), 41-51 https://doi.org/10.5937/IJCRSEE1902041S

[Slavova, 2020] Slavova, V. Emotional valence coded in the phonemic content– Statistical evidence based on corpus analysis. Cybernetics and Information Technologies, 20(2), 3-21. https://doi.org/10.2478/cait-2020-0012

[Slavova, 2021] Slavova, V. On the revealing the emotional valence in communication by text. Procedia Computer Science, 192, 1514-1523.

[Warriner et al, 2013] Warriner, A. B., Kuperman, V., & Brysbaert, M. Norms of valence, arousal, and dominance for 13,915 English lemmas. Behavior research methods, 2013, 45(4), 1191-1207. https://doi.org/10.1016/j.procs.2021.08.155

Reference to online resources:

[EmoBank]: GitHub. https://github.com/JULIELab/EmoBank

Authors' Information

Velina Slavova *– New Bulgarian University, Bulgaria; e-mail: vslavova@nbu.bg*

Major Fields of Scientific Research:

Filip Andonov *– New Bulgarian University, Bulgaria; e-mail: fandonov@nbu.bg*

Major Fields of Scientific Research: Decision support systems, computer linguistics