

## Emotional Valence Coded in the Phonemic Content – Statistical Evidence Based on Corpus Analysis

*Velina Slavova*

*New Bulgarian University, Department of Informatics, 21 Montevideo Str., Sofia, Bulgaria*

*E-mail: vaslavova@nbu.bg*

**Abstract:** *This study investigates the relationship between the phonemic content of texts in English and the emotional valence they inspire. The sublexical content is presented in terms of biphones composed by one vowel and one consonant. The statistical analysis of a vast corpus of emotionally evaluated sentences reveals a strong correlation between this sublexical presentation and the evaluations of valence provided by the readers. An initial test performed with other valence-rated prose texts makes believing that the feature observed within the corpus can be useful for the emotion classification of texts.*

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

### 1. Hypothesis and goal

The work presented here explores the hypothesis that a relationship between the phonemic content of the language and emotional valence exists. This is related to a phenomenon known as phonological iconicity.

There is mounting evidence of non-arbitrary sound-symbolic phonetic patterns in language. The known *bouba/kiki* effect of non-arbitrary mapping between the speech sounds and the visual shape of objects, investigated by Köhler [21] in 1929 was the first that questioned the theory of linguistics signs proposed by Saussure [11], suggesting that the words have arbitrary phonetic content. Since then, the most intensively studied sound-symbolic phenomenon is phonological iconicity, i.e., vocally emulating referents' physical features such as shape and dimension. During the last years, iconicity is intensively studied to examine the cognitive mechanisms that underlie the relationship between the phonetic content and the meaning of the words (e.g., [14, 17, 19, 29, 37], and many other studies). Concerning the universal character of the sound-symbolic features, a recent experimental study by D'Anselmo et al. [10] showed that there exist sound-symbolic phonological features that are language-independent. It is to be pointed out that the great majority of examples of sound-symbolic properties appear not to be consciously noticeable – their relatedness to words' meaning has been discovered by means of statistical data-analysis – often of large-scale language corpora (see [12] for an overview).

The connection between sound symbolic features and emotions has not been so intensively investigated but also starts to gain attention. The fact that emotions represent a fundamental parameter of human nature has not been rejected until now. However, much is unclear concerning the relationship between the mechanisms underlying emotions and the cognitive abilities related to conceptualization and meaning. The results of the work presented here are in support of the hypothesis that emotions and meaning are deeply related.

There exist two main models of emotions – a discrete model of emotion categories and a continuous model of emotion dimensions. Emotion categories, such as anger, disgust, fear, happiness, sadness, and surprise, as defined in the works of Ekman [15], have become accepted by the greater scientific community and as a foundation for the development of emotion recognition systems. The other largely accepted model of emotions presents them in a continuous space [26], usually described as referencing three dimensions: Valence, Arousal, and Dominance (VAD). Valence indicates whether emotions are pleasant (positive) or unpleasant (negative), arousal, the degree to which they are exciting, and dominance is a rating of one's own status in relation to an emotion-causing event. Although less directly applicable in the domain of affective computing than the discrete model, the VAD model, as more general, shows a number of advantages and has become largely used in machine emotion-recognition practices and in the domain of sentiment analysis. That prompted several efforts to transform data annotated with emotional categories to VAD-scaled data (e.g., [6, 9]).

Concerning emotions and their correlates in terms of brain activity, several studies based on brain-imaging techniques identify clear neural signatures of emotions involved in the brain's multimodal fashion. Numerous pieces of evidence from functional Magnetic Resonance Imaging (fMRI) show the existence of specific brain regions activated only by unpleasant words and regions activated only by pleasant words (e.g., [22]). A recent study [23] of brain-impaired subjects has shown that both emotional valence and basic emotions are related to semantic memory, that is – to conceptual meaning. Such results lead to suppose that some indicators of emotional valence appear in the “meaning-communication protocol” – that is – they are detectable in speaking, listening, writing, and reading. The questions related to the speech prosody are approximately well studied and intensively used in the systems for speech emotion recognition. It is, however, difficult to analyze how emotion is conveyed through written texts, as there are no prosodic features incorporated, but words alone (with meaning).

The statistical analysis presented here 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 gains automatically its corresponding phonetic content [1, 24, 34]. Recognition of lexical meaning proceeds largely in accord with the processing of presented spoken word-sequences – that is, as perceived speech.

It was hypothesized in my previous works [31, 32] that emotions are deeply incorporated in the cognitive mechanisms of conceptualizing the world. On the

presumption that the emotional dimension of valence has a neuropsychological basis that is associated with lexical semantics and thus influences the words phonemic forms, I undertook an exploration of the relation between the sublexical phonemic level of the English language and the emotional dimension Valence. Next in this paper, I investigate the hypothesis that the phonemic content of the language incorporates features related to valence.

## 2. Investigating the sublexical level

The systematic study of the phonological effects on emotion originates from the Russian Formalists where the sound in poetry was first examined in a systematic fashion (see, e.g., [18, 28]). Further investigations of English poetry and prose genres led to the detection of relationships between single phonemes and readers' experienced emotional state [36]. Recently, Adelman and colleagues [2] conducted data-analysis of words from English, German, Dutch, Spanish and Polish emotion evaluated corpora. They discovered, for all these languages, that correlations exist between the phonological properties of individual phonemes (depending, too, on the phoneme's position within the words) and the valence of these words.

Recent text-emotion studies have found that the sublexical level of German poetic texts is a reliable source to predict their emotional effects and aesthetic qualities. In these studies, Ariani, Ullrich, and colleagues [3, 4, 33] examine the language phonological content presented into layers, applying the reasoning of the Gestalt psychology. The Gestalt principles applied in literary studies may shortly be expressed as follows: a text is a whole-body composed by gradual phonemic layers – from phones to syllables, words, sentences, etc. – and all of these compositional layers have an influence on the perception of the reader. At the sublexical level investigated by Ariani and colleagues, both single phonemes and sub-syllabic segments – onset, nucleus, and coda are shown to provide information that is relevant for the investigation of emotion-related phonological iconicity.

The exploration of the emotion-effects of syllables composed of two phonemes reveals important findings. Kawahara and Shinohara [20] have 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 has demonstrated a tripartite trans-modal sound-symbolic relationship between auditory speech-sounds, visual shapes, and emotions. Having in mind this result, I have explored [30] the relationship between the analogically defined sublexical level of texts in natural language (English) and valence. The investigated phoneme couplings were of the inverse type – vowel-consonant (v-c). Two phonological types of v-c have been investigated – vowel-abrupt (e.g., /ot/) and vowel-sonorant (e.g., /i:m/). The statistical result has shown that the frequencies in the texts of subsets of both types of v-c phoneme pairings are correlated with valence. The analysis has revealed that the vowel-consonant couples provide a sublexical source for estimating valence, but as their density in the texts is quite poor, a larger set of sublexical elements has to be investigated to better cover the texts' phonemic content.

The approach used in the present study is in alignment with the Gestalt reasoning applied in literary studies. The phonemic content of the text is thought in layers, supposing that the phonological forms perceived at these layers have an emotional impact on the reader. I reasoned that if the phonemic coding of emotion exists and contributes to communicating, the sublexical layer that transmits emotion-related information should be embodied in easily pronounceable and well-perceivable phonemic structures – such as consonants and vowels in couples. Having the results regarding the vowel and consonant phoneme pairs mentioned above, in this study, I investigate a sublexical layer consisting of *biphones* – that is, pairings of phones composed of one vowel and one consonant, in either order.

To investigate the complete sublexical layer of biphones, models for all the possible biphones have been obtained using the full set of phonemes as defined in the English language – the set *V* of vowels (/æ/, /ɒ/, /ʌ/, /ʊ/, /i:/, etc., including the diphthongs /aɪ/, /aʊ/, /eɪ/, /eə/, etc.) and the set *C* of consonants (/ŋ/, /d/, /k/, /g/, /tʃ/, /dʒ/, etc.). The set of biphone-models composed by one vowel ( $v \in V$ ) and one consonant ( $c \in C$ ) was obtained by the Cartesian products  $V \times C$  and  $C \times V$ . That provided the 1056 possible combinations of vowels and consonant of the two order-types – 1. v-c: /a:dʒ/, /əʊw/, /ʌk/, etc., and 2. c-v: /ʃu:/, /mɔ:/, /hæ/, etc.

The words' sublexical representation is further considered as composed by such biphones, as shown in Fig.1.

Word	<i>love</i>	<i>hate</i>	<i>which</i>	<i>abhor</i>	<i>something</i>	<i>afraid</i>	<i>surely</i>
Phones	lʌv	heit	wɪtʃ	əb'hɔ:(r)	'sʌmθɪŋ	ə'freɪd	'ʃʊ:(r)li
Biphones	lʌ   ʌv   heɪ   eit	wɪ   ɪtʃ   əb   hɔ:   ɔ:r	sʌ   ʌm   θɪ   ɪŋ	əf   reɪ   eɪd   ʃʊ:   ɔ:r   li			

Fig. 1. Examples of words' phonemic content presented as composed by biphones

The sublexical layer being investigated is not a result of syllabification as the process of obtaining assemblies of phonemes by decomposing the words' content follows the linguistics' rules. The representation by means of biphones leads to a more granular layer of phone-gatherings, which are contained in the syllables and sometimes overlap them. It should be noted that the manner of obtaining the layer of biphones is universal, as vowels and consonants exist in all languages independently of the languages' particular phonetic content and their respective rules of syllabification.

Further in this study, the phonemic units that are submitted to statistical analysis are the occurrences of biphones, as present at the sublexical layer of biphones (Fig. 1). The assumption is that each of them has a distinct perceptual effect related to valence.

To investigate the relatedness of this sublexical layer and emotional valence the study proposed here relies on a large corpus of valence-evaluated texts.

### 3. Data description

The emotion-evaluated texts used in this study are from the EmoBank corpus ([6], [8]). EmoBank involves two previously existing corpora – a manually annotated sub-corpus of the American National Corpus, and the data set SemEval-2007, consisting

of news headlines drawn from major newspapers such as the New York Times, CNN, and BBC News. From the standpoint of universality, EmoBank comprises data collected over the past decade and covers numerous genres.

EmoBank contains 10,549 emotion-evaluated sentences (examples in Table 1), extracted from 136 different documents, classified into 7 genres – blog, essays, fiction, letters, newspaper, SemEval, and travel-guides.

Each sentence in EmoBank is evaluated from the point of view of two perspectives: writer perspective and reader perspective, in terms of the three emotion dimensions – valence, arousal, and dominance. For the writer’s perspective, a number of linguistic clues supporting the annotators in their rating decisions are presented. For the reader’s perspective, it is asked what emotion would be evoked in an average reader. Accordingly, writer emotion refers to how someone feels while producing an utterance, whereas reader emotion relates to how someone feels right after reading or hearing this utterance. For each sentence, five annotators generated VAD ratings. The task was available for workers located in the UK, the US, Ireland, Canada, Australia and New Zealand. The data in EmoBank is presented after a filtering procedure for the inter-annotators agreement and the introduction of specific metrics (for details, see [7]). The Standard Deviation (SD) of the assessments is also offered in the corpus.

Table 1. Examples of the text-content and data presented in EmoBank.

ID sentence	Text of the sentence	Valence Reader		Valence Writer	
		Mean	SD	Mean	SD
hotel-california_30738_30790	Adrienne shook her head and made a sound of disgust	1.33	0.47	1.80	0.75
littleshelter2_1003_1081	It hurts us especially now to see a scrawny dog... a wet, starving kitten	1.60	0.49	1.80	0.40
The_Black_Willow_4372_4422	“Damn you,” he cried, “will you never leave me be?”	1.67	0.75	2.00	0.00
112C-L015_1820_1880	We wish you and your family a new year full of joy and love	4.00	0.89	4.60	0.49
112C-L014_1227_1282	We’re proud of our children and all that they overcome	4.20	0.75	3.80	0.75
captured_moments_28753_28863	For a perfect moment, Emil and Tasha and I were one entity, laughing until our lungs hurt	4.60	0.49	4.00	0.82

Further in this work, the Mean of readers’ assessments of valence provided in EmoBank is considered as the Valence-score of the sentence or simply – the valence of the sentence.

To facilitate the intuitive understanding of the data submitted to statistical analysis, I further use a shifted scale of valence such that the neutral mark fits 0, this way presenting the valence axis in positive and negative parts.

#### 4. Obtaining sublexical metadata

In order to represent the sublexical level of the evaluated texts from EmoBank as biphones, the corpus was stored locally and further treated in six main steps, as

illustrated in Fig. 2. After having decomposed the sentences into words, 157,974 word-occurrences were identified in EmoBank.

Next, the phonetic transcriptions of a large vocabulary of words forms, in the form of strings of IPA phoneme symbols, were downloaded using several on-line dictionaries. Some differences in the used IPA symbols necessitated an additional data homogenization to the rules accepted in British English. A manual search was performed using the Oxford learner’s online dictionary for a number of words (about 5000, a big part of which were verb conjugations). Thus, a vocabulary of 16,392 phonetically transcribed words was obtained.

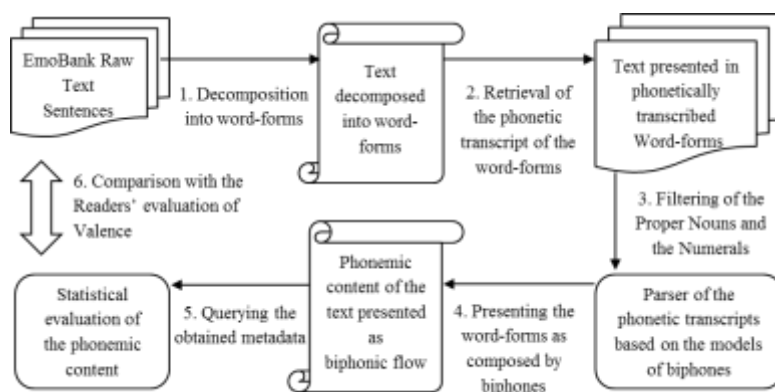


Fig. 2. General scheme of the approach

As a third step, the sentences in the corpus have been checked for word-occurrences representing proper nouns (including e.g. the names of months and abbreviations such as UN) or other words used as proper nouns (e.g., “Times” as in “New York Times” as opposed to “times” in “several times”). Such uses of words were manually retrieved and excluded from the statistical analysis. The numerals (as part of speech – in linguistics terms) written with arithmetical symbols were excluded too, the assumption being that such words are not necessarily chosen by the author and hence do not reflect her emotional intent. After this filtering, 141,950 word occurrences, that is – 90% of EmoBank’s texts’ content, were subjected to further treatment.

The downloaded corpus was locally organized in a relational database. The set of the 1056 models of biphones obtained as explained in Section 2 has been stored in a separate table and used to identify, by the developed parser, the biphones in the phonetic transcription of the words. This showed that the corpus sentences contain 805 different biphones. The used means of data structuring allows performing the obligatory step of data-verification based on comprehensible sentences’ decomposition as shown in Fig. 3.

Sentence ID	WhatToHongKong_3191_3232																	
Words	<i>Many</i>	<i>stores</i>	<i>will</i>	<i>pack</i>	<i>and</i>	<i>ship</i>	<i>purchases</i>											
Transcript	'meni	'sto:z	wɪl	pæk	ænd	ʃɪp	'pɜ:tʃəsɪz											
Biphones	me	en	ni	tɔ:	ɔ:z	wɪ	ɪl	pæ	æk	æn	fɪ	ɪp	pɜ:	ɜ:tʃ	tʃə	əs	sɪ	ɪz

Fig. 3. Example of a sentence from EmoBank as composed by biphones

In this way, the words' phonetic transcriptions were presented in terms of biphones for each word and, consequently – a content of biphones for each sentence. The sublexical units submitted to further counting and density-evaluation in the text-data were the biphones. Further, the phonological image of text-data as a sequence of biphones is called a *biphonic flow*, be that of a single sentence or of the aggregated word-content of any set of sentences.

## 5. Assembling emotion-evaluated texts for statistical analysis

The first step in composing a convenient source of text-data for statistical analysis is the filtering of the sentences. To take only sentences that inspire valence, the sentences with a valence-score zero have been discarded. The second filtering criterion reflects the valence-stimulus effect of the sentences, measured by means of the inter-annotators agreement. The extent of readers' agreement has been accorded in bounds, called agreement thresholds, based on the available in EmoBank value of the SD of the valence-ratings. The filtering gave data for statistical analyses shown in Table 2.

Table 2. Data after filtering the sentences with thresholds on the SD of the readers' ratings

Data submitted to statistical analysis	Readers' agreement SD < 1	Readers' agreement SD < 0.7
Number of sentences	6,466	5,312
Number of occurrences of transcribed words	88,676	72,46
Number of occurrences of biphones	212,244	173,371
Number of distinct biphones	784	775

The statistical investigation that follows is based on frequency measures of the biphones in the filtered sentences. Two problems have arisen: the first is related to the length of the text-data used for the analysis. To represent a reliable source, the text has to be lengthy in order to contain enough (and different) biphones. This suggests the grouping of sentences in bigger volumes of text-data. If the sentences in such an assembled textual source are with nearly equal valence, the “collective valence” of the group can be accorded as the mean of the valence-scores of these sentences.

The second problem is related to the granularity of the “collective valences” of such sentence-groups. Obviously, if the groups are too huge, the granularity of the representation of their valence on the valence axis will be poor, causing trouble for the application of regression and correlation methods.

These two contradictive problems have been solved as follows: I estimated the convenient length for a text on the basis of one word containing on average 1.6 biphones and one sentence containing 7 words, as it is in the sentence provided in Fig. 3. Thus, in theory, the full set of 784 biphones could be contained (without duplicating any) within a sequence of 69 such sentences, i.e., 483 words in total. A “normal” text would need to be roughly 7 times that length, that is – around 3300 words in order to have a reasonable chance of containing a representative subset of the biphones found in the filtered corpus.

For composing texts of that required length, the tactic was sorting the sentences according to their valence and then splitting the sorted sentence-list into fractions called *valence-layers* (Fig. 4a). All filtered sentences are included in these valence-layers and each sentence belongs to only one layer. Further, the sentences from each of these layers are considered as a distinct textual source called *artificial document*. An artificial document comprises sentences that have valence-scores in a given range independently of the genre to which each sentence belongs.

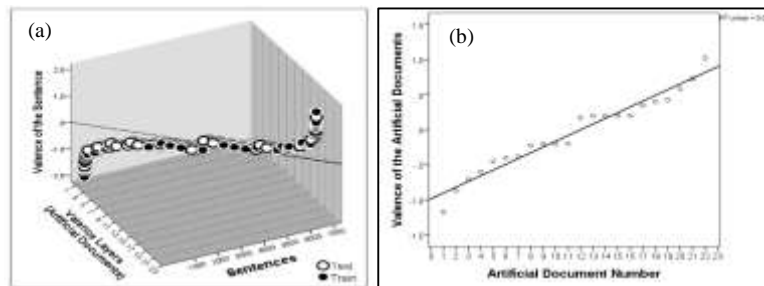


Fig. 4. Manner of composing the artificial documents corresponding to 22 valence-layers (a); the valence of the artificial documents (b)

Next, each artificial document was accorded a measure of valence, called *Document's Valence*, calculated as the mean of the valence-scores of the comprised sentences. The regression analysis shows that the artificial documents' valence is well fitted by a linear function (Fig. 4b) This result meets the objectives, as it allows applying methods of linear regression and correlation analysis.

## 6. The predictive capability of the phonemic content within the corpus

This statistical examination aims to ascertain if the statistical parameters derived from the biphonic flow of texts composed from one-half of the sentences in EmoBank are relevant for texts composed by the other half of the sentences. The corpus have been split into valence-layers and the sentences in each layer have been spread into two equal parts, each seen as distinct artificial documents – one *Train-document* and one *Test-document* (Fig. 4a).

Two parameters have been varied in order to detect factors influencing the interdependency between the biphonic flow and the perceived valence. The first, related to the statistical significance of the text-data, was the length of the artificial documents. Successive statistical trials have been performed based on artificial documents with decreasing length, obtained by splitting the corpus into  $k=12, 16, 18, 20$  and  $22$  valence-layers. That layers were organised into  $N=24, 32, 36, 40,$  and  $44$  artificial documents ( $N=2k$ ). For each occasion of splitting, the sentences were accorded auto-incremented numbers and the Train-documents were assembled by the odd sentences whereas the Test-documents by the even sentences, thus randomizing the participation of each sentence in the test-part or in the train-part of the corpus (Fig. 4a).



For each split of the corpus into layers measures have been calculated, called “*biphonic weights*”, expressing the relatedness of the biphones with valence. This is performed using the following procedure: First, the participation score ( $\text{ScoBiph}_{ij}$ ) of each biphone in each Train-document is calculated using the equation

$$(1) \quad \text{ScoBiph}_{ij} = \frac{\text{NBiph}_{ij}}{\text{NWord}_j},$$

where  $\text{NBiph}_{ij}$  is the Number of occurrences of the Biphone  $i$  in the Train-document  $j$ , and  $\text{NWord}_j$  is the Number of transcribed Words in this document.

Next, the Biphonic Weight ( $\text{WBiph}$ ) for each biphone is calculated in terms of the correlation between its participation scores in the series of the artificial documents calculated following (1) and the corresponding valences of the Train-documents:

$$(2) \quad \text{WBiph}_i = r(\text{ScoBiph}_{i,j=1,k}, \text{ValenceTrainDoc}_{j=1,k}),$$

where  $r$  is Pearson’s  $r$  between the series  $\text{ScoBiph}_{i,j=1,k}$  of the  $k$  participation scores of the biphone  $i$  in the  $k$  Train-documents and  $\text{ValenceTrainDoc}_{j=1,k}$  is the series of the valence of these  $k$  Train-documents.

The correlations of the participation scores of the biphones with valence vary (in the interval  $[-0.9, +0.8]$ ) from very strong to unimportant. Calculated this way, the biphonic weights can, of course, be negative. The biphonic weights express the degree and the direction of the relationship of a given biphone with valence, having that the positive correlation signifies that the presence of the biphone influences the valence in a positive direction.

Next, using the biphonic weights obtained from the Train-documents, I verified on the Test-documents (the other half of the corpus) the predictive capability of the biphones. To do this, I merely applied the weights of the biphones, as obtained from the Train documents, to the biphonic flow of the Test-documents. The obtained for each document sum of biphonic weights, called Document’s weight –  $\text{WDoc}$  has been calculated using the equation

$$(3) \quad \text{WDoc}_j = \frac{\sum_{i=1}^{805} n_i \cdot \text{WBiph}_i}{\text{NWord}_j},$$

where  $\text{WBiph}_i$  is the weight of the biphone  $i$ ,  $n_i$  is the number of times that the biphone  $i$  appears in the biphonic flow of the Test-document  $j$  and  $\text{NWord}_j$  is the number of transcribed words in the Test-document  $j$ .

The Predictive capability ( $\text{Pred}$ ) of the syllabic weights for emotional valence is evaluated in terms of the correlation between the obtained this way weights  $\text{WDoc}_j$  of the Test-documents and their corresponding valence:

$$(4) \quad \text{Pred} = r(\text{WDoc}_{j=1,k}, \text{ValenceTestDoc}_{j=1,k}),$$

where  $r$  is Pearson’s  $r$  between the series  $\text{WDoc}_{j=1,k}$  of the weights of the  $k$  Test-documents’ and the series  $\text{ValenceTestDoc}_{j=1,k}$  of the valences of these  $k$  documents.

The result from the trials show that the predictive capability of the biphonic weights is very high – between 0.84 and 0.97 (Table 3, all  $p$ -values are less than 0.001). In these trials, the longer artificial document (of both types – Train- and Test) consisted of approximately seven pages, and the shorter, of approximately three pages.

Table 3. Results for the predictive capability of the weights of the biphones within the corpus

Corpus divided into N artificial documents:	Valence-agreement between readers: threshold at SD < 0.7		Valence-agreement between readers: threshold at SD < 1.0	
	words per Artificial Document (Avg)	correlation between the Test-documents' weights and Test-documents' valence	words per Artificial Document (Avg)	correlation between the Test-documents' weights and Test-documents' valence
24	2898	0.96	3519	0.97
32	2174	0.91	2575	0.93
36	1932	0.92	2299	0.96
40	1739	0.89	2068	0.93
44	1559	0.84	1873	0.93
	Linear dependence of the prediction on the number of words	0.87	Linear dependence of the prediction on the number of words	0.69

This result makes it evident that the biphonic content of the texts in EmoBank is strongly related to valence.

Table 3 shows the results for the predictive capability of the biphonic weights when varying the two mentioned parameters: 1) readers' agreement on valence, and 2) the length of the artificial documents. As may be seen, in both agreement conditions, the predictive capability depends on the length of the artificial documents. The predictive capability decreases with decreasing the length, showing, unsurprisingly, that when the text-data is reduced, its content becomes less relevant for statistical purposes. The predictive capability of the biphonic weights is more sensitive to the decreasing length of the artificial documents when the sentences are with a higher between-readers agreement.

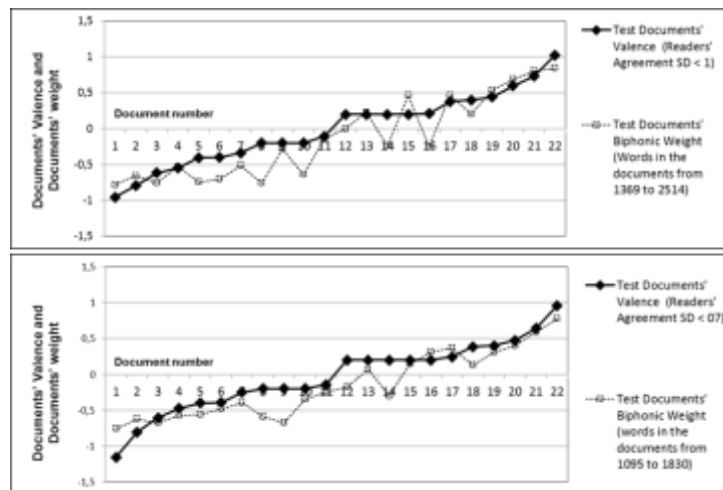


Fig. 5. Plots – results of the check performed with 44 artificial documents (22 Test-documents) in the two readers' agreement conditions

The plots in Fig. 5 show the case of 44 artificial documents (22 Train and 22 Test documents) in the two readers-agreement conditions. When the agreement threshold is stronger (SD < 0.7), the obtained artificial documents are reduced in

length when compared with the weaker agreement (as the number of sentences that correspond to the stronger agreement is smaller). As visible from the plots in Fig. 5, in the case of stronger readers-agreement, the line displayed by the documents' weights better tracks the documents' valence than in the case of long documents with weaker readers-agreement. This suggests that the analysis based on sentences with stronger readers-agreement has provided text-data that contains in a more refined way the sublexical features related to valence.

## 7. Estimating the valence of experimental texts out of corpus

It is known that the feature-values extracted from a corpus are data-sensitive and depend on the corpus content. As explained, EmoBank consists of selected sentences from several originally provided documents in several genres, and, in order to identify some phonemic patterning, the analysis has been based on composed artificial documents. In these artificially assembled documents, there is no meaningful context or any structured linguistic message involved. In contrast, ordinary, real-world texts have the normal structure of a language message and the sentences have a meaning within the text's overall context. Moreover, texts of a narrative nature rarely contain multiple successive pages charged homogeneously with one and the same valence. However, if valence-encoding of biphones is an intrinsic, non-arbitrary aspect of language, then non-arbitrary relationships between biphonic content and valence should be captured, at least partly, in real-world texts.

The aim of the test presented next is to export the obtained corpus-based weights of the biphones out of corpus and to check their validity on real-life texts (Fig. 6). The test is based on a comparison of readers' evaluation of valence conveyed by prose texts and the biphonic content of these texts. The test was conducted within an educational project at the New Bulgarian University. One difficulty concerned the problem of finding long texts that are homogeneous in valence. We designated 20 texts, each about one page long, taken from works by Leigh Bardugo, Charles Dickens, Neil Gaiman, Robin Hobb, Derek Landy, Brandon Sanderson, and J. R. R. Tolkien. They were incorporating various parts – descriptions, dialogues, etc. These texts are named hereafter “experimental texts”. Compared to the artificial documents used for the EmoBank-derived result, the experimental texts were at least three times shorter than the length estimated as statistically reliable – they contain on average 534 words (min 207, max 1016).

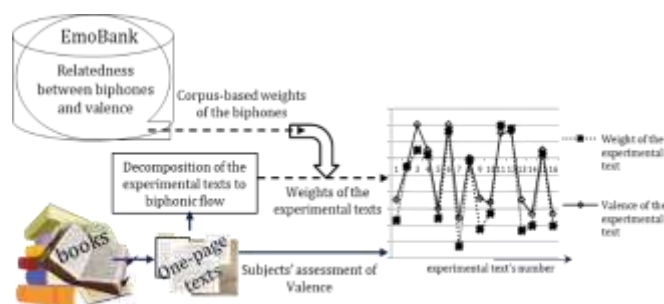


Fig. 6. General scheme of the test

The participants in this small experiment were residents of different countries, were native speakers of different languages (Arabic=2, Bulgarian=10, English=5). For the test presented here, we relied on highly proficient English language users. The non-native speakers of English were professors in English as a Foreign Language (EFL), linguists specialized in EFL, computer scientists with covered high scores of EFL, and students (four) having followed lectures given in English.

The task was to evaluate each text overall and not its parts. We submitted four texts (that we judged as different with regards to valence) to each participant to read silently and evaluate the valence of each. All participants received in written form instruction and evaluated whether the text evokes an emotion. The scale of the assessments was the following ordered set of discrete marks: *Very negative*, *Negative*, *Does not arouse emotion*, *Positive*, *Very positive*.

Obviously, the participants evaluated the valence when simply reading the texts – their attention was not centered on the phonemic content. The assessments were based on the meaning of the texts. The 20 experimental texts were evaluated by our 17 participants, where each text was evaluated by at least 3 subjects (Max 6, Mean 4.7). We received 89 marks that vary from very negative to very positive. The test is based on quite a small sample and the advanced methods of reflecting inter-annotators’ agreement did not apply. In order to work with numeric values, I considered the frequencies (Fig. 7a) of each of the marks that every text obtained as features of the text and applied to the text-to-marks frequency matrix a standard procedure of dimension reduction. Thus, I obtained valence-evaluations expressed as numeric values and further used them as valence-scores of the texts. As seen from the plot in Fig. 7b, the valence-scores of the experimental texts are fitted by a linear function.

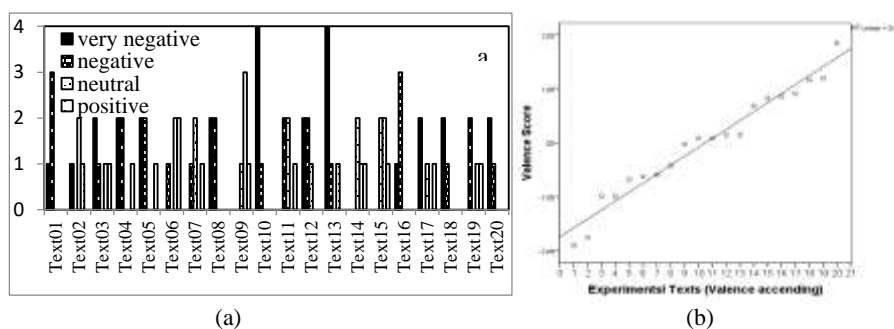


Fig. 7. Frequencies of the marks of the experimental texts (a); Valence scores of the experimental texts (b)

As Fig. 7a illustrates, many experimental texts were rated as invoking very similar valence. I took further the texts for which the subjects agreed, namely – those for which the range of the assigned marks is less or equal to 3 (in the ordered set “very negative”, “negative”, “neutral”, “positive” and “very positive”, Fig. 7a).

Further, the experimental texts were decomposed into their biphonic content. The experimental texts contain 44,557 word occurrences with 39,037 transcribed words (proper nouns and numerals excluded) decomposed to 25,482 biphoneme-occurrences (636 distinct biphones).

To check the validity of the discovered in EmoBank relationship between biphones and valence, a set of “*valence-indicative*” biphones was extracted from the corpus data. More precisely, I considered as valence-indicative 165 biphones (representing 20% of the biphones in the corpus) that had a correlation with p-values less than 0.05 with the valence of the artificial documents, as calculated following (2). From these biphones, 89 were negatively correlated with valence, and 76 – positively correlated (Appendix).

Further, these EmoBank-derived weights of the biphones were applied to the biphonic content of the experimental texts. The weights of the valence-indicative biphones were straightforwardly summed over the content of the experimental texts to calculate their Experimental text’s weights (WEText) using the equation

$$(5) \quad \text{WEText}_j = \frac{\sum_{i=1}^{165} n_i \cdot \text{WBiph}_i}{\text{NWord}_j},$$

where  $\text{WBiph}_i$  is the EmoBank-derived weight of the valence-indicative biphone  $i$ ,  $n_i$  is the number of times that the indicative biphone  $i$  appears in the biphonic flow of the experimental text  $j$  and  $\text{NWord}_j$  is the number of transcribed words in the experimental text  $j$ .

The weights WEText, calculated using the weights of the biphones as derived from EmoBank follow quite well the valence assessed by the subjects in the experiment, as illustrated in Fig. 6.

To evaluate the valence-to-biphone interdependence, I performed a correlation and regression analysis between the experimental texts’ weights calculated following (5) and the valence-scores assessed by the subjects to the experimental texts.

Despite the large difference between the structure of the corpus-based artificial documents and the structure of the experimental texts, and despite the very small length of the experimental texts, the correlation between the experimental texts’ weights and the subjects’ evaluation of valence was important: Pearson’s  $r = 0.73$ ,  $p < 0,02$ .

The plot of the linear regression of the *Weights* of the experimental text on *Valence* is provided in Fig. 8. As illustrated, the result of this exportation of EmoBank derived properties of the biphones has “classified” the experimental texts with 90% accuracy.

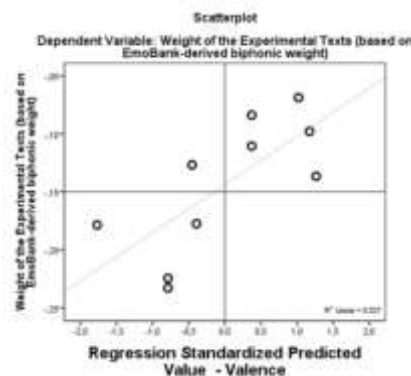


Fig. 8. Relatedness between the weights of the experimental texts calculated using EmoBank-derived weights of valence-indicative biphones and subjects’ assessment of valence

This check was performed for weights of the biphones derived from corpus splits into 22 valence-layers for both SD agreement thresholds ( $SD < 1$  and  $SD < 0.7$ ). The results for the correlation between the weights of the experimental texts and their valence practically do not differ. The regression analysis gave equivalent results (the “classification” plots are quasi-identical). This indicates that the valence-indicative biphones, as designated following a criterion based on a reliable  $p$ -value of their correlation with valence in EmoBank, represent a robust sublexical feature of the available in EmoBank text-data.

While short of perfect, this result is promising with regard to valence-classification tasks. The sample used in this test was too small and it wouldn't be credible to draw general conclusions based on it. However, the three constituents of this test: EmoBank's texts and readers, the experimental texts, and the subjects in the experiment, are completely independent. The common component that could lead to this result is the meaning expressed by the language-coded content of the text-data. This suggests that, at least for the examined texts and population, there is a relation between three components: the texts' meaning, the inspired by the text valence and the properties of the text's biphonic content.

## 8. Word-based estimation of the valence of the experimental texts

Biphones occur in words and words inspire emotion. It is to be checked is the observed biphonic effect related to occurrences of distinct emotionally charged words in the experimental texts. To do this, I used two recently published data sources: 1) the “Glasgow Norms” [27] containing normative ratings for emotion of 5,553 words, and 2) the collected by Warriner and colleagues [35] “Norms of Valence, Arousal, and Dominance” (NVAD), encompassing several data-sources and providing ratings for 13,915 English words. I assigned valence-scores to the words in the experimental texts using these two dictionaries. It should be noted that such dictionaries contain typically singular grammatical forms and, most often – the normal forms of the verbs. For example, the normal form of the verb “to grasp” is provided, but the syllabically decomposed for our analysis verb-form “grasped”, is not. Phenomena related to the morphological components of the word-forms and their corresponding valence are still not well studied, so the two dictionaries were used to find strict word-form's fit – procedures of stemming or detection of morphological building blocks of the words were not applied to enlarge the word's coverage based on words' lemmas. Glasgow Norms covered 2,298 word occurrences in the experimental texts (819 distinct words where 808 intersect the 2,048 biphonically decomposed words) and NVAD covered 2,777 word occurrences in the experimental texts (1,102 distinct words where 1,081 intersect the 2,048 biphonically decomposed words). The biphonically decomposed words represent 88% of the words in the texts (proper nouns and numerals excluded), whereas the dictionaries cover as follows: Glasgow – 21% and NVAD – 26% of the words.

Using the Mean of the words' valence-ratings as they are given in Glasgow Norms and in NVAD, I calculated correspondingly two word-based estimations of

the valence of the experimental texts that were evaluated with subjects' agreement (a range of the marks less or equal to 3). These estimations were obtained as the average of the words' (as occurrences) valence-scores for each text. The two dictionary-based estimations of valence correlate strongly – at 0.85 ( $p < 0.001$ ). Importantly, both estimations did not show relatedness with the valence assessed by the subjects – using Glasgow Norms the correlation is  $r = 0.12$  ( $p = 0.72$ ) and using NVAD:  $r=0.15$  ( $p = 0.67$ ).

Thus, I propose that the features of the biphonic flow provide a considerably more relevant indication about the valence of the entire text, possibly because the biphonic flow covers the phonetic content of the quasi-entire set of the used words.

## 9. Discussion

From this result, it is not possible to explain the principle by which the observed phonemic type of emotion-coded signal has “entered” the lexical language-content. In all cases, the texts' biphonic content shows the described property. To my knowledge, there is no ready-to-hand explanation within the sound-symbolism literature as to the phenomenon reported here. The key question awaiting scientific enlightenment is: how did words come to incorporate these sound-patterns?

It is, obviously, necessary to perform reliably designed large-scale experiments built appropriately in order to examine several crucial points. One is – is the observed phenomenon dependent on the mother tongue of the reader or it is related, as suggested here, to the inferred by him meaning. This could confirm or reject that words' lexicological meaning has an emotional ingredient that has influenced the phonemic content that was attributed to the words during the long process of language evolution. Another important “psycho-phonemic” phenomenon that should be investigated is – do the biphones *per se* inspire some valence. That would necessitate, for example, composing stimuli of pseudo-words containing solely “positive” or “negative” biphones and submit them, as pronounced, to valence-classification by subjects. And, obviously, it should be investigated what are the phonological properties of the biphones that show such a relation with valence and are these properties valid for other languages.

Such steps would contribute to the clarification of the observed phenomenon.

The broad hypothesis to which this result leads should be investigated further in several related domains. Are emotional valence and meaning interrelated in the complex processes of our cognitive functioning, and how? It is well accepted that for social species, the ability to coordinate the actions of multiple individuals involves communicating emotions, including by means of sounds. Recent advances in brain imaging allowed discovering that nonhuman primates, for example, have distinct and clear-shaped brain responses to the vocalizations produced by their conspecifics (e.g., [25]). Although it is still not clear as to the conceptual and emotional meaning of these sounds, the fact that they serve to coordinate the group's actions indicates that these sounds are semantics-discriminative, so they can reasonably be described as “proto-phonemes”. Maybe the signatures of the emotional pattern at all levels of

language would be present if the language has evolved to convey an amalgamation of conceptual meaning and emotion. Then phonemes and their combinations, as lower-level speech units, would have inheritable properties conveying both meaning and emotion. The said capacity could explain the result exposed in this study.

For the purposes of text emotion recognition and sentiment analysis, the described sublexical approach and result need further development, including their possible applications for other languages.

## 10. Conclusions

The result of this statistical investigation shows that emotional information is incorporated in the phonemic content of words and is detectable in written form. The texts are laden with emotional undertone even when their words' lexicological meaning is not taken into consideration and the words are assumed as mere phonetic constructs. This leads to an important form the point of view of cognitive linguistics conclusion – this study's result confirms that the phonemic content of words is not arbitrary from the standpoint of valence.

The valence-to-biphones relation has been detected in texts from multiple genres and includes sources originating from newspapers and travel guides where writers' words are not necessarily chosen to inspire emotion. This leads to a second conclusion, namely that phonemic emotional information exerts its effect at the ordinary level of language communication.

Phonetically conveyed emotion presents, figuratively speaking, a slight signal for statistical purposes and, in order to extract consistent information using it, it is necessary to treat lengthy text-samples composed of valence-homogeneous language data. Narrative texts are rarely of that nature – they do not consist of sentences that enter within a tiny range of valence, but an assembly of sentences that may or may not inspire any valence at all. However, the test with the one-page long prose texts and their evaluation permits believing that the reported phonemic feature is useful in natural language processing – in tasks related to emotion classification of texts and in the domain of sentiment analysis.

**Acknowledgments:** The experimental test described here was conducted at the New Bulgarian University, within a regular course – practical training in internet programming and databases, part of the educational program in computer science. I express my gratitude to my colleague Filip Andonov and to the Ph.D. student Marouane Saula for their participation in the text-processing tasks, as well as to the students Lazar Dilov, Vasil Primov and Steriana Kostova for developing a biphonic parser and organizing the access to the experimental texts. Many thanks to the project-collaborator Richard Traub, who, in the aim to find an explanation of the result, did a supplementary wide-ranging analysis of the domain of emotions, suggested valuable sources and copy-edited early manuscripts of the proposed here text.

I express my gratitude to three unknown reviewers whose valuable suggestions to the description of the study's results dedicated to the domain of cognitive linguistics helped me to formulate the described here essentials of the approach to the data, the statistical method and its outcomes.



## References

1. Abbassi, E., K. Kahlaoui, M. A. Wilson, Y. Joannette. Processing the Emotions in Words: The Complementary Contributions of the Left and Right Hemispheres. – *Cognitive, Affective, & Behavioral Neuroscience*, Vol. **11**, 2011, No 3, pp. 372-385.
2. Adelman, J. S., Z. Estes, M. Cossu. Emotional Sound Symbolism: Languages Rapidly Signal Valence via Phonemes. – *Cognition*, Vol. **175**, 2018, pp. 122-130.
3. Aryani, A., M. Conrad, A. M. Jacobs. Extracting Salient Sublexical Units from Written Texts: “Emophon,” A Corpus-Based Approach to Phonological Iconicity. – *Frontiers in Psychology*, Vol. **4**, 2013, p. 654.
4. Aryani, A., M. Kraxenberger, S. Ullrich, A. M. Jacobs, M. Conrad. Measuring the Basic Affective Tone of Poems via Phonological Saliency and Iconicity. – *Psychology of Aesthetics, Creativity, and the Arts*, Vol. **10**, 2016, No 2.
5. Bradley, M. M., P. J. Lang. Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings (Technical Report No C-1). Gainesville, FL: NIMH Center for Research in Psychophysiology, University of Florida, 1999.
6. Buechel, S., U. Hahn. Emotion Analysis as a Regression Problem – Dimensional Models and Their Implications on Emotion Representation and Metrical Evaluation. – In: *Proc. of 22nd European Conference on Artificial Intelligence*, IOS Press, 2016, pp. 1114-1122.
7. Buechel, S., U. Hahn. Readers vs. Writers vs. Texts: Coping with Different Perspectives of Text Understanding in Emotion Annotation. – In: *Proc. of 11th Linguistic Annotation Workshop*, 2017, pp. 1-12.
8. Buechel, S., U. Hahn. EMOBANK: Studying the Impact of Annotation Perspective and Representation Format on Dimensional Emotion Analysis. – In: *Proc. of 15th Conference of the European Chapter of the Association for Computational Linguistics*, Vol. **2**, Short Papers, 2017, pp. 578-585.
9. Buechel, S., U. Hahn. A Flexible Mapping Scheme for Discrete and Dimensional Emotion Representations: Evidence from Textual Stimuli. – In: *Proc. of 39th Annual Meeting of the Cognitive Science Society (CogSci 2017)*, 2017, pp. 180-185.
10. D’Anselmo, A., G. Prete, P. Zdybek, L. Tommasi, A. Brancucci. Guessing Meaning from Word Sounds of Unfamiliar Languages: A Cross-Cultural Sound Symbolism Study. – In: *Frontiers in Psychology*. Vol. **10**. 2019. 593 p.
11. De Saussure, F. *Course in General Linguistics* (Trans. Roy Harris). London: Duckworth, 1916/1983.
12. Dingemanse, M., D. E. Blasi, G. Lupyan, M. H. Christiansen, P. Monaghan. Arbitrariness, Iconicity, and Systematicity in Language. – *Trends in Cognitive Sciences*, Vol. **19**, 2015, No 10, pp. 603-615.
13. Dolan, R. J. Emotion, Cognition, and Behavior. – *Science*, Vol. **298**, pp. 1191-1194.
14. Edmiston, P., M. Perlman, G. Lupyan. Repeated Imitation Makes Human Vocalizations More Word-Like. – *Proceedings of the Royal Society B: Biological Sciences*, Vol. **285**, 2002, No 1874.
15. Ekman, P. Basic Emotions. *Handbook of Cognition and Emotion*, 1999, pp. 45-60.
16. Fuster, J. *Cortex and Mind*. New York City, Oxford University Press, 2003.
17. Imai, M., S. Kita. The Sound Symbolism Bootstrapping Hypothesis for Language Acquisition and Language Evolution. – *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. **369**, 2014, No 1651.
18. Jakobson, R. *Linguistics and Poetics*. – In: *Style in Language*. MA, MIT Press, 1960, pp. 350-377.
19. Jones, J. M., G. Vigliocco. Iconicity in Word Learning: What Can We Learn from Cross-Situational Learning Experiments? – In: *Proc. of Annual Meeting of the Cognitive Society*, 2017.
20. Kawahara, S., K. Shinohara. A Tripartite Trans-Modal Relationship among Sounds, Shapes and Emotions: A Case of Abrupt Modulation. – In: *Proc. of Annual Meeting of the Cognitive Science Society*, Vol. **34**, 2012, No 34.

21. Köhler, W. *Gestalt Psychology*. New York, NY: Liveright, 1929.
22. Maddock, R. J., A. S. Garrett, M. H. Buonocore. Posterior Cingulate Cortex Activation by Emotional Words: fMRI Evidence from a Valence Decision Task. – *Human Brain Mapping*, Vol. **18**, 2003, No 1, pp. 30-41.
23. Macoir, J., C. Hudon, M. P. Tremblay, R. J. Laforce, M. A. Wilson. The Contribution of Semantic Memory to the Recognition of Basic Emotions and Emotional Valence: Evidence from the Semantic Variant of Primary Progressive Aphasia. – *Social Neuroscience*, 2019, pp. 1-12.
24. Marinkovic, K., R. P. Dhond, A. M. Dale, M. Glessner, V. Carr, E. Halgren. Spatiotemporal Dynamics of Modality-Specific and Supramodal Word Processing. – *Neuron*, Vol. **38**, 2003, No 3, pp. 487-497.
25. Ortiz-Rios, M., P. Kuśmierk, I. DeWitt, D. Archakov, F. A. Azevedo, M. Sams, J. P. Rauschecker. Functional MRI of the Vocalization-Processing Network in the Macaque Brain. – *Frontiers in Neuroscience*, Vol. **9**, 2015.
26. Russell, J. A., A. Mehrabian. Evidence for a Three-Factor Theory of Emotions. – *Journal of Research in Personality*, Vol. **11**, 1977, No 3, pp. 273-294.
27. Scott, G. G., A. Keitel, M. Becirspahic, B. Yao, S. C. Sereno. The Glasgow Norms: Ratings of 5,500 Words on Nine Scales. – In: *Behavior Research Methods*. 2018, pp. 1-13.
28. Shklovsky, V. *Theory of Prose*. Transl. B. Sher. Elmwood Park, IL: Dalkey Archive, 1990.
29. Sidhu, D. M., P. M. Pexman. Five Mechanisms of Sound Symbolic Association. – In: *Psychonomic Bulletin & Review*. Vol. **25**. No 5. 2018, pp. 1619-1643.
30. Slavova, V. Towards Emotion Recognition in Texts – A Sound-Symbolic Experiment. – *International Journal of Cognitive Research in Science, Engineering and Education*, Vol. **7**, 2019, No 2, pp. 41-51.
31. Slavova, V., A. Soschen. On Mental Representations: Language Structure and Meaning Revised. – *International Journal Information Theories & Applications*, Vol. **2**, 2015, No 4, pp. 316-325.
32. Slavova, V., H. Sahli, W. Verhelst. Multimodal Emotion Recognition – More Cognitive Machines. – In: L. Fernando, J. de Mingo Lopez, C. K. Markov, K. Ivanova, I. Mitov, Eds. *New Trends in Intelligent Technologies: Supplement to the Journal Information Technologies and Knowledge*, Sofia, Institute of Information Theories and Applications FOI ITHEA, 2009, pp. 70-78.
33. Ullrich, S., A. Aryani, M. Kraxenberger, A. M. Jacobs, M. Conrad. On the Relation between the General Affective Meaning and the Basic Sublexical, Lexical, and Inter-Lexical Features of Poetic Texts – A Case Study Using 57 Poems of HM Enzensberger. – In: *Frontiers in Psychology*. Vol. **7**. 2017, 2073.
34. Van Orden, G. C., J. C. Johnston, B. L. Hale. Word Identification in Reading Proceeds from Spelling to Sound to Meaning. – *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. **14**, 1988, No 3.
35. Warriner, A. B., V. Kuperman, M. Brysbaert. Norms of Valence, Arousal, and Dominance for 13,915 English Lemmas. – *Behavior Research Methods*, Vol. **45**, 2013, No 4, pp. 1191-1207.
36. Whissell, C. Phonosymbolism and the Emotional Nature of Sounds: Evidence of the Preferential Use of Particular Phonemes in Texts of Differing Emotional Tone. – *Perceptual and Motor Skills*, Vol. **89**, 1999, No 1, pp. 19-48.
37. Winter, B., M. Perlman, L. K. Perry, G. Lupyán. Which Words are Most Iconic?. – *Interaction Studies*, Vol. **18**, 2017, No 3, pp. 443-464.
38. EmoBank Corpus [8]. Internet Resource.  
<https://github.com/JULIELab/EmoBank>

## Appendix. Valence-related biphones

biphone	gʌ	nəʊ	nɒ	hɪ	əʊn	feɪ	ɪd	u:m	ɔ:tf	ʌd	ɔɪd	vəʊ	wɒ
weight	-0.82	-0.78	-0.77	-0.74	-0.73	-0.72	-0.69	-0.69	-0.69	-0.69	-0.68	-0.67	-0.67
biphone	ɒt	leɪ	fiə	vɔɪ	ʌt	ʌdʒ	eθ	wəʊ	bæ	viə	ed	fe	dæ
weight	-0.67	-0.66	-0.66	-0.66	-0.66	-0.66	-0.65	-0.65	-0.65	-0.65	-0.65	-0.64	-0.63
biphone	dʒɔ:	nʌ	pɔɪ	ʃu:	bʌ	vz	dəʊ	æk	kʊ	dʒɜ:	li:	dəʊ	dɪ
weight	-0.63	-0.63	-0.63	-0.62	-0.62	-0.60	-0.60	-0.58	-0.58	-0.58	-0.58	-0.57	-0.57
biphone	dai	ðəʊ	hi:	ɜ:d	ɒb	θɪ	rʌ	kɪ	ɪm	iəd	wu:	ɜ:r	de
weight	-0.57	-0.56	-0.56	-0.55	-0.55	-0.55	-0.55	-0.55	-0.54	-0.54	-0.54	-0.54	-0.53
biphone	ʌg	wə	ði:	sʌ	vai	dʒʌ	lɪ	ʌθ	ðæ	əʊv	pʌ	aʊt	iəl
weight	-0.53	-0.53	-0.53	-0.53	-0.52	-0.52	-0.52	-0.51	-0.51	-0.51	-0.51	-0.51	-0.50
biphone	rɔɪ	ɜ:t	ers	ʌf	haʊ	ðəʊ	ɒr	ɔ:d	ʌz	bɜ:	hu:	ʃʊ	ʊtf
weight	-0.49	-0.48	-0.48	-0.48	-0.48	-0.47	-0.47	-0.47	-0.47	-0.46	-0.46	-0.46	-0.46
biphone	bɔɪ	θʌ	ʃʌ	dʒə	peɪ	ɔɪt	bɪə	və	ɒt	iθ	ʃɑ:		
weight	-0.46	-0.45	-0.45	-0.44	-0.44	-0.44	-0.44	-0.43	-0.43	-0.43	-0.43		
biphone	jɔ:	ʊər	ve	dʒɔɪ	ef	əʊg	gɪ	gʊ	veɪ	u:tf	θæ	pɒ	u:f
weight	0.79	0.74	0.73	0.72	0.72	0.70	0.70	0.70	0.68	0.68	0.67	0.66	0.65
biphone	aɪf	nə	æm	faʊ	ju:	ɑ:dʒ	tʃʊə	el	ɒb	tʃi	ʃə	gɑɪ	tʃɔɪ
weight	0.64	0.64	0.62	0.61	0.60	0.60	0.59	0.58	0.58	0.58	0.57	0.56	0.56
biphone	u:n	zi:	be	i:tf	fʌ	ɜ:v	əʊp	tə	æŋ	vi	ʌv	i:j	ʃɪ
weight	0.55	0.55	0.54	0.54	0.54	0.54	0.53	0.52	0.52	0.51	0.51	0.51	0.51
biphone	ri:	ku:	ɜ:k	ertʃ	dʒu	fæ	ætʃ	ʊəz	ih	nu:	ziə	fə	bə
weight	0.50	0.50	0.50	0.49	0.49	0.49	0.49	0.49	0.48	0.48	0.48	0.48	0.48
biphone	jeə	ɔ:g	wɪ	lʌ	arw	zɜ:	u:d	i:ʃ	u:z	reə	wəʊ	ɪʒ	ʊət
weight	0.47	0.47	0.47	0.47	0.46	0.46	0.46	0.46	0.46	0.45	0.45	0.45	0.45
biphone	æn	tʃə	tʃiə	dʒɒ	ʃai	aɪv	sʊ	ɪv	kɔɪ	dʒæ	ɑ:n		
weight	0.45	0.45	0.44	0.44	0.44	0.43	0.43	0.43	0.43	0.43	0.42		

NB. These biphones and their weights (correlation  $r$  with valence) were obtained by EmoBank corpus-split into 22 layers having a readers-agreement threshold at  $SD < 1$ .

*Received: 18.02.2020; Second Version: 27.04.2020; Accepted: 12.05.2020*