HOW DEEPLY ARE EMOTIONS ENCODED IN LANGUAGE COMMUNICATION AND IS THIS DETECTABLE IN TEXT

Velina Slavova, Filip Andonov

Abstract: *The study aims at exploring the relatedness between the emotional valence (positive-negative emotion) conveyed by conventional texts and their phonemic content. In the proposed approach, the content of a text is considered as composed of layers, assuming that the sub-lexical phonological elements involved contribute to the transmission of an emotional tone. The sub-lexical layer investigated here is composed of "biphones" - pairings of a vowel and a consonant in either order. A previous analysis of metadata derived from a corpus of sentences evaluated for emotion found that the frequencies of the biphones are strongly correlated with emotional valence. The same method of representing the sub-lexical layer was applied to real-world texts. The validity of the relationships derived from the corpus was tested on arbitrarily downloaded texts on diverse topics published by several news agencies. The statistical analysis shows that the derived relationship between emotional valence and the sub-lexical content of biphones is present in structured real-world, ordinary texts. In conclusion: language communication incorporates phonemic patterns that indicate emotional valence. The overall body of a text carries sub-lexical information about its emotional tone. This approach and result are useful for the fields of sentiment analysis and text emotion recognition. The overall result raises the hypothesis that the cognitive mechanisms that shape the phonemic forms used to mentally build a structured language expression reflect emotional valence. The phenomenon needs further investigation whether it depends on the native language of the reader and how, or if at all biphones per se inspire positive or negative emotion in the listener.*

Keywords: *emotional valence; language; sound symbolism; statistical analysis; text emotion recognition; sentiment analysis.*

DOI: *https://doi.org/10.54521/ijita28-03-p04*

1. Introduction

This work investigates the existence of a relationship between the phonemic content of structured language expressions and the emotional valence (positive-negative emotion) of what is expressed.

The fact that emotions are fundamental to human nature has never been contradicted. There are two main models of emotion. The first model consists of discretely mediated categories called “basic emotions” (e.g., anger, fear, happiness, etc., from the listing of Ekman, 1999). This model has long been regarded as intuitively obvious and used as the foundation for developing an emotion recognition technology. The second largely accepted model, proposed by Russel and Mehrabian (1977) and applied in the present study, portrays emotional phenomena according to their discernible attributes, interpreted as emotion-dimensions – Valence, Arousal, and Dominance (VAD). Valence indicates the positive or negative (pleasant or unpleasant) aspect of the emotion, arousal indicates the degree of excitement, and dominance shows one’s status in regard to the emotion-causing event.

Results from different domains show the existence of a biologically underwritten functional layer related to emotions, deeply involved in the cognitive processes. Brain imaging studies allowed identifying clear neural signatures of basic emotions (e.g., Saarimäki et al., 2015), the existence of specific brain regions activated only by unpleasant or only by pleasant words (Maddock, Garrett, and Buonocore, 2003), and results from brain impairments showed that emotional valence and basic emotions are related to semantic memory (Macoir, Hudon, Tremblay, Laforce, & Wilson, 2019). Studies in psycholinguistics report that words for positive and negative emotion tend to be recognized faster (Scott, Keitel, Becirspahic, Yao, & Sereno, 2018), and results in the domain of child development show that emotional valence facilitates the acquisition of abstract words (Ponari, Norbury, and Vigliocco, 2018; Kim, Sidhu & Pexman, 2020). Such findings suggest that valence is associated with levels of cognitive processing in language related to lexical semantics.

The analysis presented next concerns emotional processing in reading. It is known from brain imaging 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 (see Abbassi, Kahlaoui, Wilson, and Joanette, 2011; Marinkovic et al, 2003; Orden, Johnston, and Hale, 1988). Processing lexical meaning at the brain level is similar to hearing sequences of spoken words.

1.1. Language and sound symbolism

A basic assumption of formal linguistics is that the meanings of words and their phonetic characteristics are not related. This assumption, proposed by De Saussure (1916), has been challenged by a diversity of research, providing evidence of non-arbitrary phonological regularities in the languages. One of the best-known examples of this phenomenon, termed phonetic (or sound) symbolism, is described by Köhler (1929) with the known *bouba/kiki* effect of non-arbitrary mapping between the speech sounds and the visual shape of objects. The recent advances in cognitive linguistics, psycholinguistics, and related fields have shown that non-arbitrary mappings between form and meaning are far more widespread than previously believed (e.g., Blasi, Wichmann, Hammarstrom, Stadler, & Christiansen, 2016; Monaghan, Shillcock, Christiansen, and Kirby, 2014). A center of interest of non-arbitrariness is iconicity: the resemblance between form and meaning (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Dingemanse, Perlman, and Perniss, 2020; Sidhu, Vigliocco, and Pexman, 2020 for overviews). Several experiments and studies have shown that adults, children, and pre-linguistic infants are sensitive to sound symbolism (e.g., Asano et al. 2015; Imai and Kita, 2014; Kovic, Plunkett, and Westermann, 2010; Yang, Asano, Kanazawa, Yamaguchi, & Imai, 2019). Such results prompted the hypothesis that language evolution and language acquisition are related to sound symbolic phenomena.

Following the extended analysis presented in Dingemanse et al., (2015), the sound symbolic phenomena are not perceived at a conscious level, but they have been discovered through statistical data analysis of language corpora.

1.2. Emotion and phonology

The phonological effects on emotion were first described by the Russian formalists in the domain of poetry and subsequently investigated systematically (Jakobson, 1960 for an overview). Further studies of Whissell (1999) detected relationships between single phonemes and the experienced emotion conveyed by poetry and prose in English. Recently, Ariani and colleagues performed several text-emotion studies of German poetry (Aryani, Conrad, and Jacobs, 2013; Aryani, Kraxenberger, Ullrich, Jacobs, & Conrad, 2016; Ullrich, Aryani, Kraxenberger, Jacobs, & Conrad, 2017), and found that the sublexical representation contains features that can predict emotional effects and aesthetic qualities. The authors applied the reasoning of Gestalt psychology and examined the text as a body composed of gradual layers – from phones to syllables, words, sentences, etc., and found that these perceived layers influence the reader, consciously or not.

Recently, Adelman, Estes, and Cossu (2018) conducted a data analysis on emotion-evaluated corpora of English, German, Dutch, Spanish, and Polish words. The authors found a correlation between individual phonemes, their position in words, and the emotional valence of the words. The authors suggest that this phenomenon is related to a person's survival in cases of announced danger.

Kawahara and Shinohara (2012) investigated the effect of pairs of consonant-vowel phonemes when listening to pseudo-words and found a sound symbolic phenomenon that relates speech sounds with visual shapes and emotions. Slavova (2019) explored the relationship between valence and analogically defined sub-lexical level based on phoneme couplings of the inverse type (vowel-consonant). The statistical result indicated that these pairings were correlated with valence. Further, to better cover the sub-lexical content of the texts, all types of pairings of vowels and consonants were investigated (Slavova, 2020). The result based on an emotion annotated corpus (shortly presented hereafter) showed a strong relatedness of this sublexical representation with emotional valence.

2. Methods

In this work we explore written language expressions with the assumption that their phonemic content incorporates a "secret code" that conveys the emotion involved in the message. The approach applied here, proposed in Slavova (2019, 2020) follows the Gestalt reasoning according to which the content of a text is seen in layers, assuming that the phonological forms perceived at the sub-lexical layer have an emotional impact on the reader.

The steps which were performed to obtain the result reported here about the relationship between phonemic units and emotional valence were:

1. A phonological unit was chosen – a "biphone" – and the sublexical layer for investigation was presented as composed by biphones;
2. Textual material evaluated for emotional valence was extracted from a huge corpus and was treated to obtain appropriate metadata, and statistically analyzed;
3. The strong dependencies found in the corpus were verified in real-world texts arbitrarily downloaded from the websites of a number of news agencies.

2.1. Approach to the sub-lexical presentation and description of the corpus used to detect the relationship between biphones and valence

If conveying emotional valence through speech plays an important role in communication, some phonemic "markers" of valence should be present in the phonemic stream. All languages have vowels and consonants and all languages use sequences of these. A pair of phones – a vowel and a consonant represents a basic elementary unit of speech – a good candidate for a phonemic marker of emotion. This general reasoning led to the choice of the units used as a methodological device for the investigation. Moreover, the mentioned studies of Kawahara and Shinohara (2012) and Slavova (2019, 2020) have shown that these pairings are related to emotion.

The statistical analysis presented here is based on the manner of sub-lexical representation proposed in Slavova (2020). The set of all pairings of a vowel and a consonant was obtained by the Cartesian products V×C and C×V, where V is the set of all vowels and C is the set of all consonants as they are defined in the English phonetics. That provided all 1056 possible pairings of the two order types – 1. Vowel - consonant: /ɑːdʒ/, /əʊw/, /ʌk/, etc. and 2. Consonant - vowel: /ʃuː/, /mɔː/, /hæ/, etc. Further in this work, a pair consisting of a vowel and a consonant, in either order, is called a *biphone*.

Table 1

Examples of sentences presented in EmoBank with their readers' assessment of valence

Text of the sentence	**Readers Perspective Valence**	
	Mean	**SD**
Adrienne shook her head and made a sound of disgust.	1.33	0.47
This is torture.	1.6	0.49
Stepping into the living room, he said, "A beautiful house."	4.2	0.75
I love what you are doing.	4.4	0.49

The decomposition to biphones leads to a granular layer of phone assemblies which are contained in syllables and on occasion overlap them, as illustrated in Table 2. The units that are submitted to statistical analysis following this method are the occurrences of biphones with the assumption that each biphone has a distinct specific perceptual effect related to valence.

To evaluate the informational load of the biphones with regard to valence, the emotion-evaluated corpus EmoBank (Buechel and Hahn, 2017a,b), annotated with Valence, Arousal, and Dominance was used (Slavova,2020).

The corpus consists of more than 10,000 sentences (examples in Table 1), taken from 136 different documents belonging to 7 genres (Figure 1). The filtering procedures and the specific metrics, applied to construct EmoBank are presented in details in Buechel and Hahn, (2017.a.).

The approach applied for the analysis (Slavova, 2020) was based on an amalgamation of sentences from all seven genres to ensure a universal character of the data.

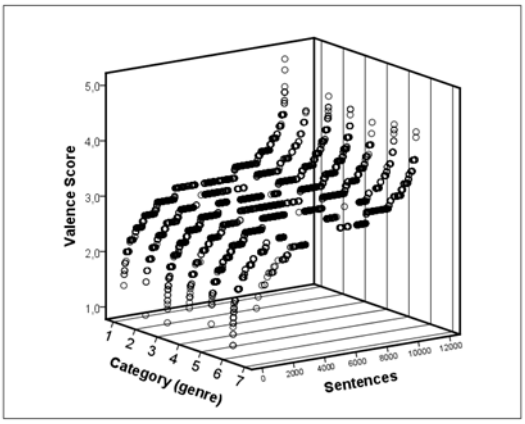


Figure 1. Valence scores of the sentences in EmoBank by category. 1. Blog: 2. Essays; 3. Fiction; 4. Letters; 5. Newspaper articles; 6. Newspaper headlines; 7. Travel guides

2.2. Obtaining phonemic metadata for statistical analysis

Technically the analyzed texts are stored locally and further treated to produce the necessary metadata. The sentences are decomposed into words and the phonetic transcription of each word is retrieved using online dictionaries. This provides a lexicon of transcribed words (about 20,000 different word-forms so far). An example of a decomposed sentence is given in Table 2.

Table 2

Example of a sentence from EmoBank, presented as composed by biphones

Words	Better				still		go	shopping				with		an	expert		
Transcripts	ˈbetə(r)				stɪl		gəʊ	ˈʃɒpɪŋ				wɪð		æn	ˈekspɜː(r)t		
Biphones	be	et	tə	ər	tɪ	ɪl	gəʊ	ʃɒ	ɒp	pɪ	ɪŋ	wɪ	ɪð	æn	ek	pɜː	ɜːt

To perform the extraction of sub-lexical features the transcriptions are decomposed into biphones. This way, as shown in Table 2, a sub-lexical representation of each sentence in terms of biphones is obtained, called further a "biphonic content". The occurrences of proper nouns and numerals are intentionally discarded for the statistical analysis because they are unavoidable in any emotional context.

As illustrated in Figure 2, certain biphones participate in EmoBank with quite a different frequency in positive and in negative sentences. The approach underlying the used statistical methods is based on density measures of the biphones in a given text. The statistical examination of the biphonic content of the sentences from EmoBank showed that there is a strong interdependency between the density measures of the biphones and the readers' evaluation of valence (Slavova, 2020). The statistical analysis that follows is based on the dependencies between biphonic content and valence derived from EmoBank.

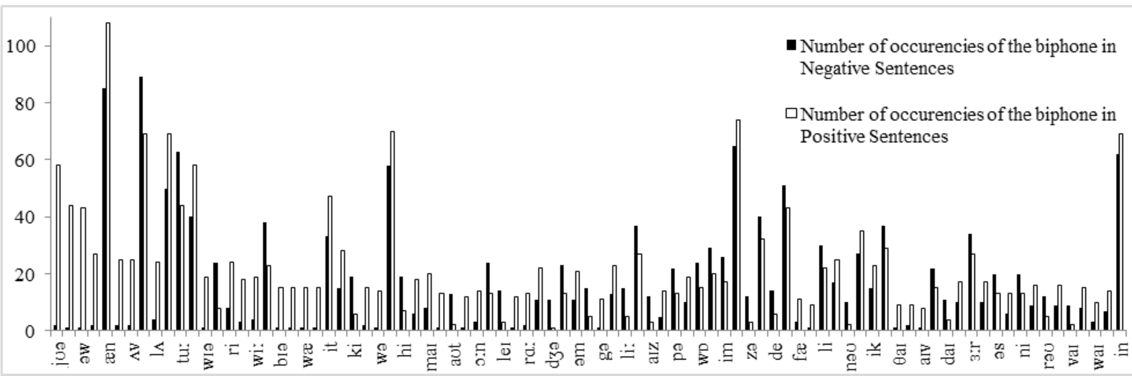


Figure 2. Number of occurrences of biphones in 500 positive and 500 negative sentences from EmoBank

2.3. Investigating the valence-phonemic inter-dependency on real-world texts

It is known that the feature values extracted from a corpus are data sensitive and depend on the corpus content. We suppose that the sub-lexical transmission of valence is an intrinsic aspect of language, and thus the discovered relationships between biphonic content and valence should be observed, at least partially, in real-world texts.

To export the parameters of the relatedness discovered in EmoBank, a “weight” *WBiph* was attributed to each biphone. EmoBank consists of 10,549 isolated sentences, selected from 136 different documents, so in order to identify the phonemic patterning the analysis was based on distributing the sentences into statistically relevant portions – valence layers (see Slavova, 2020 for details). We estimate here the weights of the biphones based on a split of EmoBank into 22 layers of valence, each considered as a distinct “artificial document” of about 3300 words (about three pages). The weight WBiph of each biphone is obtained as a correlation coefficient (Pearson) of the interdependence between the series of the 22 participation scores of this biphone in the 22 valence-layers and the series of valences of the 22 “artificial documents” (the mean of the valence of their involved sentences). The plot of the weights of the 746 biphones present in EmoBank, calculated as explained, is presented in Figure 3. We further use these weights in the analysis of the real-world texts.

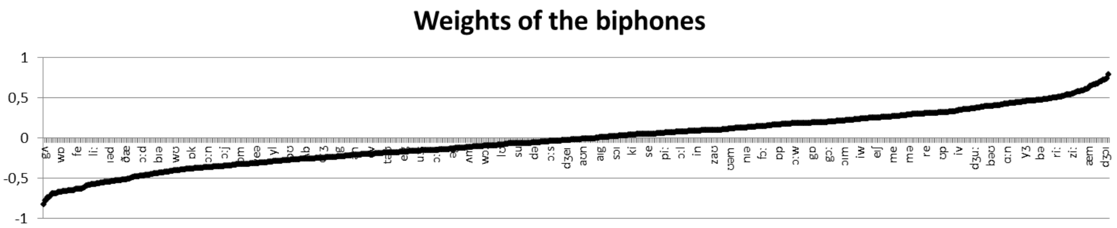


Figure 3. *Weights of the biphones*

In the metadata used to evaluate the weights in the “artificial documents” assembled by the sentences in EmoBank there is no structured linguistic message involved, but they are homogeneous in valence. On the contrary, the real-world texts, have the normal structure of a linguistic message that conveys meaning. Several consecutive pages of structured real-world text are rarely charged with the same valence. However, if the dependencies derived from the corpus are universal in nature, they will be detectable in real-world texts.

The examination presented next aims at checking if the discovered dependencies between biphonic content and valence are universal for the English language. A preliminary test, reported in Slavova (2020), compared the evaluation of the subjects of the study of valence conveyed by fictional texts and the biphonic content of these texts. This test gave a promising result, but was based on quite a small sample.

The statistical analysis presented here is based on Internet News (Figure 4). The aim is to examine a larger sample of various, randomly selected texts.

Hundred news items published on the Internet were downloaded using search engines that look for positive or negative words in the headlines. The headlines were in the center of several studies of media discourse, all of them showing the importance of the headlines for attracting, inspiring emotion, and influencing the opinion of the readers. It was further assumed that the headline aims at announcing the emotional tone of the text that follows it.

The words to look for in the headlines were highly charged with valence. They were taken from the two data sources of emotional ratings of English words: 1. the “Glasgow Norms” (Scott et al., 2018) containing normative ratings for emotion dimensions of 5,553 words and 2. the “Norms of valence, arousal, and dominance” (NVAD) (Warriner, Kuperman, and Brysbaert, 2013), providing ratings for 13,915 English words. The list of emotion keywords for the Internet search was composed of the 150 most positive (love, peace, happiness, freedom, etc.) and the 150 most negative (horrible, kidnapped, malaria, terror, etc.) words from these two sources.

The download was performed in the course of two distant short periods – for the duration of a week in January 2020 (N=45) and for the duration of a week in

January 2021 (N=55). Half of the downloaded texts were retrieved automatically by a keyword search engine, and the other half was retrieved by manual search (using the Google Search Engine). The headlines retrieved in this way are, for example: “Copenhagen's new Metro is a thing of beauty”, “Heavy winds disrupt traffic, kill one”, “A photographic history of men in love”, etc. Most of the texts (N=72) were from major agencies such as CNN, Reuters, Fox News, The Independent, etc. Various topics were covered – crime and justice, environment, business, economy, sports, politics and society, science, etc.

In this test, the statistical inquiry was whether the phonemic content of the text of the news item can predict the emotional tone indicated in its headline.

First, the texts were divided into two groups of emotional tone – positive and negative. The keyword in the headline was used as an indication for the emotional tone of the text, where 3 texts were classified to the inverse group because of the presence of a negation in the headline statement (e.g., “Tayla Harris: 'I would be happy to go through it all again but I'd rather not”). Thus, the headlines “announced” two groups of emotional tone – positive (N=43) and negative (N=57).

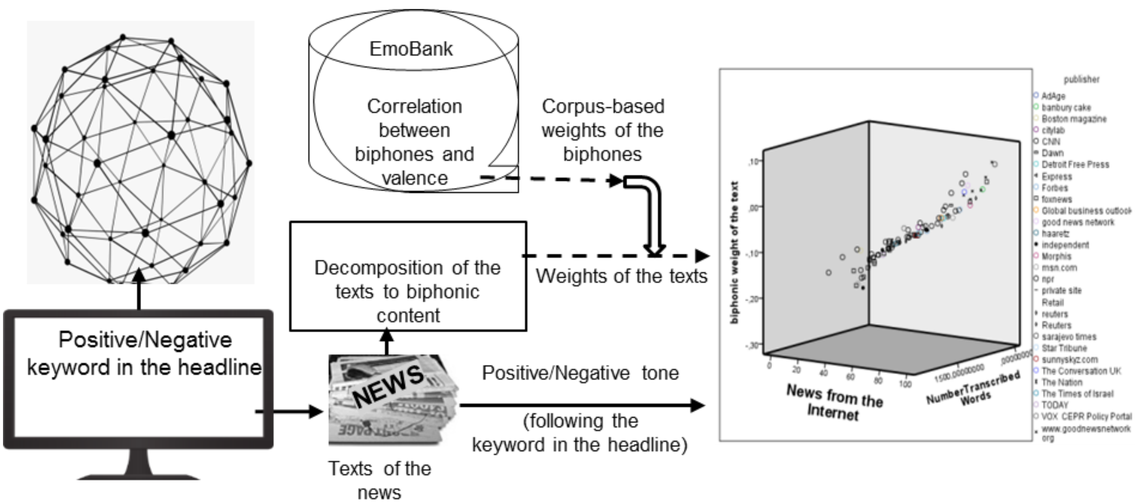


Figure 4. General scheme of the test with Internet news

The downloaded texts were decomposed into words, where the proper nouns and the numerals were protected from further decomposition. The word content of the texts was decomposed into biphones (175,312 occurrences of biphones were found). An example of a sentence from a news item on the topic of Science, decomposed to words and to biphones with their corresponding weights, is given in Appendix.

The investigated texts are very dissimilar in length – the shortest text contained 89 words (proper nouns and numerals excluded) and the longest – 2,367. Having this big variance in the length of the investigated sub-lexical content, to cover the short texts as much as possible, all the participating biphones were taken into consideration regardless of the p-value of their correlation with valence (derived from EmoBank). This was not the case in the preliminary test (Slavova 2020), where all 20 texts used for the experiment were about one page long and the evaluation was based on the biphones that displayed highly reliable correlation in EmoBank.

The measure used to evaluate the sub-lexical indication for valence of the texts downloaded from the Internet was the “weight” of each news item calculated as a sum (normalized) of the weights WBiph of all the included biphones that compose it:

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

where $WBiph_i$ is the EmoBank-derived weight of the biphone *i* (i = 1..746), n_i is the number of times that the biphone *i* appears in the biphonic content of the news item *j* and $NWord_j$ is the number of transcribed and decomposed words in the news item *j*.

Two word-based estimations of the emotional tone were calculated to verify whether the biphonic content is related to the occurrence of distinct emotionally charged words in the texts. The calculation was based on the Mean values of the valence-ratings of the words as given in Glasgow Norms and NVAD. The word-based estimations of valence of the news items were obtained as the average of the valence scores of those words for each text that have emotion ratings in the two dictionaries. The greater part of the words available in Glasgow Norms and NVAD are singular grammatical forms and the normal forms of the verbs, so ratings of valence could not be accorded to all words in the investigated texts. The dictionaries covered as follows: Glasgow norms – 20% and NVAD – 28% of the words in the investigated texts, whereas the biphonic decomposition covered 84% of the words (proper nouns and numerals were excluded).

3. Results

A usual regression procedure was applied to the weights $WENews_j$ of the texts where the dependent variable was "tone of the headline" (dummy coded). The weights of the texts correlated with the tone of the headlines at r = 0.562, p-value < 0.0001. The weights of this arbitrary set of news items showed a noticeable difference in the statistical profile for the two emotional tones (ANOVA: Mean squares 0.004 within groups, 0.150 between groups, F = 44.725, p < 0.0001).

The plots in Figure 5 visualize the statistical differences derived following the biphonic content of the texts for the two groups of emotional tone. When the mean of the weights of the 100 texts is taken as a zero-baseline, based on their weights, 76% of the texts are correctly distributed as positive or negative. There is no difference in the ratio of correct estimations between the negative and positive groups of news items. The decomposition of these various and randomly downloaded texts into biphones, followed by application of the derived from EmoBank weight for each biphone, provided a quantitative estimation that corresponds very well to the emotional tone, as "declared" by their headlines.

A question of special attention is the length of the texts. As explained, the weights of the biphones were statistically derived from valence-homogeneous assemblies of sentences from EmoBank, equivalent to about 3 pages of text.

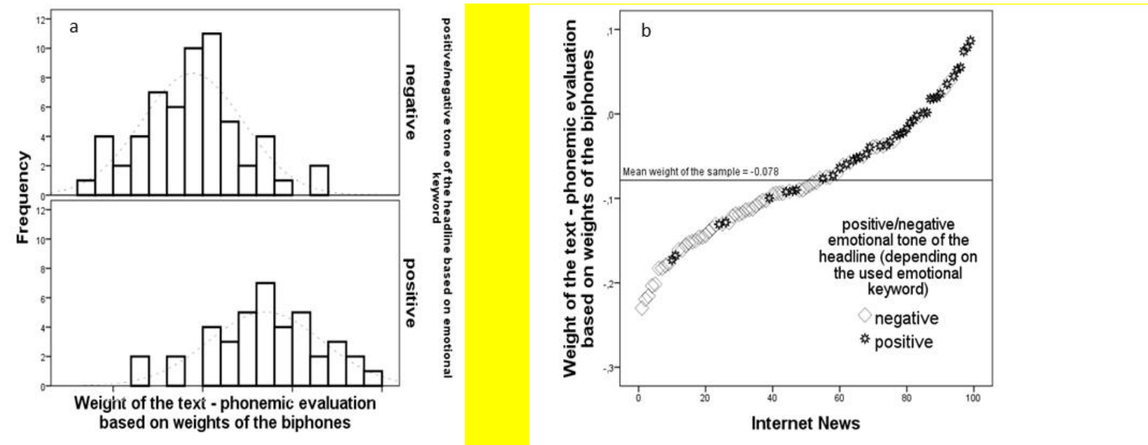
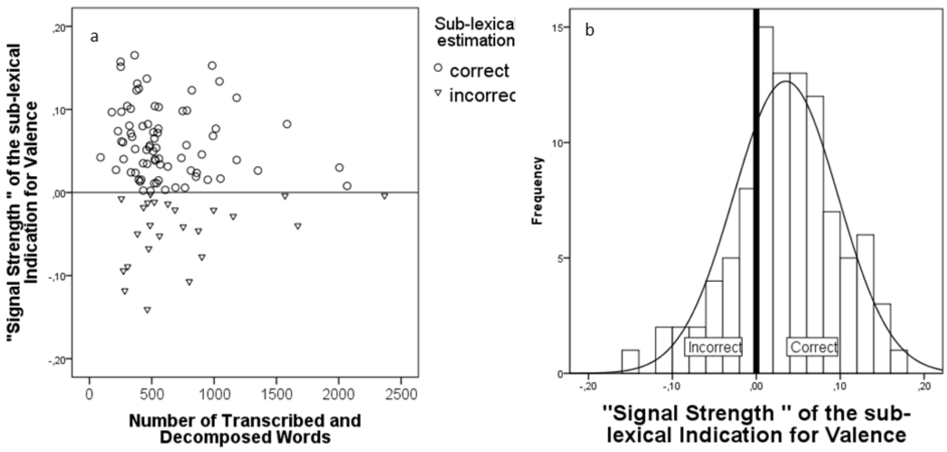


Figure 5. Histograms (a) and a plot (b) of the weights of the texts (rank-ordered by weights) and emotional tone of the headline

The investigated news items carry sub-lexical information involved in the biphonic content which, as seen in Figure 5.b., discriminates quite well the texts following their weights. The weighs of the news items are normalized over the number of transcribed words in the texts and should not be sensitive to the length of the text. The differences in the weights of the news items show differences in the density and the weights of the positive and the negative biphones involved. Thus the weights reflect, figuratively speaking, the incorporated at sub-lexical level “Signal Strength”, carried by the biphones involved. To evaluate whether this “strength of the signal” is independent of the length of the text, we introduce a measure of “strength” of the biphonic information that indicates valence - SigStren, defined as follows:

$$SigStren_j = \pm(WENews_j - Mean(WENews_j, j = 1..100)) \qquad (2)$$

where $WENews_j$ is the weight of the text as calculated following (1), $Mean(WENews_j, j = 1..100)$ is the mean of the weights of all the investigated texts, $SigStren_j$ is the measure of the “strength” of sub-lexical information indicating valence for the text *j*. If the text is wrongly "classified", the difference calculated by (2) is set to negative, otherwise it has a positive value. This measure reflects the discriminative aspect of the estimation of valence based on sub-lexical representation of the text. The introduced measure is normally distributed (Figure 6.b).



a. b.

Figure 6. a. "Signal Strength" indicating valence at the sub-lexical level depending on the length of the news items. b. Histogram of the introduced measure of "Signal Strength" indicating valence at sub-lexical level

The performed correlation analysis suggests that the length of the texts in our sample and the "strength of the sub-lexical signal" SigStren are not related (correlation between the length and the strength of indication r = - 0.1; p-value < 0,2). The plot that illustrates this lack of dependency is given in Figure 6a.

This result suggests that the weights of the biphones derived from EmoBank represent a sufficiently robust indication of valence and can be applied for valence estimation of short texts. However, more data analysis is needed to explore the limits of the approach. The overall result obtained so far shows that the explained approach can be used in text processing tasks to provide speedily statistics for estimation of valence.

The word-based estimations, using Glasgow Norms and NVAD, gave much less clear groupings and correlations with the emotional tone announced in the headlines. It has to be noted again that the dictionaries used cover only partially the word content of the texts. The correlation between the word-based weights of the texts and the emotional tone "announced" in the headlines (dummy coded) were: Glasgow Norms: r = 0.374, and NVAD: r = 0.308. The plots in Figure 7 illustrate what has been said. Thus, the sub-lexical biphonic content covering the overall set of words in the text (numerals and proper nouns excluded) provides a more reliable estimation of valence.

The comparison of the word-based estimation with the result based on the biphonic content of the texts leads to the conclusion that the phonemic pattern related to valence, available in the overall set of the used words, leads to a more discriminative estimation of the valence of the texts.

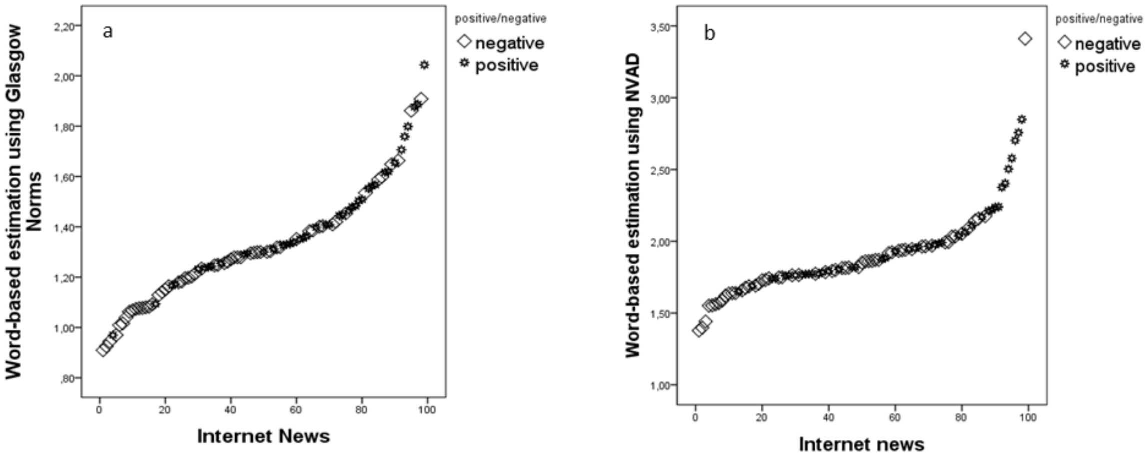


Figure 7. Plots of the word-based weights (rank-ordered by weights) and the emotional tone of the texts (a. Glasgow norms; b. NVAD).

4. Discussions

The biphonic content of the texts shows relatedness with valence. How did words come to incorporate these sound patterns? It is not possible, unfortunately, to derive from this result an explanation for the principle by which the observed sub-lexical phenomenon has "entered" the lexical language content.

As the measures used in the analyses of relatedness with valence are based on "artificial documents" assembled by sentences, one may expect that the emotional valence perceived by the EmoBank readers is prompted by the meanings of these sentences. That leads to the assumption that the set of all the words used in lengthy valence-homogeneous texts does not have an arbitrary phonemic content, but a phonemic content that reflects the emotional valence involved. But words are to indicate meaning.

The relation between sounds, emotion, and meaning has recently started to be investigated using brain imaging techniques. The first neuroimaging evidence for the potential for emotions underlying the sound of words was provided in a recent study by Aryani, Hsu, and Jacobs, (2018), investigating brain response to sub-lexical arousal of presented spoken words (the sub-lexical arousal was calculated based on features extracted from the acoustic representation of the words). The performed analysis suggested the existence of a unifying neural network of affective sound. The results showed that subjects are sensitive to arousal-specific emotional information in the sound of words even when the focus of attention is not directed on that aspect. The examination of the brain imaging results led the authors to propose that emotional dimension arousal links sound to meaning. Until now, to our knowledge, there are no similar brain imaging studies that examine relationships between sound, meaning and valence.

Further, a related investigation by Aryani, Isbilen, and Christiansen (2020), based on a "bouba-kiki" experiment discovered that the extent to which a word is matched with a rounded or spiky shape depends on the level of sub-lexical arousal elicited by its sound. The study showed that the dependency between the arousal involved in the sounds and the meaning is generalizable to new sets of pseudo-words. These findings, the authors say, highlight the role that emotion may play in language development and evolution by grounding associations between abstract concepts and linguistic signs.

The observed here emotional signs at the sub-lexical levels suggest, too, that the language has evolved by exchanging between the individuals a consolidation of both the conceptual meaning and the emotion. Such a phenomenon could explain the results exposed in the present study.

Is the cognitive process of forming a conceptual meaning related to the emotional valence and how? This question has still not found a clear answer in the fields of scientific knowledge, but further investigations and analyses should be able to provide it.

To better clarify the results of the presented study we started to perform further investigations that could provide additional data to answer the following questions:

1. Is the phenomenon dependent on the native language of the reader;
2. Do the biphones per se inspire positive or negative emotion in the listener.

The investigation of such questions would contribute to the clarification of the observed valence-related phonemic phenomenon.

The described sublexical approach and result are useful for the fields of sentiment analysis and text emotion recognition and need further investigation. The methods in these domains use mainly feature extraction techniques based on words in emotion evaluated corpora (see Acheampong, Wenyu, and Nunoo-Mensah, 2020 for a recent review). However, the interest in phonetic features revealed in recent years led to involving the rates of phoneme occurrence as input features to constructing emotion classifiers. A recently developed emotion recognition system (Liu, Rehman, Wu, Cao, and Hao, 2021) confirms the prominence of phonemic features and the effectiveness of their use.

5. Conclusions

The body of a structured linguistic message is charged with an emotional undertone which is detectable when the words used are assumed as mere phonetic constructs. This property of the language is detectable in texts.

A previous statistical investigation made evident the existence of a connection between the valence of textual material and its phonemic content presented as composed by biphones.

We performed a test based on randomly downloaded news items to check if the discovered relationship between biphones and valence is valid outside the corpus. When applying the proposed sublexical decomposition, the quantitative estimation based on their biphonic content matches to a high degree the emotional tone indicated by the headlines. The observed phenomenon is present and detectable even in short structured texts.

The results put forward the hypothesis that, during language evolution, the process of assembling phonemes into speech units that indicate meaning is related to emotion. This would suggest that emotional valence and meaning are profoundly connected in the mechanisms of cognizing and expressing the world.

Acknowledgements

Parts of the text processing tasks were accomplished within a training course in programming at the New Bulgarian University. The authors express gratitude to the students in computer science Lazar Dilov, Vasil Primov, Dimitar Georgiev, Irina Hristova, Vesselin Stoyanov, Denitza Slavcheva, and Martin Lozev, who, during the successive years of work on this project, developed modules of automation that helped the investigation.

Many thanks to the project collaborator, Richard Traub, who, in the attempt to find tracks that explain the observed statistical picture, did an extended analysis of the published research in the domain of emotions and suggested directions of reasoning.

Conflict of interests

The authors declare no conflict of interest.

References

Abbassi, E., Kahlaoui, K., Wilson, M. A., & Joanette, Y. (2011). Processing the emotions in words: The complementary contributions of the left and right hemispheres. Cognitive, Affective, & Behavioral Neuroscience, 11(3), 372-385. https://doi.org/10.3758/s13415-011-0034-1

Acheampong, F. A., Wenyu, C., & Nunoo-Mensah, H. (2020). Text-based emotion detection: Advances, challenges, and opportunities. Engineering Reports, 2(7), e12189. https://doi.org/10.1002/eng2.12189

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

Aryani, A., Conrad, M., & Jacobs, A. M. (2013). Extracting salient sublexical units from written texts:"Emophon," a corpus-based approach to phonological iconicity. Frontiers in psychology, 4, 654. https://doi.org/10.3389/fpsyg.2013.00654

Aryani, A., Hsu, C. T., & Jacobs, A. M. (2018). The sound of words evokes affective brain responses. Brain sciences, 8(6), 94. https://doi.org/10.3390/brainsci8060094

Aryani, A., Isbilen, E. S., & Christiansen, M. H. (2020). Affective arousal links sound to meaning. Psychological science, 31(8), 978-986. https://doi.org/10.1177/0956797620927967

Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A. M., & Conrad, M. (2016). 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

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. Cortex, 63, 196-205. https://doi.org/10.1037/aca0000033

Blasi, D. E., Wichmann, S., Hammarstrom, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound-meaning association biases evidenced across thousands of languages. Proceedings of the National Academy of Sciences, USA, 113, 10818-10823. https://doi.org/10.1073/pnas.1605782113

Buechel, S., & Hahn, U. (2017.a). Readers vs. writers vs. texts: Coping with different perspectives of text understanding in emotion annotation. In Proceedings of the 11th Linguistic Annotation Workshop (pp. 1-12). https://doi.org/10.18653/v1/w17-0801

Buechel, S., & Hahn, U. (2017.b). EMOBANK: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers (pp. 578-585). https://doi.org/10.18653/v1/e17-2092

De Saussure, F. Course in General Linguistics (trans. Roy Harris). London: Duckworth. (1916/1983).

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. Trends in Cognitive Sciences, 19, 603-615. https://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., Perlman, M., & Perniss, P. (2020). Construals of iconicity: experimental approaches to form–meaning resemblances in language. Language and Cognition, 12, 1-14. https://doi.org/10.31234/osf.io/9qb6a

Ekman, P. (1999). Basic emotions. Handbook of Cognition and Emotion, 45-60.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. Philosophical Transactions of the Royal Society B: Biological Sciences, 369, 20130298. doi:10.1098/rstb.2013.0298 https://doi.org/10.1098/rstb.2013.0298

Jakobson, R. (1960). Linguistics and poetics. In Style in language (pp. 350-377). MA: MIT Press.

Kawahara, S., & Shinohara, K. (2012). 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). retrieved from: https://escholarship.org/content/qt47b452vw/qt47b452vw.pdf

Kim, J. M., Sidhu, D. M., & Pexman, P. M. (2020). Effects of emotional valence and concreteness on children’s recognition memory. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.615041

Köhler, W. Gestalt Psychology. New York, NY: Liveright. 1929.

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. Cognition, 114, 19-28. https://doi.org/10.1016/j.cognition.2009.08.016

Liu, Z. T., Rehman, A., Wu, M., Cao, W. H., & Hao, M. (2021). Speech emotion recognition based on formant characteristics feature extraction and phoneme type convergence. Information Sciences, 563, 309-325. https://doi.org/10.1016/j.ins.2021.02.016

Macoir, J., Hudon, C., Tremblay, M. P., Laforce, R. J., & Wilson, M. A. 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, 1-12. https://doi.org/10.1080/17470919.2019.1577295

Maddock, R. J., Garrett, A. S., & Buonocore, M. H. Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. Human brain mapping, 2003, 18(1), 30-41. https://doi.org/10.1002/hbm.10075

Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. Neuron, 38(3), 487-497. https://doi.org/10.1016/s0896-6273(03)00197-1

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language?. Philosophical Transactions of the Royal Society B: Biological Sciences, 369. doi:10.1098/rstb.2013.0299 https://doi.org/10.1098/rstb.2013.0299

Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(3), 371. https://doi.org/10.1037/0278-7393.14.3.371

Ponari, M., Norbury, C. F., & Vigliocco, G. (2018). Acquisition of abstract concepts is influenced by emotional valence. Developmental science, 21(2), e12549. https://doi.org/10.1111/desc.12549

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

Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., ... & Nummenmaa, L. (2015). Discrete neural signatures of basic emotions. Cerebral cortex, 26(6), 2563-2573. https://doi.org/10.1093/cercor/bhv086

Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2018). The Glasgow Norms: Ratings of 5,500 words on nine scales. Behavior research methods, 1-13. https://doi.org/10.3758/s13428-018-1099-3

Sidhu, D. M., Vigliocco, G., & Pexman, P. M. (2020). Effects of iconicity in lexical decision. Language and Cognition, 12, 164-181. https://doi.org/10.1017/langcog.2019.36

Slavova, V. (2019). Towards emotion recognition in texts – a sound-symbolic experiment. International Journal of Cognitive Research in Science, Engineering and Education, 7(2), 41-51. https://doi.org/10.5937/ijcrsee1902041s

Slavova, V. (2020). 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

Ullrich, S., Aryani, A., Kraxenberger, M., Jacobs, A. M., & Conrad, M. (2017). 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. Frontiers in psychology, 7, 2073. https://doi.org/10.3389/fpsyg.2016.02073

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. Behavior research methods, 45(4), 1191-1207. https://doi.org/10.3758/s13428-012-0314-x

Whissell, C. (1999). 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, 89(1), 19-48. https://doi.org/10.2466/pms.1999.89.1.19

Yang, J., Asano, M., Kanazawa, S., Yamaguchi, M. K., & Imai, M. (2019). Sound symbolism processing is lateralized to the right temporal region in the prelinguistic infant brain. Scientific Reports, 9, 1-10. https://doi.org/10.1038/s41598-019-49917-0

Authors Information

***Velina Slavova** - New Bulgarian University, department of Computer Science, vslavova@nbu.bg*

***Major Fields of Scientific Research**: AI, Cognitive Science*

***Filip Andonov** - New Bulgarian University, department of Computer Science, fandonov@nbu.bg.*

***Major Fields of Scientific** Research: multi-criteria optimization, data mining, text processing, Python language*

APPENDIX

Example of a sentence from the sample of investigated news items, decomposed to biphones.

Sentence: ***The study discovered levels of happiness reported by those who were married was higher than those who do not have spouses, but only when their husband or wife was in the room.***

word	position	Transcript	sequence	biphone	weight of the biphone
The	0	'ðə	1	ðə	-0,0132
study	1	ˈstʌdi	1	tʌ	-0,2027
study	1	ˈstʌdi	2	ʌd	-0,6903
study	1	ˈstʌdi	3	di	-0,2530
discovered	2	dɪ'skʌvəd	1	dɪ	-0,5706
discovered	2	dɪ'skʌvəd	2	ɪs	-0,2265
discovered	2	dɪ'skʌvəd	3	kʌ	0,2546
discovered	2	dɪ'skʌvəd	4	ʌv	0,5098
discovered	2	dɪ'skʌvəd	5	və	-0,4344
discovered	2	dɪ'skʌvəd	6	əd	0,0184
levels	3	'levəlz	1	le	-0,0652

levels	3	'levəlz	2	ev	-0,1996
levels	3	'levəlz	3	və	-0,4344
levels	3	'levəlz	4	əl	0,3949
of	4	əv	1	əv	0,2993
happiness	5	ˈhæpinəs	1	hæ	-0,1290
happiness	5	ˈhæpinəs	2	æp	-0,0982
happiness	5	ˈhæpinəs	3	pi	0,2829
happiness	5	ˈhæpinəs	4	in	0,0870
happiness	5	ˈhæpinəs	5	nə	0,6417
happiness	5	ˈhæpinəs	6	əs	-0,1985
reported	6	rɪ'pɔ:tɪd	1	rɪ	0,0490
reported	6	rɪ'pɔ:tɪd	2	ɪp	0,1365
reported	6	rɪ'pɔ:tɪd	3	pɔː	-0,2023
reported	6	rɪ'pɔ:tɪd	4	ɔːt	-0,0588
reported	6	rɪ'pɔ:tɪd	5	tɪ	-0,3971
reported	6	rɪ'pɔ:tɪd	6	ɪd	-0,6941
by	7	baɪ	1	baɪ	0,0428
those	8	ðəʊz	1	ðəʊ	-0,4743

those	8	ðəʊz	2	əʊz	-0,1777
who	9	huː	1	huː	-0,4635
were	10	'wɜː	1	wɜː	-0,1798
married	11	ˈmærid	1	mæ	0,0969
married	11	ˈmærid	2	ær	-0,1412
married	11	ˈmærid	3	rɪ	0,0490
married	11	ˈmærid	4	ɪd	-0,6941
was	12	'wɒz	1	wɒ	-0,6704
was	12	'wɒz	2	ɒz	-0,5978
higher	13	ˈhaɪə(r)	1	haɪ	-0,0051
higher	13	ˈhaɪə(r)	2	ɪər	0,0900
than	14	ðæn	1	ðæ	-0,5137
than	14	ðæn	2	æn	0,4481
those	15	ðəʊz	1	ðəʊ	-0,4743
those	15	ðəʊz	2	əʊz	-0,1777
who	16	huː	1	huː	-0,4635
do	17	duː	1	duː	-0,3941
not	18	nɒt	1	nɒ	-0,7714

not	18	nɒt	2	ɒt	-0,6677
have	19	hæv	1	hæ	-0,1290
have	19	hæv	2	æv	0,0367
spouses	20	spaʊsiz	1	paʊ	-0,1550
spouses	20	spaʊsiz	2	aʊs	-0,0579
spouses	20	spaʊsiz	3	si	-0,3622
spouses	20	spaʊsiz	4	iz	-0,1679
but	21	bʌt	1	bʌ	-0,6159
but	21	bʌt	2	ʌt	-0,6566
only	22	ˈəʊnli	1	əʊn	-0,7329
only	22	ˈəʊnli	2	li	-0,2955
when	23	wen	1	we	-0,0762
when	23	wen	2	en	-0,0199
their	24	ðeə(r)	1	ðeə	-0,3096
their	24	ðeə(r)	2	eər	-0,2291
husband	25	ˈhʌzbənd	1	hʌ	-0,3145
husband	25	ˈhʌzbənd	2	ʌz	-0,4656
husband	25	ˈhʌzbənd	3	bə	0,4756

husband	25	ˈhʌzbənd	4	ən	-0,1522
or	26	ɔː(r)	1	ɔːr	0,4091
wife	27	waɪf	1	waɪ	0,1600
wife	27	waɪf	2	aɪf	0,6446
was	28	'wɒz	1	wɒ	-0,6704
was	28	'wɒz	2	ɒz	-0,5978
in	29	ɪn	1	ɪn	0,0955
the	30	'ðə	1	ðə	-0,0132
room	31	ruːm	1	ruː	-0,4045
room	31	ruːm	2	uːm	-0,6925