

How arbitrary is language? A cross-cultural research on phono-emotional iconicity

Devashish Kamble, Matteo Melis, Anastasiia Salova, Ecesu Ürker

February 2022

1 Aim of the Research

The project aims to investigate phono-emotional iconicity. The question at hand is whether emotionally valenced words are arbitrary across languages or whether they involve phonological regularity. The research contributes to the ongoing discussion about linguistic arbitrariness by providing evidence for or against emotional iconicity.

2 Prior Research

2.1 Background

In 1916, Ferdinand de Saussure stated that "the sign is arbitrary", whereby the relationship between the sign and its referent is entirely conventional, identical to the relationship between the sounds that constitute a word and its semantic meaning. However, many subsequent studies have shown that, instead, the sounds of words are systematically related to their meaning. (Adelman J.S., et al., 2018; Myers-Schulz B., et al., 2013).

A well-known experiment is the "takete-maluma" effect illustrated by Köhler (1929) in which the majority of people tend to agree that "takete" is the angular object and "maluma" is the round one (Adelman J.S., et al., 2018).

While a recent research proposed by Yu C. S. P. (et al., 2021) investigates the "gleam-glum" effect, which states that words with the /i:/ phoneme (gleam) are perceived as conveying a much more positive valence than words with the /ʌ/ phoneme (glum). The authors found a valid embodied theory that associates the sound produced with the musculature pattern activates while showing positive emotion, congruent with the phoneme /i:/, or negative emotion, congruent with the phoneme /ʌ/ (Yu, C. S. P., et al., 2021).

In 2003, Wiseman and van Peer (Auracher J., et al., 2010), found a significant difference in emotional valence in nasal and plosive sounds. Participants chose the latter (p, b, t, and d) as more appropriate to express happy feelings, and the former (m,n) to express sad feelings. However, there are also few experiments that have found no correlation whatsoever between plosive and nasal sounds and the emotional valence of phonemes (Kraxenberger M., and Menninghaus W., 2016).

2.2 What is sound iconicity?

In the current literature, the following concepts are used in the sound-meaning relation:

1. Phonosemantic is a general word to refer to every sound-meaning relation (Auracher J., et al., 2010).
2. Sound symbolism, which refers to the conventional relation between the sign (sound) and its referred meaning (e.g., glitter, glow, glimmer, etc).
3. Sound iconicity, which refers to the connection between the sound and its meaning, and can be subdivided in three domains:
 - (a) Corporal sounds (e.g coughing)
 - (b) Onomatopoeic sounds
 - (c) Acoustic symbolisation of non-acoustic phenomena, which is what we are interested in (Hinton et al., 1994).

2.3 Why should sound iconicity exist?

There are three different theoretical framework approaches that propose different hypotheses on sound iconicity (Auracher J., et al., 2010):

- Anthropology proposes the vocal theory, which claims that language developed from instinctive reactions to “distress, elation and courting” (Mayeux & Kandel, 1985).
- Psychology starting from the results proposed by Köhler (1929) in the “takete-maluma” effect, they are interpreted as a frequency code. So there is an everyday experience of non-arbitrary connection that leads to a cognitive mapping of visual and auditory input.
- Neuroscience instead of connections created by experience, suggests an inherited connection between different areas in the brain. For example, in many languages, the words to express “littleness” require a narrowing of the vocal trait. From this Ramachandran & Hubbard (2001) hypothesized that a primitive vocabulary could have been evolved in a vocabulary of tongue/palate/lip movements.

3 Key Questions

The fundamental question is, undeniably, *‘how arbitrary is language?’* As mentioned before, the conventionality between a sign and its referent has been a debatable topic for decades. Here, phono-emotional iconicity was taken as a narrower aspect of the principle of linguistic arbitrariness; and the actual research aims to provide conclusive evidence, through varying methodologies, as to whether the phonological properties of the word affect its emotional valence.

Overall, the literature tends to implicate phono-emotional iconicity. Yet, encouraged by the existing research, we decided to expand on the previous findings and adopt a different strategy that would both utilize four unrelated languages for pattern extraction — thus, creating a more solid foundation for any regularity/universality claims, and make use of the pseudowords (derived from recurrent patterns in positively and negatively valenced words) as the more reliable testbed — canceling out the semantic meaning as a component contributing to the final judgment on the emotional valence.

‘Will the pseudowords get evaluated according to their phonology-motivated positive or negative emotional valence?’ is the primary question. We hypothesize that the emotional valence of the pseudowords will get judged depending on their phonological properties. In line with our first hypothesis, the data collected from the participants’ input must clearly indicate consistency between the participants’ judgments following the creation of the regularized pseudowords

and valence, implied by initial clustering. Alternatively, phonological properties of the pseudoword may not affect the evaluation of their emotional valence. In this case, there is no strong emotional-coloring or consistency between the final judgments and implied valence.

4 Intended Results

We are using the pseudowords that sound English-like and are generated from the phonemes of four different languages, which in turn are obtained from clustering emotionally valenced words to test our above mentioned hypothesis.

In the process of where the participants of respective languages rate the words, we perform a data reduction step wherein based on a certain threshold of the Likert scale we obtain our valenced words. But it may so happen that the resultant words are very few, in such a case we might adjust the threshold and derive additional words.

During the phase of extracting phonemes and categorizing them into positive and negative clusters we might end up with very few or none as the phonemes can be present in both categories. In this case, we can lower the bands to accommodate more phonemes keeping in mind their significance. If we fail to include more phonemes through the clusters, dismally that would bring us to the end of the experiment and we conclude that language is arbitrary or we tweak our procedure to further continue working.

Another circumstance could be that suppose clustering works but phonemes extracted are not suitable for English and the likelihood of this happening cannot be ignored. In this scenario, we plan to change to language which is apt for the given phonology and also distinct from the four languages used in the experiment.

Based on the data analysis procedure, we expect to find a low p-value that would suggest the difference in the valence to be as high as possible and in turn prove our hypothesis stating phonological properties affect the perception of emotionally valenced words thus affirming the non-arbitrariness and universality of languages. In the case where the ratings provided by the participants are truly comparable, it would be indicated by a high p-value and disprove our hypothesis showing that phonological properties are uncorrelated to the consciousness of emotional words.

5 Procedure

The initial step will be dataset acquisition. EMOTE database for emotional words in English (Grühn, 2016) will be used as a starting point for obtaining emotionally valenced words. This particular dataset has subjective emotional valence ratings of more than 2000 words which would provide us with many words to start with. Then, the words in the dataset will be translated into the languages that will be used in the study: Turkish, Marathi, Ukrainian, and Italian. To ensure emotional valence of the translated words in these languages, in an online experiment, the native speakers of these languages will evaluate the emotional valence of the words using a Likert scale [1-7] where 1 means “extremely negative” and 7 “extremely positive”. At the end of the experiment, there will be a part where, if they find it necessary, the participants can add and rate emotional words that are missing from the dataset. Any word provided by 20% of the participants will be added to the database.

The second step will be database reduction. For each language, the words that have average ratings greater than 6 and lower than 2 will be selected to obtain positive and negative words, respectively. Then, using the International Phonetic Alphabet, the selected words will be translated into their phonological counterparts.

The subsequent step will be clustering. Using scikit-learn library (Pedregosa et al., 2011) for recognizing patterns and clustering with Python, the statistically recurrent patterns of sounds across the words will be extracted for both positive and negative words. The sounds that are clustered significantly only in one type of word (positive/negative) will be used in the study.

Using the sounds that are extracted during clustering, pseudowords that represent the sound patterns that are found in the positive and negative emotion words will be created. Since the final experiment will utilize English speakers, the pseudowords will be formed only using the clustered sounds that belong to the English phonological system. During pseudoword generation, we will take into account the phonological rules of English and generate English-sounding words. At the end of this process, we will obtain two sets of English pseudowords: one with the phonological properties of positive words and the other with negative words.

The next step will be creating sentences to test participants. Short English sentences (eg. He/She/They is/are ----) will be created by placing pseudowords in the gap. In an online experiment, English speakers will be asked to rate the emotional valence of the sentences with a Likert scale [1-7] where 7 will correspond to “extremely positive” and 1 to “extremely negative”. To measure the consistency of the responses of the participants, each participant will be presented with each word three times with different subjects.

As the last step, we will conduct data analysis. There will be one independent variable (the phonology of the words) with two levels (positive/negative). In this step, the change in ratings of participants depending on the phonological characteristics of the words will be compared. To do so, we will initially calculate the mean rating for each word across and within participants. Then, we will compare the ratings of positive and negative words using a t-test and our significance level will be 0.05.

6 References

- ❖ Adelman, J. S., Estes, Z., & Cossu, M. (2018). Emotional sound symbolism: Languages rapidly signal valence via phonemes. *Cognition*, 175, 122-130.
- ❖ Auracher, J., Albers, S., Zhai, Y., Gareeva, G., & Stavniychuk, T. (2010). P is for happiness, N is for sadness: Universals in sound iconicity to detect emotions in poetry. *Discourse processes*, 48(1), 1-25.
- ❖ Grün, D. (2016). An English Word Database of EMOtional TErms (EMOTE). *Psychological Reports*, 119(1), 290–308. Hinton, L., Nichols, J., & Ohala, J. J. (1994). Introduction: Sound–symbolism processes. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 1–12). Cambridge, England: Cambridge University Press.
- ❖ Kraxenberger, M., & Menninghaus, W. (2016). Mimological reveries? Disconfirming the hypothesis of phono-emotional iconicity in poetry. *Frontiers in Psychology*, 7, 1779.
- ❖ Myers-Schulz, B., Pujara, M., Wolf, R. C., & Koenigs, M. (2013). Inherent emotional quality of human speech sounds. *Cognition & emotion*, 27(6), 1105-1113.

- ❖ Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830.
- ❖ Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A window into perception, thought and language. *Journal of Consciousness Studies*, 8, 3–34.
- ❖ Yu, C. S. P., McBeath, M. K., & Glenberg, A. M. (2021). The gleam-glum effect:/i:/versus/λ/phonemes generically carry emotional valence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

7 Team Member Contributions

The idea of the project and its outline has been discussed and created during the class hours by the four of us.

- Prior research: **Matteo Melis**
- Key questions: **Anastasiia Salova**
- Intended results: **Devashish Kamble**
- Procedure: **Ecesu Ürker**