Sound of Emotion – Synesthetic Sound Iconicity

Jan Auracher

Keywords: phonosemantics, sound iconicity, multisensory processing

Abstract

The term *Synesthetic Sound Iconicity* refers to sound-meaning relationships which are inherent and thus neither arbitrary nor language-specific. Being universal and allowing for easy-to-conduct detection of semantic properties in texts, Synesthetic Sound Iconicity has the potential for quantitative content analysis and inter-language comparison of text corpora at comparatively low cost. Results from empirical research have provided sound evidence that the ratio of plosive vs. nasal phonemes allows for a text’s emotional tone (as seen by a majority of readers) to be deduced. In this paper, I report results from a recent research project aimed at extending and generalizing previous findings. To this end, the emotional tone (assessed by over 400 readers) and phonetic structure of 16 German poems were analyzed to find correlations. Results suggest a strong interrelationship between emotional tone and frequency of phonemes with specific articulatory properties.
1. Where does Synesthetic Sound Iconicity come from?

Within the realm of phonosemantics, that is the association of single phonemes with semantic concepts, Synesthetic Sound Iconicity refers to sound-meaning associations that are not based on imitation (i.e., not onomatopoeia) (Hinton et al., 1994). Recent studies, in which the relative occurrence of specific phonemes were linked to the emotional tone in texts, have shown a significant connection between plosive sounds and positive-active emotions, and between nasal sounds and negative-passive emotions. This phenomenon has been surprisingly consistent across language families, including Pharaonic Egyptian (~ 2000 BC), classic Chinese (~200 to 400 AD), and modern German and Slavic languages (Wiseman & van Peer, 2003; Albers, 2008; Auracher et al., 2011). The robustness of this effect suggests that it is a universal tendency in human communication to communicate an emotional state by sub-consciously using a higher frequency of specific phonemes.

Observations like this provide solid evidence for the existence of Synesthetic Sound Iconicity. However, there is still no paradigm which provides a sufficient explanation for the expressiveness of sound patterns. One line of argument for explaining the linkage between the sound of language and its emotional tone has been brought forward by Tsur (1992) and Miall (2001), who suggest that contrasts in semantic concepts, such as bright vs. dark and high vs. low, can be expressed through physiological dimensions of phonemes. For example, the categorization of vowels as open vs. closed or front vs. back corresponds to the relative position of their first and second formant, causing the vowels to sound either more diffuse or more compact (Jakobson & Waugh, 1979). Similarly, all phonemes can be categorized according to their articulatory properties. Miall and Tsur argue that phonemes of different categories are systematically associated with specific semantic concepts, and thus can be applied to emphasize an underlying
Empirical evidence for this view comes from recent research by Reilly et al. (2008). The authors report two experiments in which subjects were asked to associate colors, varying in hue (wavelength) and lightness, with tone pitch (sound frequency). In the first experiment, subjects were presented with color dots and pure (sinus) tones. In the second experiment, subjects were asked to choose names for pictures of fantasy animals. The authors found that participants applied the same kind of connection for both associations; they positively linked bright colors (long wavelength) to high-frequency sounds, and dark colors (short wavelength) to low-frequency sounds. The authors argue that these results suggest that the origin of language is based on the mapping of acoustic properties onto the physical properties of objects. They assume that iconic associations based on daily experiences, for example the negative correlation between tone pitch and mass, have been extrapolated to other characteristics that usually accompany mass, such as size and, moreover, more emotional associations like predation. Thus, although in the course of evolution symbolic (arbitrary) sound-meaning connections became dominant, language may originally have been rooted in the iconic attribution of names to objects (LeCoux, 1978).

2. Other research on cross-modal integration

Observations that names or attributions of objects often mimic their physical appearance have a long tradition. Examples are the preference of closed vowels (small articulation tract) for expressing smallness, as in ‘teeny-weeny’ (Ramachandran & Hubbard, 2001), or the association of plosive consonants with spiky forms and continuant consonants with rounded forms (e.g., Köhler, 1929; Westbury, 2005). Ramachandran and Hubbard (2001) claim that this form of phonosemantics is based on neuronal cross-modal connections similar to those that produce synesthesia. In fact, several scholars have suggested that synesthetes apply the same unambiguous
associations between visual and auditory stimuli as non-synesthetes (Marks, 1974, 2004; Cytowic & Eagleman, 2009; Cohen-Kadosh et al., 2009). Thus, Cytowic and Eagleman (2009) conclude that sensory domains, such as hearing and seeing, are orderly and systematically interconnected. Recent research on the cognitive integration of multimodal stimuli has increasingly focused on the influence of semantic (in-)congruence of multi-sensory input (for an overview see Doehrmann & Naumer, 2008). Results show that perception of incongruent vs. congruent combinations of cross-modal stimuli has an influence on cognitive processing, as indicated by observations from behavioral and neuroimaging studies (Laurienti et al., 2004; Spence, 2007; Molholm et al. 2004; Beauchamp et al., 2004; Noppeney et al. 2007; Werner & Noppeney, 2010).

Research on the neural basis of phonosemantics has been conducted with miming content (onomatopoeia) in Japanese. Osaka (2004) compared neural activation patterns of onomatopoetic expressions for pain with those for nonsense words. He found increased activation in the anterior cingulate cortex, a structure generally associated with the experience of pain. Similarly, Arata et al. (2010) found that motor and pre-motor areas of the brain react differently to onomatopoetic words resembling action than they do to normal (not onomatopoetic) verbs and adjectives. These findings suggest that sound has the potential to trigger activity in brain areas that are primarily related to other – non-acoustic – domains of perception. This raises the question of whether such cross-modal connections are applied in natural language use.

3. Hypotheses

Based on the above insights, it was assumed that articulatory characteristics of phonemes are associated with meaning, and in particular with emotional meaning. If the phonetics of language indeed tends to mimic semantic concepts, one can hypothesize that front vowels, which have a higher or
brighter sound and are associated with bright colors, tend to connote happy feelings, whereas back vowels, which have a deeper or darker sound and are associated with dark colors, tend to connote negative and sad feelings. Another relationship which has repeatedly been found in research on phonosemantics associates plosives with spiky shapes and continuants with round shapes (e.g., Westbury, 2005). Ramachandran and Hubbard (2001), in an attempt to explain this phenomenon, reason that “the sharp changes in visual direction of the lines [in the spiky shape] mimic the sharp phonemic inflections [of the plosive sounds]” (p. 19). To put it another way, the interruption and explosive release of the air stream in the pronunciation of plosive sounds mimics the break or discontinuity of lines in spiky shapes. Related findings have been made by Auracher et al. (2011), who found a positive relation between plosives and high arousal. Extrapolating from these findings, it follows that the association of continuants with rounded shapes might also link them to more harmonic feelings of low arousal. Thus, it is hypothesized that...

1. a high frequency of plosives in a text is associated with a high degree of arousal, whereas a high degree of continuants is associated with a low degree of arousal, and
2. a high frequency of front vowels is associated with positive emotions, whereas a high frequency of back vowels is associated with negative emotions.

4. Research

The following presents results of a recent research study designed to explore correlations between the phonetic structure of a text and its emotional tone, as predicted in the above hypotheses. To this end, the frequency of phonemes in a text, categorized by their articulatory properties, was
compared with the text’s assessment by readers. Sixteen German poems, selected randomly from a collection of poems, were read by over 400 readers. Each reader read one poem and subsequently assessed it for eight emotional expressions using a five-point Likert scale. Pairs of emotional expressions (e.g., aggressive and angry, euphoric and happy) were allocated to each quarter of a two-dimensional model for emotions (Larsen & Diener, 1992), with axes for valence (positive vs. negative) and arousal (high / “active” vs. low / “passive”) (Figure 1).

To analyze the phonetic structure of the poems, they were first transcribed into the phonetic alphabet SAMPA (http://www.phon.ucl.ac.uk/home/sampa/index.html). All phonemes were categorized according to their articulatory place (front or back) and manner (consonants: plosive or continuants, vowels: closed or open); this resulted in eight categories (from a 2x4 matrix). Finally, for each poem, the relative occurrence of each cluster of
phonemes was correlated with the poem’s emotional expression values. The results reveal a significant relationship between the frequency of phonetic clusters and emotional tone in poems.

4.1 Design and results

Participants

Participants for this study were recruited in cafeterias, libraries, and other public places at the University of Munich. All participants either held a university degree or were currently enrolled in the university. The poems were read by an average of 31 readers each (50% female and male). Female-male distribution per poem was close to equal, with the greatest disparity being 18 female and 11 male readers for poem 7, and 14 female and 18 male readers for poem 8. Age varied greatly among participants; average age was 27 years, with a minimum of 18 and a maximum of 99. However, over 80% of participants were under 30 and 90% under 40. Age and sex were equally distributed among poems, with a maximum median of 29 for males and 25 for females (poem 1) and a minimum of 22 for males and 21 for females (poem 12). A test for significance (GLM) comparing the answers of male and female readers revealed an influence greater than chance (F = 2.13, df = 12, p < .05). However, this was mainly due to differences for the item fear (F = 8.766, p < .01), which was later excluded from the analysis. Age had not significant influence on the assessment of the poems.

Participants were also asked how often they read poetry: never, seldom, sometimes, or often. Nearly two thirds of participants answered that they had either never or only seldom read poetry. Another quarter of participants reported that they only sometimes read poetry. A MANOVA test found that this variable exerted no significant influence on the answers of participants.

Material

The research used German poetry as the linguistic stimulus. This is because
poetry, compared to other genres of written language, allows authors to express inner states in a dense language. Poetry therefore seems to be particularly suited for the aims of the research. All poems were taken from the 1921 edition of “A Book of German Lyrics” by von Friedrich Bruns (available at: http://www.gutenberg.org/ebooks/8565). The poems used in the study were selected randomly. A list of poems used, together with their index numbers (which are used in this paper), can be found in Table 1.

Table 1: Poems used in the study

<table>
<thead>
<tr>
<th>#</th>
<th>Poem (author)</th>
<th>#</th>
<th>Poem (author)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Du bist wie eine Blume (H. Heine)</td>
<td>9</td>
<td>Ein Gleiches (J.W. v. Goethe)</td>
</tr>
<tr>
<td>2</td>
<td>Säerspruch (C.F. Meyer)</td>
<td>10</td>
<td>Es ragt ins Meer der Runenstein (H. Heine)</td>
</tr>
<tr>
<td>3</td>
<td>April (Th. Storm)</td>
<td>11</td>
<td>Harfenspieler (J.W. v. Goethe)</td>
</tr>
<tr>
<td>4</td>
<td>Der Wirtin Töchterlein (L. Uhland)</td>
<td>12</td>
<td>Sommernacht (D. v. Liliencron)</td>
</tr>
<tr>
<td>5</td>
<td>Die Drei (N. Lenau)</td>
<td>13</td>
<td>Morgenlied (L. Uhland)</td>
</tr>
<tr>
<td>6</td>
<td>Die Stadt (Th. Storm)</td>
<td>14</td>
<td>Ewig jung ist nur die Sonne (C.F. Meyer)</td>
</tr>
<tr>
<td>7</td>
<td>Leise zieht durch mein Gemüth (Th. Storm)</td>
<td>15</td>
<td>Der letzte Baum (F. Hebbel)</td>
</tr>
<tr>
<td>8</td>
<td>Neujahrsglocken (C.F. Meyer)</td>
<td>16</td>
<td>Auf eine Lampe (E. Mörike)</td>
</tr>
</tbody>
</table>

Results

Assessment of poems

All statistical analysis was done with the program IBM SPSS Statistics 19. First, the position of each poem within a two-dimensional model of emotions was assessed. A factor analysis (Varimax with Kaiser normalization) confirmed the expected four factors (the four quarters of the coordinate plane), with two emotional expressions per factor (Figure 1). Correlations within the factors were relatively high, with a mean loading of .81 (maximum: .86 (euphoric); minimum: .73 (calm)). The four factors explained 72% of the overall variance. Two items (fear and excited) did not load for any factor and were therefore excluded from further analysis. Mean values per factor were calculated and each poem was allocated to one of the factors according to its highest mean value (Table 2). Results reveal that
seven poems were rated highest for negative-passive emotions. Two poems each were allocated to the factors positive-active and positive-passive, and no poem had its highest mean value for the factor negative-active. For five poems, the mean values did not reveal a clear dominance of any of the four factors. This, unfortunately, means that factors were unequally represented, which limits the explanatory power of the research.

Table 2: Assessment of poems per factor – mean values (standard deviation)

<table>
<thead>
<tr>
<th>#</th>
<th>positive active</th>
<th>positive passive</th>
<th>negative active</th>
<th>negative passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.77 (.12)</td>
<td>2.42 (.99)</td>
<td>1.08 (.22)</td>
<td>3.27 (1.0)</td>
</tr>
<tr>
<td>2</td>
<td>2.9 (.82)</td>
<td>3.47 (1.0)</td>
<td>1.25 (.56)</td>
<td>2.2 (.92)</td>
</tr>
<tr>
<td>3</td>
<td>3.59 (.90)</td>
<td>3.23 (1.0)</td>
<td>1.36 (1.1)</td>
<td>1.9 (.82)</td>
</tr>
<tr>
<td>4</td>
<td>1.66 (.86)</td>
<td>1.97 (.86)</td>
<td>1.31 (49)</td>
<td>3.85 (.92)</td>
</tr>
<tr>
<td>5</td>
<td>1.15 (.29)</td>
<td>2.18 (.97)</td>
<td>1.84 (.89)</td>
<td>3.73 (.94)</td>
</tr>
<tr>
<td>6</td>
<td>1.66 (.69)</td>
<td>2.43 (.88)</td>
<td>1.30 (46)</td>
<td>3.57 (.82)</td>
</tr>
<tr>
<td>7</td>
<td>3.36 (.94)</td>
<td>3.59 (.87)</td>
<td>1.17 (.75)</td>
<td>2.21 (1.1)</td>
</tr>
<tr>
<td>8</td>
<td>2.67 (1.3)</td>
<td>2.42 (1.2)</td>
<td>1.9 (1.11)</td>
<td>2.63 (1.0)</td>
</tr>
</tbody>
</table>

Analysis of phonetic structure

Subsequently, the relative occurrence of phonemes was assessed for each poem and averaged for each factor (Figure 2). Phoneme clusters were:

- consonants: front-plosives, back-plosives, front-continuants, and back-continuants
- vowels: front-open, back-open, front-closed, and back-closed

As stated above, an association between back vowels and negative feelings, and between front vowels and positive feelings, was hypothesized. Further, it was expected that plosives, in particular front-plosives, would be associated with a high degree of arousal (“active”), while continuants would be associated with a low degree of arousal (“passive”). The results are
consistent with this perspective (figure 2). Although, due to the small number of poems used, no value of significance was calculated, it is notable that, without exception, all poems which were assessed as expressing a high level of arousal had higher frequencies of front-plosives than did other poems. A similar observation can be made for front and back vowels. Poems expressing a positive-active feeling contained more front vowels, while poems expressing a negative-passive feeling contained more back vowels. The distinction is less clear-cut when positive-passive poems are included. Though these contained more closed-front vowels, as predicted by the hypothesis, results for open-front and for back vowels (open and closed) did not conform to expectations. Thus, the frequencies of back and front vowels seem to be related not primarily to positive vs. negative feelings, but rather to position along the hybrid axis of positive-active vs. negative-passive.
5. Discussion

In this paper, findings from a research study on Synesthetic Sound Iconicity have been presented. Based on previous research, it was hypothesized that plosives are associated with feelings of high arousal, whereas continuants are associated with feelings of low arousal. Similarly, it was hypothesized that sad feelings are positively associated with back vowels, while happy feelings are more often expressed with front vowels. To test these hypotheses, sixteen poems were randomly chosen from a book of German poetry and their emotional tone was assessed by readers. Reader assessments were compared with the relative occurrence of phonemes, categorized according to articulatory characteristics. Results in general confirmed the hypotheses, suggesting that the relationship between sound and meaning is not as arbitrary as generally assumed.

Though the differences in phoneme frequency among the groups of poems are rather small, they are relatively consistent over all poems, suggesting that there is a systematic connection between the articulation of phonemes and their potential to express feelings. While this research involved the rating of 16 randomly chosen German-language poems by over 400 readers, similarly extensive studies have not been done for other languages. Still, taken together with previous findings (Albers, 2008; Auracher et al., 2011), the results of the current study support the view that language is – at least partly – iconic. One reason for iconic features in language may lie in the history of its development. There is a tradition of claiming that language evolved when our forefathers started to name objects by imitating aspects of their appearance with articulatory movements. (For a recent discussion of this hypothesis, see Ramachandran & Hubbard, 2001.) This argument is based on the assumption that there are orderly and non-arbitrary associations across sensory modalities, which allow for the expression of aspects of visual perception, such as shape or color, through variations in sound.
Several studies have found evidence for such synesthetic associations (for an overview see Cytowic & Eagleman, 2009), suggesting that brain areas for specific sensory modalities are connected with each other. Such neural pathways could explain why Synesthetic Sound Iconicity is still found in modern language today, for example in the expression of feelings in poems. Certainly more research in this field needs to be done. Thus far, insights into the relationship between emotional tone and phonetic structure have been based almost solely on poetic language. What we don’t yet know is whether the results can be transferred to other domains of language use. This question is closely linked to the search for a neural basis of Synesthetic Sound Iconicity. Though some research has been conducted with onomatopoeic expressions in Japanese (Arata, 2010; Osaka, 2004), this researcher is not aware of studies that apply neuroimaging to non-imitative phonosemantics (i.e., to Synesthetic Sound Iconicity). More detailed analysis of phonetic structure is also called for. Does it matter, for example, whether the phonemes occur in stressed or unstressed syllables? Finally, one might want to ask whether insights from Synesthetic Sound Iconicity can be applied to automatic text analysis. Is it possible, for example, to detect works or passages expressing sad or happy feelings from a text corpus consisting of several thousand texts, only by assessing the frequency of certain phonemes?

Acknowledgments

This study was financially supported by the Japanese Ministry of Education and Science (科学研究費補助金). I wish to thank Dr. Uwe Reichel from the University of Munich for the phonetic transcription of the poems. I also wish to thank Ms. Aysin Nar (MA) for the distribution of the questionnaires.
References


