Symmetry as Iconicity: 
The Lexicalizations of ‘Breasts’

Vladimir Borissov Pericliev

Bulgarian Academy of Sciences, Bulgaria

Abstract

A word is said to be *iconic* if there is a perceived similarity between its form and its meaning. Research on lexical iconicity has shown that it pervades most sections of vocabulary structure but is scarce in regards to mimicking abstract properties of referents. In this paper, we propose the abstract property of symmetry as iconic, illustrating it with the lexicalizations of the meaning ‘(woman’s) breasts’ in world languages. Statistical analysis shows that the frequency of symmetric lexemes corresponding to the symmetric notion ‘breasts’ in the world languages by far exceeds the frequency that could be expected by mere chance. We also adduce further evidence from
symmetric languages, of phonological and semantic nature, that further supports our proposal of the existence of symmetric iconicity. In particular, some languages violate a regular sound change just in order to preserve the iconicity of the original word, others develop by analogy symmetric synonyms, and still others may falsely reanalyze a symmetric form in two identical parts just in order to reflect their perception of the meaning as symmetric. The presented arguments make a strong case for the proposed notion of symmetric iconicity.

Keywords: lexical iconicity, visual symmetry, lexical symmetry, violation of sound changes, false reanalysis

1. Introduction

A word is said to be *iconic* if there is a perceived similarity between its form and its meaning. If there is no similarity between form and meaning, the word is arbitrary, or conventional.

There are different types of iconicity (cf. e.g., Hinton et al. 1994a), the most popular probably being *imitative iconicity* involving onomatopoeia, in which the string of sounds of the word directly imitates the sounds produced by the referent of the word (e.g., animal sounds like *meow, moo*, or the sounds made by objects upon contact with other objects like *bang, crack, crash*). *Synesthetic iconicity* expresses non-acoustic information by speech sounds. Synesthesia deals with cross-modal sound-meaning associations. In synesthetic sound symbolism certain vowels, consonants, or their combinations systematically represent sensory, motor, or affective experiences or characterize aspects of the spatio-temporal unfolding of an event. A celebrated example is the association of the sound [i] with small size and sharpness and [u] with largeness, softness and heaviness (cf. Sapir 1911, 1927; Jespersen 1933; Ohala 1984). *Conventional iconicity*
Vladimir Borissov Pericliev 131

deals with frequent pairings of phonemes and aspects of meanings in specific languages (e.g., *gl-* in English *glisten, glimmer*, etc.). Yet another type of iconicity is *ideophones* (or *mimetics*), which are marked words depicting sensory imagery (e.g., Dingemanse 2012), as e.g., the Japanese *goro* ‘heavy object rolling’ and *koro* ‘light object rolling’, where the voicing of an initial consonant indicates the size of the object involved.

Though in contemporary linguistics from structuralism on it is generally believed that vocabulary is largely arbitrary, the last few decades witnessed an ever-growing body of literature on lexical iconicity (cf. e.g., Hinton et al. 1994b; Voeltz & Kilian-Hatz (eds.) 2001; Ahlner & Zlatev 2010; Perniss et al. 2010; Wichmann et al. 2010; Dingemanse 2011, 2012; Monaghan et al. 2014; Dingemanse et al. 2015; Blasi et al. 2016; Osoba 2018; Arokoyo & Lagunju 2019; Motamed et al. 2019; Sobola 2019; Cabrera 2020; Johansson et al. 2020; McLean 2020 and the extensive references therefrom) showing that iconicity is not limited to specific pockets of vocabularies, but rather encompasses practically all aspects of the vocabulary of the world languages. Thus, a wide variety of referents’ properties are used in the construction of iconic words: imitation of sounds emitted by the referent, salient properties of the referent (such as shape, size, colour, texture), expression of affection, evaluation and social status, etc. (Hinton et al. 1994a: 10–11, Dingemanse 2012: 663–664, McLean 2020). These cover practically all aspects of vocabulary, with the notable exclusion of the iconic employment of abstract referent’s properties. Hinton et al. (1994a: 10) e.g., write in this context that “Only abstract relational notions (such as categories of even and odd numbers) seem to be sparse in sound-symbolic representation.” Actually, we are not aware of *any* example of lexical iconicity involving such abstract notions.
The goal of the present paper is to propose the abstract property of symmetry as an iconic form-meaning mapping: a visual symmetry of the referent is reflected in a symmetric form of the word denoting it. The phenomenon is illustrated with the lexicalizations of the meaning ‘breasts’ in world languages.

The paper is organized as follows. Section 2 explains symmetry in visual patterns and in word forms and gives some examples of “symmetric words”, i.e. words denoting symmetric objects and having symmetric forms. It introduces the ASJP¹ database (Wichmann et al. 2013) from which our examples are taken and shows that the frequency of symmetric words denoting breasts is far greater than could be expected by chance. Section 3 examines further circumstances of phonological and semantic nature substantiating the existence of a linkage between form and meaning in symmetric words for ‘breasts’ in specific languages. Finally, in Section 4, we recapitulate the results and place symmetric iconicity among other types of iconicity.

2. Symmetric Iconicity in the Lexicalizations of ‘Breasts’

In this section, we show that the frequency of symmetric word forms associated with the symmetric meaning ‘breasts’ by far exceeds the frequency that should be expected by mere chance. We interpret this preference as iconicity: native speakers in the respective languages conceive ‘breasts’ as symmetric and accordingly choose symmetric forms to reflect this fact.

¹ The following abbreviations are used in this paper: ASJP (Automated Similarity Judgment Program), PMP (proto-malayo-polynesian).
Let us first look at the notion of symmetry. In everyday parlance, *symmetry* is used to mean something which is well-proportioned, well-balanced and generally beautiful (Weyl 1952: 3). In mathematics, symmetry has a more precise meaning concordant with the basic Euclidean transformations in plane geometry. Two types of symmetry, which will be of interest to us, are reflectional and translational symmetry. Thus, an object has *reflectional symmetry* (also called *mirror-image, bilateral or left-right*) if there is a line going through it which divides it into two pieces that are mirror images of each other. An object has *translational (or repetition) symmetry* if it can be translated (moving every point of the object by the same distance) without changing its overall shape.

Symmetry can be found in a wide variety of objects, both natural and artificial. Some objects have reflectional, some repetition symmetry, and some both types. The notion ‘breasts’ we are investigating here has both. Thus, with ‘breasts’, one breast can be viewed as a mirror-image of the other or one breast can be viewed as a repetition of the other.

As it is well-known from Gestalt psychology (e.g., Köhler 1929) and contemporary cognitive science (e.g., Wagemans 1997, Tyler 2002, Bertamini 2010) symmetry is a salient property of objects and the human visual system efficiently and robustly detects symmetry. So the question naturally arises whether this symmetry is accordingly reflected in natural language vocabularies.

Visual symmetry can be directly imitated in strings of sounds (or their graphical representations), even though sound and vision are different modalities. Thus, mirror-image symmetry in word forms is found in palindromes (words that read the same from left to right and from right to left.). Translation, or repetition, symmetry is present in various types of segment or syllable repetition. More specifically, we define mirror-image symmetry as palindromes, excluding single
sound word forms, which do not seem to fit our intuitions of symmetry. We define translational (repetition) symmetry in word forms as follows. If a form divides into two sub-strings, which are identical, it has repetition symmetry (e.g., ka-ka). Also, we allow some slight perturbations of this ideal pattern, in which the word-form would also be perceived as symmetric. For our (conservative) approach we assume a perturbation of just one additional segment, occurring either at the beginning (e.g., a-ka-ka), the middle (ka-r-ka) or end (ka-ka-r) of the word string comprising two identical sub-strings.

Below are examples illustrating the variety of symmetric patterns lexicalizing the meaning ‘(woman’s) breasts’ in different languages. The examples come from the ASJP database (Wichmann et al. 2013). (The ASJP phonetic abbreviations used are as follows: C = voiceless palato-alveolar affricate, 7 = voiceless glottal stop, N = velar nasal, 5 = palatal nasal, x = voiceless and voiced velar fricative, 3 = high and mid central vowel, rounded and unrounded).

*Reflectonal symmetry in forms meaning ‘breasts’*:

- **[uu]** Ansus (Austronesian, Western New Guinea, Indonesia)
- **[ele]** Liuwa (Niger-Congo, Zambia)
- **[pap]** Yitha Yitha (Australian, South Australia)
- **[dud]** Zapotec Mitla (Oto-Manguean, Mexico)
- **[amama]** Ogbronuagum (Niger-Congo, Nigeria)
- **[korok]** Lamalera (Austronesian, Indonesia)
- **[titit]** Marshallese (Austronesian, Marshall Islands)
- **[ususu]** Saaroa (Austronesian, Taiwan)
- **[Cu7uC]** Itzaj (Mayan, Guatemala)
- **[tabubat]** Tashelhit Ida Usemal (Afro-Asiatic, Morocco)
*Translational symmetry in forms meaning ‘breasts’*:

- [nee] Nea Nooli (Austronesian, Solomon Islands)
- [CuCu] Quechua Imbabura (Quechuan, Ecuador)
- [popo] Niupo Gelao (Tai-Kadai, China)
- [goNgo] Gaagudju (Australian, Northern Territory)
- [otete] Pilaga (Guaycuruan, Northern Argentina)
- [koykoy] Allentiac (Huarpian, Chile)
- [aloalo] Pukapuka (Austronesian, Northern Cook Islands and other regions)
- [t3pt3p] Burum Mindik (Papuan, Morobe Province, Papua New Guinea)
- [hatahata] Vaeakau Taumako (Austronesian, Solomon Islands)
- [raqoraqo] Bilua (Papuan, Solomon Islands)
- [birmbirm] Woiwurrung (Australian, Victoria)

As an empirical basis for our study we use the ASJP database (Wichmann et al. 2013), including 100-item list of basic meanings in over 7,000 languages. Our study excludes creoles, pidgins, mixed and constructed languages, which yields 6,809 languages. Since the 100-item wordlists are not exhaustive for all languages, we focused on forty basic meanings out of the 100-item lists. We thus base our study on the 40-item list, which includes the meanings: blood, bone, breast, come, die, dog, drink, ear, eye, fire, fish, full, hand, hear, horn, I, knee, leaf, liver, louse, mountain, name, new, night, nose, one, path, person, see, skin, star, stone, sun, tongue, tooth, tree, two, water, we, and you (sg).

We now turn to investigating whether the frequency distribution of symmetric forms is random across the 40 meanings (our null hypothesis) or there is a universal trend of preference of some meanings to be associated with symmetric forms (our alternative hypothesis). To accomplish this task we use the chi-square goodness-of-fit test, measuring the statistical significance of the difference
between the actually occurring frequencies and those that could be expected by mere chance. If this difference is close to zero or very small, the null hypothesis of randomness should be accepted; if the difference is large, the null hypothesis of randomness should be rejected and the alternative hypothesis of preference accepted. Thus, there are 232,840 word forms for all meanings in our 40-item list, 10,104 of which turn out to be symmetric (with either reflectional or translational symmetry, to be referred to further on as just “symmetric”). The probability $p_s$ of a symmetric form to be associated with some meaning is therefore $10,104/232,840 = 0.0433$. This probability $p_s$ is used to compute the expected number of symmetric forms to be associated with a specific meaning by the formula $n \times p_s$, where $n$ is the total number of word-forms, both symmetric and non-symmetric, associated with some meaning in our database. Then the expected frequencies of symmetric forms for each meaning are computed and the chi-square test applied for determining significance (for each meaning, the squared difference between observed and expected frequencies are divided by the expected frequency, all results are summed and the resultant number tested in a table of critical chi-square values to assess significance). In our particular case, the test yields a highly significant result ($\chi^2 df = 39, p = 0.000$), or the probability is zero with accuracy up to three decimal numbers, meaning that the hypothesis of randomness should be rejected and the one of the existence of a universal trend of preference accepted. The basic culprit for this non-randomness is the symmetric meaning of ‘breasts’, which conspicuously prefers symmetric forms for its lexicalization with actual frequency of 1,187 symmetric forms against expected frequency of only 253. From 5,840 lexemes for ‘breasts’ in our database 20% are symmetric while less than 5% should be expected by chance. (We note in passing, that there is another symmetric category from the 40 examined ones, viz. ‘tooth’, which is
also associated greater than chance frequency of symmetric forms, though in a less marked manner, with 391 observed against 280 expected occurrences.)

Within the framework of the observed universal tendency, some language families also manifestly prefer symmetric forms for expressing ‘breasts’. Notable among these is Austronesian, which is a very large family and hence could potentially bias the outcome of the test described above. To eliminate this bias, we entirely excluded Austronesian from the database and rerun the statistic test. Again, the frequency distribution of symmetric forms across meanings turned out to be significantly non-random, for ‘breasts’ the observed frequency being 444 and the expected only 179. By way of another example of a family with exclusively symmetric forms for ‘breasts’ we may mention Quechua, all of whose 38 lexicalizations are symmetric. A family like Indo-European has 28 symmetric forms, only two of which are associated with European (and extinct) languages (Latin [mama] and Old Church Slavonic [susu]), though it partly compensates this with a number of such forms occurring in informal speech in European languages outside our database (e.g., English boob, chi-chis, tits; Spanish mama, chichi, Bulgarian cici, etc.).

3. Further Justification

Different types of iconicity require different justifications. Onomatopoeia e.g., usually requires little justification since the word-form directly imitates the sound emitted by the referent. In other types of iconicity the linkage between form and meaning is somewhat less transparent, but still tangible. E.g., for a familiar case of synesthethic iconicity, Martinet (1965: 231) writes that the familiar symbolic
equations [i] = smallness and [u] = largeness have an obvious physiological foundation (small vs. large resonating cavity used for their production), hence one need not be a great scholar of articulatory phonetics to understand the reason for such identifications. Here, again, we have imitation, though of a less direct manner, viz. one pertaining to the action of speech organs and the referent. In still other cases, the form-meaning similarity is more difficult to understand and hence more problematic. For example, Marchand (1959: 147) notes the prevalence of /k/-/p/ and /k/-/b/ in words referring to “protuberant forms”, such as knap and knob, but acknowledges that we cannot ascertain why the mapping would be iconic. The similar problem we encounter in justifying the iconicity of some ideophones (Ahlner & Zlatev 2010, Dingemanse 2011). A case in point is e.g., the Japanese ideophones koro-koro ‘small object rolling continuously’ and goro-goro ‘large object rolling continuously’ (Kita 1997), where it is not clear in what sense the meaning contrast ‘small vs. large’ is similar to the sound contrast [k] vs. [g]. The linkage between form and meaning is even vaguer in ideophones expressing mental states (cf. Ikegami & Zlatev 2007).

The symmetric iconicity we propose in this paper is well justified. First, it is imitative in nature in the sense that the symmetry of the word form apparently mimics the symmetry of the referent, and this is essential to iconicity. And, secondly, there is a statistically significant preference of association of the symmetric notion of breasts with symmetric forms, which can be reasonably interpreted as iconicity, in the sense that the forms are so chosen as to fit the meaning. As further evidence strengthening our thesis below we offer some additional phonological and semantic considerations attesting to this argument in regards to some specific languages.
3.1. Symmetry-Preserving Sound Changes

Linguists from Grassmann on have observed that regular sound changes often do not affect onomatopoeic words (for illustrations from Huastec, Nez Perce, Finnish and Modern Greek, cf. e.g., Hinton et al. 1994b). The rationale behind this linguistic strategy is apparently to preserve the iconicity of the word by preserving its original sound shape, which is conceived by the speakers of the respective languages as resembling the referent. Therefore presenting similar examples in the lexicalizations of ‘breasts’ can be viewed as further support for their iconicity.

Blust (2013: 625) reports the sound change *u > i in the final syllable of many Austronesian languages of the South Halmahera-West New Guinea group (e.g., Gimán, Buli, Waropen), in Bobot (Bonfia), of the central Moluccas, and Wetan of the Lesser Sundas. Consider data from Warembori (Donohue 1999), a South Halmahera-West New Guinea language, spoken by some 600 people on the north coast of Papua, Indonesia.

<table>
<thead>
<tr>
<th>Word</th>
<th>Original Form</th>
<th>Meaning</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>ki</td>
<td>*kutu ‘louse’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>muni-ro</td>
<td>*bunuq ‘kill’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>mini-no</td>
<td>*inum ‘drink’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>kuni</td>
<td>*tunu ‘cook’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>akuni (itr)</td>
<td>*tunu ‘cook’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>kuni</td>
<td>*tunu ‘burn’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>mani-ro</td>
<td>*manuk ‘bird’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>ki-ro</td>
<td>*kutu ‘louse’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>kati</td>
<td>*qasu ‘smoke’</td>
<td>(PMP)</td>
<td></td>
</tr>
<tr>
<td>mi</td>
<td>*kamu ‘you (pl)’</td>
<td>(PMP), but</td>
<td></td>
</tr>
<tr>
<td>ke-tutu-ro</td>
<td>*susu ‘breasts’</td>
<td>(PMP)</td>
<td></td>
</tr>
</tbody>
</table>
As seen above, Warembori consistently follows the sound change *u > i in the final syllable, with the exception of the last example, viz. the one involving the meaning ‘breasts’. Warembori violates this sound change in order to preserve the symmetry. The speakers of the language conceive ‘female breasts’ as symmetric and this requires a correspondingly symmetric form as a “proper” expression of the property.

3.2. Symmetry-Preserving Lexical Changes

Some lexical changes may also be indicative of iconicity. Below we mention two such cases.

First, some languages prefer symmetric forms for designating the symmetric notion ‘breasts’, as seen from their construction of synonyms which are also symmetric. Below are some examples:

- [mem, xux] Mongolian (Altaic)
- [dada, susu] Buginese (Austronesian, Sulawesi)
- [gugu, nuanua] Wedau (Austronesian, Papua New Guinea)
- [tamb~utamb~u, mimi] Ngawun (Australian, Cape York, Queensland)
- [fatafata, huhu] Tongan (Austronesian)
- [susu, boNaboNo] Minigir (Austronesian, Papua New Guinea)
- [omo, nono] Bamu (Papuan, Papua New Guinea)
- [ipi, mimi] Pitjantjatjara Yankuntjatjara (Australian, Central Australia)
- [5u5u, CiCi] Laraos (Quechuan, Peru)

It is interesting to note that Proto-Austronesian is reconstructed with as many as six distinct synonyms for ‘breasts’, all of which are symmetric in form [*susu, *zuzu, *dada, *5u5u, *NuNu, *nunu], showing an expressed preference for correspondence between form
and meaning. We note that under this heading we should not consider languages with symmetric synonyms, which are however historically derived from one another, e.g., [*susu, *huhu] ‘breasts’ Proto-Polynesian, where the second form is the results of the regular lenition s → h in Polynesian.

Secondly, reanalysis, or folk etymology, provides precious insights into the linguistic intuitions of native speakers. A symmetric word-form in a language may be decomposed into two identical parts, which is iconic of a symmetric notion denoting an object comprising two identical parts. The following are examples from the authoritative Austronesian Basic Vocabulary Database (Greenhill et al. 2008) for words denoting ‘breasts’, graphically showing the decomposition:

[tot-tot] Mentawai (Austronesian, Northwest Sumatra-Barrier Islands)
[dub-dub] Surigaonon (Austronesian, Central Philippine)

Remarkably, some languages may decompose the original word from the parent language in two independent parts, as [so so] Bicol (Buhinon) (Austronesian, Central Philippine) (from PMP *susu).

4. Conclusion

Previous research has shown that iconicity practically encompasses all aspects of human language lexicons with the notable exception of abstract properties of referents. In this paper, we propose symmetric iconicity as such a type of iconicity: symmetric notions are expressed by symmetric word forms. We illustrated this iconicity with the lexicalizations of the notion ‘breasts’ in the world languages. It was
shown that the frequency of symmetric word forms for this meaning by far exceed the frequency that could be expected by chance. We also discussed some phonological and semantic circumstances in specific languages that also attest to the thesis that the speakers of the respective languages indeed conceive of these words as iconic. Thus e.g., some languages may violate a regular sound change just in order to preserve the iconicity of the original word, others may develop by analogy symmetric synonyms, and still others may falsely reanalyze a symmetric form in two identical parts just in order to reflect their perception of the meaning as symmetric. All these, taken collectively, are strong arguments in favour of the soundness of the proposed notion of symmetric iconicity.

Finally, a word about the place of symmetric iconicity amongst other types of iconicity. Symmetric iconicity is synesthetic because it involves a cross-modal mapping, viz. a mapping between visual images and sounds. At the same time, in sharp contrast to most examples of this sound symbolism, the linkage between form and meaning is not intangible, but is rather direct, the symmetry in the designated object being reflected in, or mimicked by, the symmetry of the word form; in other words, symmetric iconicity is also imitative in nature. As already mentioned, an essential difference between symmetric iconicity and all other types of iconicity is the semantic realm of this iconicity, viz. involving the abstract property of symmetry. Finally, like ideophones, it concerns whole words, not individual segments (or clusters of segments).

References


Perniss, P. et al. 2010. Iconicity as a General Property of Language:
Evidence from Spoken and Signed Languages. *Frontiers in Psychology* 1, 227.


