Bartosz T. Wojciechowski

Japanese Ideophones – towards a Systematic Phono-Semantic Description

ABSTRACT

Ideophones are not unique to Japanese, nor typologically uncommon, but, since largely absent from Europe-centered languages, they pose a number of questions to linguists traditionally not accustomed to (or simply oblivious of) them. The paper briefly outlines the definition and characteristics of ideophones in Japanese, including their internal typology and difference from onomatopoeia. However more important point to be made was whether this group of words can be described in terms of correlation between their phonological structure and meaning. Some regularities seem to exist, and attempts to outline these regularities have been made. This, in turn, brings up the age-old discussion of arbitrariness of language signs.

KEYWORDS: ideophones, mimetics, sound symbolism, giongo, gitaigo, onomatopoeia, l’arbitraire du signe, phono-semantics.

0. Introduction

One of the most striking typological characteristics of the Japanese language is the abundant presence of ideophones, i.e. lexemes in the like of pera-pera, chanto, or yukkuri, which appear in various kinds of texts, both spoken and written, but are particularly numerous in the colloquial, informal variant of Japanese. This lexical stratum appears to be particularly difficult to describe, compared to other strata or lexical groups. The key difficulties include the:

i. precise semantic definition of each lexeme;
ii. degree of adherence to particular speech registers, also functional limitations thereof;
iii. classification of ideophones as speech parts, their role in syntax;
iv. precise definition of the relationship between the form (phonological structure) and content (meaning).

This paper aims to outline the last problem. My objective is to try to address the question whether a regular description of the semantic structure of this lexical group in the Japanese language based entirely on its phonological structure is possible. It appears that there are certain constant
regularities between the phonological structure of ideophones and their meaning.
The basic assumption is that the exact phono-semantic description (i.e. description of the phonetic-semantic regularities) would enable us to foresee the meaning of a given lexical item, and, conversely, to devise a lexical item that conveys given meanings. It appears, however, that such a presumption is possible only to a certain extent, as the exact semantic definition seems to be unlikely, even with the most enthusiastic attitude to the validity of phono-semantic analysis. The degree of such certainty remains yet to be theoretically ascertained. Another theoretical value of phono-semantics would be the possibility to gain insight into the mechanisms of sound imaging (i.e. mapping MEANING → SOUND, thus SUBSTANCE → FORM) among native speakers of the Japanese language. Such a study, conducted in a broader scale, and based on various languages, would possibly enrich our understanding of how human language works. It could also provide a basis to build hypotheses on the origin of language as we know it.

1. Typology
The presence of ideophonic vocabulary has been discovered to exist in a great many of the world’s languages (perhaps we could even venture upon a universalist hypothesis that has ‘all’ instead of ‘many’ in the previous sentence), yet the reasons behind their hugely disproportionate role across the world’s languages remain unknown. We can only state that there are certain linguistic areas where languages make extensive use of this kind of linguistic devices on a much wider scale than elsewhere. These areas include e.g. Central Africa, Southeast Asia, and East Asia (here especially the Japanese and Korean languages).
The abundance of ideophones in certain linguistic areas seems to be governed by mechanisms of linguistic osmosis, similar to a Sprachbund – they often appear in neighbouring, yet unrelated languages, thus their relevance in a certain system may be attributed to linguistic convergence, not a genetic relationship. Similarly to other rare typological phenomena, the extensive use of ideophones has an insular character on the language map of the world.
Despite the common occurrence of ideophones in many world languages, they differ not only with their scope (in some languages they are marginal, in others they are one of the core lexical categories), but also with the language registers they belong to. In such languages as English or Polish, which have a rather modest repertoire of ideophones, lexemes of this kind
belong to rather lower registers: they are markedly colloquial (e.g. Polish gadu-gadu ‘small talk’, English okey-dokey, tip-top, wishy-washy), intimate (e.g. Polish łapu-capu ‘helter-skelter’, English goody-goody), jocular (e.g. Polish fiksum-dyrdum ‘crazy’, English rumpy-pumpy), childish (e.g. Polish szuru-buru ‘to wash’, English itsy-bitsy, teenie-weenie) or even vulgar (e.g. Polish pierdu-pierdu ‘to talk gibberish’, English arty-farty). Compared to this, Japanese ideophones pertain to a much wider set of registers, most importantly including the neutral one and the written language (however, they seem to be largely absent from the lexicon of the highest registers).

2. Definition and Criteria
The formulation of a concise, yet adequate, unambiguous and language-independent definition of the ideophone proves to be unexpectedly difficult. Clement Doke, who coined the term ideophone, is widely quoted to have defined it as follows:

*A vivid representation of an idea in sound (…) a word, often onomatopoeic, which describes a predicate, qualificative or adverb in respect to manner, colour, sound, smell, action, state or intensity.* (Doke 1935: 118)

This definition is, however, far from being precise. It is hard to have such an ambiguous explanation as *vivid representation* as part of a usable definition. The same applies to being *often onomatopoeic* (the difference between ideophones and onomatopoeia will be discussed later) or *describing a predicate, qualificative or adverb* (it is hardly a feature unique to an ideophone).

Many other definitions, which in turn focus on particular languages or language groups, try to define them as a specific group of speech parts or at least a lexical class.
It seems that a proper definition cannot avoid attempts to describe this group as seen from various levels of language structure. I would not attempt to formulate a general definition, but let us turn our attention to the different properties of ideophones in the Japanese language.
The most salient criterion seems to be *semantics*. From the semantic perspective we may thus define ideophones as marked language signs that reflect impressions and sensory impulses. They are further characterized by their additional emotive and perlocutory value (i.e. the sender has the intention to make the receiver feel particularly engaged emotionally in the
content and context of the given message; much more than would be the case without using such linguistic means). Their validity in the information structure of the message seems to be based on redundancy — an ideophone is more frequently a carrier of an additional emotional/perlocutory charge than a vital part in the information wireframe.

In other words, it is quite often the case that ideophones can be left out of a sentence without making it nonsensical, even without losing vital parts of the message. Using ideophones, thus, performs a phatic role — they help to build an emotional liaison between the sender and the receiver, they also convey the speakers’ will to maintain and add depth of emotion to communication. Such is, tentatively, their function on a **pragmatic** plan.

The criterion of **morphology** appears to be important as well, as far as the Japanese language is concerned. It also proves to be quite useful to sort out ideophones as a clearly separate category of the lexicon. Japanese ideophones are characterized by their remarkably strict adherence to three separate patterns of phono-morphological structure, to which the overwhelming majority of Japanese ideophones belong. They may be outlined as follows (Latin letters A and B stand for separate morae, constant parts are written using Japanese syllabary *hiragana*):

| PATTERN | EXAMPLES¹ |  |
|---------|-----------|  |
| i. AB AB | にこにこ *niko-niko* | どきどき *doki-doki* |
| ii. A っ B り | にっこり *nikkori* | どっきり *dokkiri* |
| iii. AB (っ)と | にこっと *nikotto* | どきっと *dokitto* |

³[to/with a] smile’ ‘irritation, excitement’

The iii. group exhibits the incorporated enumerative particle *to*, which is often geminated.

Ideophones may also be part of compound and derivates (Saito Hamano 1986: 51).

In the case of Japanese ideophones, **syntax** does not seem to be a vital criterion. They usually serve as a noun modifier or adverbial, modifying verbal or nominal parts of a sentence. Less frequently they also modify adjectives. Sometimes they can also act on their own as independent

---

¹ I admit that the examples cited here are selected rather subjectively for the table to illustrate without empty spaces all three main structural patterns, retaining basically the same meaning in both vertical columns. This is not supposed to suggest that all the three patterns are equally productive in terms of derivation. Just the opposite: it is quite uncommon that same-meaning lexemes are present in all three structural patterns. However, this is not to mean that these categories are not valid — they are, but in most cases empty spaces will occur in all but one cell.
predicates (with the default auxiliary verb *suru*, cf. *guzu guzu suru* ‘to act indecisively, to be slow, to move slovenly’, *unzari suru* ‘to be fed up with’), or, particularly in informal speech, they can serve as sole predicates, with auxiliary verbs omitted (cf. Saito Hamano 1986: 17-18). However, they do not act as nouns, i.e. as a subject or object. From a syntactic point of view, ideophones show properties typical to other speech parts — usually adverbs, *na*-adjectives and verbs, but never or extremely rarely as nouns. Japanese ideophones have, nevertheless, specific phonemic, phonetic, and phonotactic properties uncommon for other layers of the Japanese lexicon (by which I mean three traditionally recognized lexical strata, defined in genetic terms, i.e. *wago* – native Japanese vocabulary; *kango* – Classical Chinese loanwords, borrowings from other languages, mostly East-Asian, via Classical Chinese subsystem, and Japanese coinages using Chinese morphemes; and *gairaigo* – lexemes borrowed without the intermediary role of the Sino-xenic subsystem in Japanese, usually from European languages, like Portuguese, Dutch and German, but most often English, also native coinages with such morphemes).

The phonetic peculiarity of ideophones has led some authors (cf. e.g. Tsujimura 1996: 147) to identify not three, but four layers of the Japanese lexicon. The fourth one, apart from *wago*, *kango*, and *gairaigo*, being precisely ideophones. The criterion of such a classification (i.e. ideophones being on equal status with the other three) must however be different than in the case of the others — from the etymological point of view they are native vocabulary, not borrowed vocabulary.

However, they have certain quite unique phonetic characteristics, different from other layers (especially within the native stock), such as the presence of:

i. [p] in word-initial position;
ii. non-geminated [p] in middle (i.e. inter-vocalic) position;
iii. [aː], which otherwise features only in *gairaigo* words, and within *wago* only rarely: in few family terms and the pronoun ā ‘that way’;

---

2 Interestingly, we can observe a striking difference here with the Korean language, which otherwise has a very similar system of ideophones to Japanese. In Korean, nouns are also frequently derived from ideophones (Sohn 2001: 98-101). Note that in Japanese there is a number of onomatopoeia-motivated nouns, as is the case with probably any language.

3 To be precise, we must note that the Classic Japanese subsystem is also sometimes classified as another stratum of Japanese lexicon (cf. Huszcza et al. 2003: 126). It serves as a relic subsystem in contemporary Japanese, e.g. in titles of books, feature films, proverbs, classical citations, certain grammatical constructions etc. Such a subsystem is thus defined not in genetic, but grammatical and syntactic terms. If we were to accept such a classification, the Japanese language would have five different (yet heterogenically defined) subsystems.
iv. [N] in syllable coda after a long vowel;
vi. geminated consonant after a long vowel.

3. Ideophones and Onomatopoeia: Key Differences
I must at this point make an important distinction between various groups of Japanese vocabulary that imitate external states by means of speech sounds. The following distinction seems to be the most salient: onomatopoeia (or phonomimes) vs. ideophones. The crucial difference between them is based on what such lexemes try to emulate. The former merely imitate human auditory impulses (i.e. audible sounds) of external world, whereas the latter are phonetic expressions of non-auditory emotions and impulses. The difference is therefore also a difference in iconicity. The sound-to-meaning relation in the former group is clearly a naturalistic, mimetic one, while in the case of ideophones this relation is far from being apparent.

We can draw the distinctions further down: onomatopoeia can imitate human voices (we lack a corresponding English term, although there is a precise Japanese term 擬声語 giseigo, lit. ‘words that imitate [human] voices’) or other sounds, namely animals, natural phenomena and manmade machines (Japanese term is 擬音語 giongo, lit. ‘words that imitate sounds’).

Ideophones, in turn, can be further classified into psychomimes (Japanese 擬情語 gijōgo, lit. ‘words that imitate emotions’), i.e. lexemes that describe human mental states and impressions, and phenomimes (Japanese 擬態語 gitaigo, lit. ‘words that imitate states’), which are impressions of other senses than hearing expressed by means of language sounds. An English term adopted sometimes (e.g. Tsujimura 1996: 93) to embrace both ideophones and onomatopoeia is mimetics, but elsewhere it is quite often used to mean ideophones only. The nomenclature is still far from being consistent in this subject, to say the least.

The IDEOPHONE : ONOMATOPOEIA distinction is also important from a typological point of view — these two groups appear to have quite different properties in various languages or language areas. While the number of onomatopoeia is rather constant across languages, their function within corresponding systems (they belong to similar registers and sociolects, and even have similar sound structure) is also similar. Ideophones differ not only in number and function in different languages, but also their position within a given language, sound structure and morphology.
Ideophones are also interesting from another point of view: by analyzing cognitive mechanisms involved in their usage and functioning, we encounter numerous questions that still remain to be answered, and some hypotheses contradict traditional paradigms when reflecting about language.

The common characteristics of both groups seem to be the provenience of the lexemes. These language units differ from every other lexicon item in one important point: while every lexeme of every natural language is determined to come from certain previously existing entities (their own proto-language or a certain source-language), ideophones and onomatopoeia seem to be rooted in simply imitating the external world by means of speech sounds. We may thus dub this provenience “extra-linguistic”, “naturalistic”, or even “un-etymological”.

An important reservation must be made here. In contemporary Japanese there is a substantial group of ideophones that seem to be connected in a certain way (because I would hesitate to refer to them decidedly as to “be derived from”, as the direction of word-formation is not entirely certain) with lexemes in other groups, primarily native wago words (cf. Saito Hamano 1986: 6; 52-), e.g.:

<table>
<thead>
<tr>
<th>IDEOPHONE</th>
<th>MEANING</th>
<th>LEXEME</th>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>kira-kira</td>
<td>‘glitter, sparkle’</td>
<td>明らか</td>
<td>‘clear, plain’</td>
</tr>
<tr>
<td>koro-koro</td>
<td>‘rolling’</td>
<td>転がす</td>
<td>‘to roll’</td>
</tr>
<tr>
<td>guru-guru</td>
<td>[going around] in circles’</td>
<td>車 kuruma</td>
<td>‘car; wheel’</td>
</tr>
<tr>
<td>hissori</td>
<td>‘quiet, still, silent’</td>
<td>密か hisoka</td>
<td>‘secret, private’</td>
</tr>
<tr>
<td>tera-tera</td>
<td>‘gleamingly’</td>
<td>照らす terasu</td>
<td>‘to shine’</td>
</tr>
<tr>
<td>nuru-nuru</td>
<td>‘slimy, slippery’</td>
<td>濡れる nureru</td>
<td>‘to get wet’</td>
</tr>
</tbody>
</table>

may be associated with

4. Terminology
The systematic comparative and typological research of ideophones is relatively advanced in Bantu and Southern Semitic linguistics. These are the language areas that are comparatively rich in such phenomena, often being vital parts of the language systems used there. Linguistic terminology is therefore somewhat rooted in works concerning African linguistics. Otherwise, we must admit that it is rather far from being clear and unanimously accepted.

As for the general term for ideophones, various coinages have been proposed. These include phon(a)esthemes (John Rupert Firth, 1930), impressives (Maurice Durand, 1961), descriptives (Neil Smith, 1973), expressives (Gérard Diffloth, 1976), or phonosemics (Roger W. Wescott, 1980). Diego Collado, the Spanish author of an early Japanese grammar,
Ars grammaticae Iaponicae linguae, published in 1632, has called them *adverbia sonus*, literally ‘adverbs of sound’. In French there is a term *mimologique*. In the English language, except for *ideophones*, also *sound symbolism*, *sound iconisms*, *phonetic symbolism*, and *mimetic words* are frequently found. The Japanese-language terms (introduced above) are relatively stable and precise. They are commonly used not only in scientific texts, but also in numerous how-to books and dictionaries targeted at a general audience. I have settled upon *ideophones* for two reasons, as this term is:

1. relatively widely accepted in English-language texts
2. surprisingly precise, aggregating the key dualism of two aspects of this group, namely the unique fusion of content (Greek ἴδεα, *idea*) and sound substance (Greek φωνή, *phone*)

Despite various terms, we must stress the importance of differentiating onomatopoeia from ideophones, as they are fundamentally different language phenomena. Such a distinction is seldom made in the practical description of the Japanese language (most often, for teaching purposes) and in practical usage during Japanese classes. They are all commonly called *onomatopoeia* in English. At least, this is the case of Japanese language teaching in Poland. Such confusion can be seen as a hindrance to the proper understanding of such an interesting phenomenon as the ideophone.

5. Discussions Surrounding the Arbitrariness of Language Signs
Linguists seem to be interested in ideophones in quite a disproportionate way to the role played by this lexical group. This is the point in the human language that seems to show regular bonds between the substance of content and form of expression (in Hjelmslev’s terms), thus being not entirely arbitrary language signs. Such arbitrariness has been the cornerstone of modern thought about language and linguistics, especially in its structuralist canon, as first outlined by Ferdinand de Saussure (who introduced the well-known rule of *l’arbitraire du signe*).

Since Ancient Greek philosophy, there have been two different approaches to how language signs are motivated. They are concisely outlined in Plato’s dialogue *Cratylus*. The "naturalist" approach (also known as the φύσει theory, lit. ‘by nature’) claimed that any name existing in a language was the most suitable label for any objects it denoted. Their phonetic structure is not accidental and there is a phonetic regularity to the very nature of the
object. Note that this approach must inevitably have taken for granted the inequality of different languages. This is hardly surprising — the Ancient Greeks viewed their language as being “civilized”, whilst other languages were considered much less civilized, even uncivilized, and therefore barbaric.\textsuperscript{4} If that was not the case, it would have been impossible to explain the different sound structures in semantic equivalents in other languages. Although Ancient Greek naturalists are sometimes quoted to have uttered such generalized statements as “both in Greek and any other language”, they failed to explain this obvious contradiction.

Naturalists were opposed by what are today called “conventionalists”. They held all names to be purely conventional, arbitrary and devised in an essentially random and consensual way (this approach is also known as the νόµῳ theory, lit. ‘by law’, or θέσει theory, lit. ‘by assumption’).

Modern linguists have generally rejected the former approach, maintaining the tenacious claim of the arbitrariness of language signs. There is hardly a contemporary scientist who believes that the nature of all objects has a direct influence on how their names sound.

The above-mentioned rule of sign arbitrariness usually accepts a few minor exceptions, e.g. pure onomatopoeia and close family terms. The latter are usually coined using the first words uttered by a baby in its life, thus having a simple phonological structure — they contain the easiest to pronounce speech sounds in a given language, which are acquired first in the lives of its native speakers. Most often these include labials like [m] or [b] and open vowels like [a].\textsuperscript{5} Syllables are very often reduplicated (repeating the same syllables is typical at the babbling stage in human evolution). Thus, the coincidental similarities between unrelated languages may emerge.

There were, however, natural scientists and thinkers who — just like Ancient Greek naturalists — did not deny some sort of connection between the phonetic structure of the word and its meaning. Even if we ignore fancy hypotheses verging between poetics, aesthetics, philosophy, and mysticism, such as Mikhail Lomonosov’s poorly argued idea that “meek” phenomena and objects often have [e], [i], and [‘u] vowels and “frightening” ones have [o], [u], and [i] vowels, such a correspondence has also been noted by Wilhelm von Humboldt (1836). He claimed that language tended to choose

\textsuperscript{4} Etymologically, itself an onomatopoeia bar-bar being the imitation of incomprehensible foreign speech – to babble.

\textsuperscript{5} Cf. the similarity between the Polish lexeme baba and Japanese lexeme婆 baba, both meaning ‘old woman’. Polish and Japanese are not related (as far as the present state of science is concerned), these particular lexemes are not mutually connected either (e.g. not being loanwords from either side or from a third party).
phonetic structures that make the same impression to the ear as the very object does. The formula “impression similar to the object” is the key concept for further, now psycholinguistic, research in the field. 

Otto Jespersen, too, was deeply convinced that the arbitrariness rule is not absolute:

*Sound symbolism, we may say, makes some words more fit to survive and gives them considerable help in their struggle for existence. (...) There is no denying that there are words which we feel instinctively to be adequate to express the ideas they stand for.* (Jespersen 1922: 408; bolding mine – BTW)

After the Second World War, many linguists undertook research in phonosemantics in natural languages, not limited to ideophones. They established, for instance, that the same sound groups appear in many etymologically unrelated, yet semantically somewhat proximate lexemes. Examples include English word groups such as *crick, cramp, crack, crunch, scrunch, crash, crumple, crease, cram*, which have a common semantic element ‘to modify shape forcibly’. Such words as *fire, flame, flare, flash, flicker* have the common element of ‘flickering light’. Words as *throw, thrust, thrash, thresh, thwack, thwart, thump* have the common element of ‘thumping sound’. Lexemes like *slack, slouch, slosh, sloppy, slug, sluggard, sloth, slattern, slow, sleepy, slovenly* seem to contain somewhat negative connotations with the impression of being too slow and/or untidy.

Studies such as these were conducted by such distinguished linguists as Leonard Bloomfield (1933) and John Rupert Firth (1964). The latter, however, was fiercely opposed to attributing any semantic value to the phonetic structure.

### 6. Ideophones — Cognitive Mechanisms

To explain the phenomenon of grouping some lexemes with a similar sound profile and semantic element, a mechanism called clustering has been proposed (Margaret Magnus). It works the following way: the sound structure of the basic (prototypic) lexeme for a given semantic category determines in some way the sound structure of lexemes less prototypic and peripheral. It seems to be a factor to explain why some words stick in a language for good, whilst some disappear over time. Those which stick seem to be perceived by native speakers as being “more explicit”, “graphic”, or “picturesque”, and somehow “better meaning” than others. Details of sound structures are language-dependent, but they do show certain similarities within a given language group or subgroup.
Another psycho-linguistic mechanism that is sometimes considered to be responsible for observed form-content regularities, is synaesthesia. This means that an impulse of one human sense is perceived as an impulse of another. This term is sometimes used in clinical psychiatry, where it applies to patients who perceive e.g. colours as sounds or sounds as tastes. Synaesthesia, though, as understood as the metaphorization sensual impulses between various human senses, is actually a commonplace process in natural languages – many impulses are conceptualized as other sensory impulses. In many languages (or perhaps all) the following expressions are very common and natural:

<table>
<thead>
<tr>
<th>EXAMPLE</th>
<th>COGNITIVE PATHWAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>loud shirt</td>
<td>sight/touch → hearing</td>
</tr>
<tr>
<td>bitter wind</td>
<td>feel → taste</td>
</tr>
<tr>
<td>sweet melody</td>
<td>hearing → taste</td>
</tr>
<tr>
<td>sharp reek</td>
<td>smell → touch</td>
</tr>
</tbody>
</table>

Another mechanism that may be held responsible for the given results is the motoric theory of speech. This theory is developed within the framework of the psycho-physiology of speech, and is based on relations of cognitive impulses with speech organs. Human speech organs are directly proximate to the central nervous system, located in the human head, thus there may be a certain neuro-psychological connection between sensory impulses and thought impulses and the impressions felt while generating speech sounds.

Explanations of this kind are as follows. To give a simple example, closed vowels generated with narrowed speech organs (e.g. [i]) are associated with something small, and open ones (e.g. [a]), that have open mouth, are associated with bigger objects.

The cognitive co-relation between speech sounds and size may be dead simple: one has to open the mouth wider to gulp bigger pieces of food than smaller ones. Tentatively speaking, were humans to produce speech with a different body organ than the same one to consume food, this phenomenon would not have emerged.

Researchers have long tried to establish concrete links between phonetics and semantics. Research studies were typically conducted with the participation of a test group of people with no acquaintance of a tested language, in which pairs of semantically contrasting adjectives (e.g. big-small, cold-warm) were chosen. The subjects were to pick the correct meaning (i.e. which of the two means ‘big’ etc.). There were numerous
research studies of this kind (examples see Allott 1995: 7–9), and the
languages tested included Japanese, Polish, Hungarian, Croatian, Hindi,
Hebrew etc. The researchers generally claim that the obtained results
exceeded the predicted random outcome, thus they have proved the
existence of non-arbitrary phono-semantics. The methodology of these
research studies was often criticized, though, with the critics claiming that
the researcher can subconsciously influence the tested subjects. Moreover,
not all the language pairs (i.e. the tested language and native language of
the subjects) showed a similar correlation of concordance.
Another kind of experiment was the often cited research by Edward Sapir,
conducted in 1929. The researcher posed the following question to the
tested subjects: “imagine two tables, a large one and a small one. In a
certain language there are two words for them, one is *mal*, the other being
*mil*. Which one means ‘large table’, and which one ‘small table’?” The
results differed in different test groups, but ca. 80% pointed to *mal* as
meaning ‘large’, and *mil* being definitely smaller one. Other experiments,
however (cf. Newman (1933), cited by Allott (1995); Newman was a
student of Sapir’s), did not seem to confirm Sapir’s results. Newman tested
the sound structure of ca. 500 English words related to size.
Psychologists and psycho-linguists also conducted another kind of
interesting experiment. They showed subjects abstract shapes and colours,
asking them to devise a fitting expression (or choose from given ones), or
attribute a word to a sound they heard. Such experiments were conducted
as early as in the 1920s (Usnadze, Kohler, Wissemann). A well-known
researcher in this field is Vilayanur Ramachandran, who coined the term
“buba-kiki effect”. He showed subjects two shapes: one with clear-cut
edges and sharp, pointed ends, the other one – smooth, with oblong curves.
He then asked to attribute abstract sound sequences of *buba* and *kiki* to
both. Subjects came from different language backgrounds, but the results
were quite consistent: ca. 95-98% of tested persons said *kiki* better suits the
sharp-edged figure. These results were observed not only with adult
subjects — children reacted in a similar way, but less conspicuously (only
ca. 83%).
The same experiment has also been carried out several times during the
present author’s lectures. The results also appear to corroborate
Ramachandran’s findings. The overwhelming majority (well above 90%)
chose *kiki* to mean a sharp-edged figure.\(^6\)

\(^6\) In my presentations the figures were additionally coloured. The sharp figure was red. Red is
commonly thought to be associated with blood, fresh meat, aggression, prominent sexuality and
sharpness. The smooth shape was green, the colour generally considered to be associated with
Another explanation of ideophone regularities is offered by **evolutional biology**. Research by Morton (cited by Ohala 1997: 2) focused on sounds uttered by animals of 56 species (birds and mammals) during a face-to-face fight scene or when in danger (so called “agonistic vocalizations”). The research showed that the individual who tried to attack or threaten an opponent emitted a low-pitched sound, whilst the submissive individual (who tried to surrender and escape the fight) was prone to produce a high-pitched vocalization. The proposed explanation is as follows. During a fight the sizes of opposing individuals are all-important, thus the competitors try to present themselves as larger than they actually are, to impress the opponent. Such behaviour is well attested and commonly known, and also includes the raising of feathers and hair, ears, wings, tails and spines. Threatening behaviour that manages to convince the opponent that the attacking party is larger, more mature and powerful, allows fighting to be avoided, which is objectively beneficial to both (they simply avoid physical harm). This is probably the reason behind the wide repertoire of such behaviours.

A larger individual is naturally prone to emit a voice of lower pitch, thus lowering one’s voice leads to being associated with a larger size. Pitch is in turn related to the frequencies of vocalization resonance. As John Ohala points out,

> [t]he characteristic frequencies of the vocal tract resonances are roughly inversely correlated with the length of the vocal tract which, in turn, is correlated with the linear dimensions of the vocalizer. So acoustically high resonances should convey an impression of a small vocalizer and low resonances, of a larger vocalizer (Ohala 1997: 3).

Such an interpretation may explain research findings, both experimental and data-based, which suggest a correlation between certain vowels in words describing size in various non-related languages.

**7. Is precise Phono-Semantic Description Possible?**

Numerous research studies seem to confirm that there is certain kind of relation between sound structure and meaning. As mentioned earlier, there is strong evidence to prove such a relation within one language, but there are also credible premises that a similar relation may exist across different languages.
Both assumptions, although far from being established theoretic postulates, are quite tempting to formulate a solid pattern of relations between sound structure and meaning. Such temptation is particularly strong when only one language is involved, as its native speakers instinctively feel the semantic nuances between similar lexemes, all the more so in the case of ideophones. One example of an attempt to describe such (proposed) regularities is a description published in a very useful and popular grammar of Japanese language (Makino, Tsutsui 1995: 50-56). The authors include the following regularities:

(A) voiced consonants represent something big, heavy, dull or dirty; voiceless consonants — something small, light, sharp or pretty, e.g.
- *koro-* to *korogaru* ‘[small object] rolls’
- *goro-* to *korogaru* ‘[heavy object] rolls’
- *pota-* to *ochiru* ‘[small amount of liquid] drips’
- *bota-* to *ochiru* ‘[large amount of liquid] drips’

(B) velar consonants ([k] and [g]) tend to represent hardness, sharpness, clear-cuttedness, separation, detachment or sudden change, e.g.
- *kukkiri* to *mieru* ‘be visible clearly’
- *garatto* to *kawaru* ‘completely change’
- *pokkuri* to *shinu* ‘die suddenly’

(C) a dental fricative consonant ([s]) tends to represent a quiet state or a quiet and quick motion, e.g.
- *suru-* to *suberu* ‘slide smoothly’
- *hissori* to *suru* ‘be quiet’
- *koso-* to *nigeru* ‘escape secretly’

(D) a liquid consonant [ɽ] tends to represent fluidity, smoothness or slipperiness, e.g.
- *sara-* to *nagareru* ‘flow smoothly’
- *sura-* to *kotaeru* ‘answer with great ease’
- *tara-* to *nagareru* ‘[sweat or blood] drip continuously’

(E) nasal consonants [m], [n], (and [n]) tend to represent tactuality, warmth and softness, e.g.
- *muku-muku* shite *iru* ‘[a dog or a cat] is plump’
- *nyuru-nyuru* shite *iru* ‘be slimy’
- *nechi-nechi* shite *iru* ‘be sticky’

(H) a back high vowel [u] tends to represent something that has to do with human physiology or psychology, e.g.

---

7 Quotation abridged.
usu-usu kanzuku ‘perceive dimly’  
uzu-uzu suru ‘itch for action’  
ukkari suru ‘be off guard’

(I) a back vowel [ɔ] tends to represent something basically negative with regard to human psychology, e.g.

ozu-ozu shite iru ‘be nervous and timid’  
ota-ota suru ‘don’t know what to do’  
ome-ome to damasareru ‘be deceived in a shameless manner’

(J) a front vowel [ɛ] tends to represent something vulgar, e.g.

hebereke ‘become dead drunk’  
hera-hera to warau ‘laugh meaninglessly when embarrassed’

Such a description seems to be quite arbitrary, though. Chosen examples may also look suspicious to many readers — one can easily find counter-examples to them. Another problem is that in the Japanese language, which has no more than five distinct vowel phonemes (if we exclude five long vowels), attributing precise, highly specialized semantic value to respective vowels can hardly be credible. Moreover, some vowels are described in a stunningly general manner — cf. (H), where “something that has to do with human physiology or psychology” could describe most other ideophonic lexemes as well.

This description is also obviously noncomprehensive — it does not provide any clues to determine the meaning (even in a most general way) of any given lexeme, it does not allow an ideophone of intended meaning to be generated. Similar descriptions have also been proposed for the Korean language as well (Sohn 2001: 96–101).

A much more comprehensive and better argued phono-semantic analysis is offered in Saito Hamano 1986 (77-227). The author, providing abundant examples, draws concrete conclusions as to the generalized semantic value of respective parts of e.g. bisyllabic mimetic adverbs (224-225):

The voicing feature of initial consonants of bisyllabic mimetic adverbs concerns the weight/mass of objects. In addition, initial consonants of bisyllabic mimetic roots describe the tactile nature of objects. This semantic dimension closely parallels the physical properties of the organs that are involved in the production of initial consonants themselves. These two semantic dimensions are summarized in the following.

/p/=breakable tense surfaces light/small/fine  
/b=/breakable tense surfaces heavy/large/coarse  
/t/=lack of surface tension, subduedness light/small/fine
Initial vowels generally control the semantic dimension of the shape of the first object or movement. The vowel /e/ is an exception.

| /i/  | line                     |
| /e/  | inappropriateness        |
| /a/  | flat plane               |
| /o/  | roundish object          |
| /u/  | protrusion               |

Such an analysis could lead to a conclusion that asserts the near-complete predictability of a given mimetic lexeme. Even if such findings are not entirely accurate, this proves that the formulation of a kind of regularized phono-semantic paradigm is possible to a degree.

8. Conclusions
A contemporary researcher of phono-semantic phenomena Margaret Magnus believes that both extreme stances (i.e. naturalists’ and conventionalists’) are over-generalizations. She argues that extreme naturalists’ claim about phonetic structure determining reference is wrong, though conventionalists’ claim about semantics of a lexeme that can be completely reduced to word reference appears to be equally wrong (MAGNUS 2001: 2) and that there is absolutely no relationship between sound structure and meaning.

Experimental research studies have provided much credible evidence that language signs are not completely arbitrary; that certain aspects of semantics can be encoded in its sound structure (yet not the referent, as Magnus points out). This is true particularly in certain subsets of vocabulary. The word group in which this non-arbitrariness is particularly visible, are ideophones.

Such a corollary gives rise to more questions than it offers answers. There are still crucial questions remaining to be answered, such as:

i. is the systematic description of phono-semantic relations possible, in the way Makino & Tsutsui and Saito Hamano have tried to do in the examples above;

ii. if so, is such a description possible cross-linguistically (i.e. common for a human as an animal that uses language), or only within one language system;
iii. are different layers of vocabulary phono-semantic in various degrees; one can theoretically preview that more specific terms might have closer phono-semantic ties, but this assumption remains yet to be proved or disproved;
iv. are only onomatopoeia and ideophones phono-semantic, thus making them “naturalistic stratum” in a natural language; to what extent are other layers (e.g. words relative to size, shape, sound and colour) phono-semantic; which makes other vocabulary layers purely conventionalist.

It seems that further studies in the field may prove to be quite fruitful to our understanding of language as a phenomenon. Also important is their validity to understand the cognitive processes that shape vocabulary in a specific language, in this case — Japanese. It seems though, judging from the present state of studies in the field, that the practical importance for the semantic description of the Japanese language is much less advanced than data that provide more general statements.

**Literature**
Hida Yoshifumi, Asada Hideko. 2002. *Gendai giongo gitaigo yōhō jiten* [a dictionary of contemporary giongo and gitaigo usage]. Tōkyō: Tōkyōdō. (飛田良文・浅田秀子著 『現代擬音語擬態語用法辞典』).


AUTHOR’S PROFILE

Bartosz T. Wojciechowski
Graduated from the Adam Mickiewicz University, Japanese Studies. After a stint in corporate media in Poland, completed Ph.D. degree in Japanese linguistics at University of Warsaw. Currently, for a few years already, working at Jagiellonian University in Cracow, Department of Japanology and Sinology. Published two books on Japanese language issues. Basic interests: Japanese linguistics, sociolinguistics, writing systems (particularly Kanji/Hanja), linguistic typology.