Akita, Kimi & Mutsumi Imai. 2022. The iconicity ring model for sound symbolism. In Sara Lenninger, Olga Fischer, Christina Ljungberg & Elżbieta Tabakowska (eds.), *Iconicity in cognition and across semiotic systems*, 27-45. (Iconicity in Language and Literature 18.) Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/ill.18.02aki (preprint)

The iconicity ring model for sound symbolism

Kimi Akita and Mutsumi Imai
Nagoya University / Keio University

The 'iconicity ring model' proposed in this chapter depicts a lexicon's evolutionary path from genuine iconicity (termed 'primary iconicity') to arbitrariness to another type of iconicity (termed 'emergent iconicity') that emerges from linguistic systematicity. The model captures the universality and language-specificity of sound symbolism and its role in lexical acquisition. Iconicity loss in ideophones illustrates the shift from primary iconicity to arbitrariness, whereas ideophonization exemplifies the shift from arbitrariness to emergent iconicity via systematicity. The two types of iconicity are mixed together in individual lexicons, providing a clue to the symbol grounding problem.

Keywords:
emergent iconicity, iconicity loss, ideophonization, primary iconicity, symbol grounding problem, systematicity

1. Introduction

Language has long been considered as a system of arbitrary symbols, whereby the forms of symbols are disconnected to their meanings or bodily experiences (de Saussure 1916). However, recent developments in language sciences have challenged this view, and a plethora of evidence from different disciplines and approaches suggests that spoken language contains sound symbolism, which is typically iconicity between the phonological form and meaning of words (Perniss et al. 2010; Imai and Kita 2014; Dingemanse et al. 2020). Many languages abound with a special class of sound-symbolic words called ideophones, expressives, or mimetics, in which sound-meaning correspondence is apparent (at least to the speakers of these languages). Recent research using large-scale statistical analyses revealed that even in languages which do not have a developed ideophone lexicon (e.g., Indo-European languages like English), the relationship between the form and meaning of words is not totally arbitrary (Christiansen and Monaghan 2016; Blasi et al. 2016; Johansson et al. 2020).

Sound symbolism has been discussed in connection with the ontogenesis and phylogenesis of language. Imai and Kita (2014) proposed the sound-symbolism bootstrapping hypothesis for language acquisition, which argues that sound symbolism scaffolds word learning in infants and young children in multiple ways: (1) it helps them gain insight in that each segment of speech sound has a meaning; (2) it helps children map words to corresponding referents at the site the word is introduced, be they objects, motion, acts, or properties of things; (3) it helps children generalize words to referents they did not directly observe when the word was

introduced. These have been supported by a body of empirical studies. For example, by analyzing infants' brain response using Event Related Potential (ERP) and large-scale synchronization, Asano et al. (2015) demonstrated that at 11 months of age, an infant's brain responds differentially to visual shape-word sound associations when they match and mismatch. Imai et al. (2015) further showed that 14-month-old infants learn associations between words and referent objects more easily when the two sound-symbolically matched than when they did not. Imai et al. (2008) demonstrated that 3-year-olds were able to generalize novel verbs to the same motion done by a different actor when the sound of the verbs matched the manner of motion, but they failed to generalize verbs that did not sound-symbolically match the manner of motion. The same effect was identified in English-speaking children even though their ambient language (i.e., English) does not have an established class of sound-symbolic words (Kantartzis et al. 2011).

These findings suggest that there are inherent sound-meaning correspondences that are universally accessible to children who do not have much exposure to word learning. However, as Imai and Kita (2014) acknowledged (see also Nielsen and Dingemanse 2021), the sound-symbolism bootstrapping effect is mostly for helping children to kickstart the vast enterprise of lexical acquisition, which takes many years. After all, modern language is an immensely complex system of abstract symbols. In the literature on lexical development, researchers have begun to explore how children learn and acquire such abstract systems in relatively later stages of lexical acquisition in which iconicity is not easily detectable (Ameel et al. 2008; Saji et al. 2011; 2020). However, in light of language evolution, the question is yet to be addressed how iconicity in language has evolved from stages in which linguistic signs were mostly iconic and expressive to stages in which iconicity is largely diluted, to the extent that outside a special class of sound-symbolic words such as ideophones, speakers do not (at least consciously) sense sound symbolism for a majority of words in the lexicon.

Here, we propose the 'iconicity ring model' that depicts the evolutionary process of sound symbolism. This model maintains that there are four stages in the evolutionary pathway that linguistic signs go through, as in Figure 1.

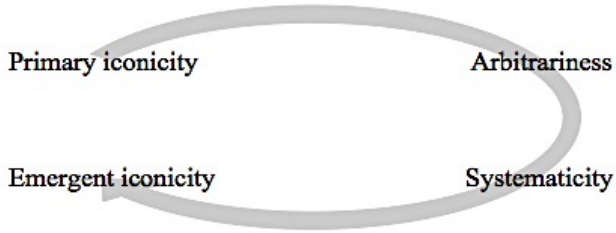

Figure 1. An iconicity ring in the evolutionary process of sound symbolism

The evolution starts from the stage of primary iconicity. In this phase, iconic forms are coined from mimicry of what human beings perceive in sensory modalities. These signs stand mostly on their own, like iconic gestures, and are hardly integrated into grammar. A lexicon consisting of sound-symbolic words has an advantage in terms of acquisition, processing, and communication as they are easy to learn and remember as well as to share common grounds with other speakers (Perniss et al. 2010; Ortega 2017; Perniss et al. 2017). However, as the need to differentiate concepts arises, the number of these iconic words increases; when the density of words in a given conceptual domain increases, the presence of similar-sounding words for similar concepts becomes a burden on the processing of these words in both retrieval and production (Gasser 2004; Monaghan et al. 2014; Christiansen and Monaghan 2016; Dingemanse et al. 2015). When this demerit exceeds the merit, words newly coming into the conceptual domain no longer carry iconicity to the concept, and non-sound-symbolic, arbitrary words gradually come to dominate the lexicon. Also, when the vocabulary of the domain grows sufficiently large, the lexicon is pushed toward regularity (Kirby et al. 2008) as a large conceptual space somehow needs to be structured. In this way, the number of non-iconic words increases, but systematicity (i.e., statistical regularity) arises at the same time.

It has been well-established in the literature on cognitive psychology that people's sense of similarity is very malleable and context dependent. For example, people judge that a dog and a doghouse are highly similar because the two are frequently associated by spatial proximity even though they share no common perceptual or conceptual attributes. People also judge that a pair of objects from the same classifier category (e.g., a cloud and a flower) are more similar than a pair drawn from different classifier categories (e.g., a bed and an apple) presumably because the two words in the former case frequently co-occur with the same classifier, even though the former and latter pairs were equally distant (i.e., conceptually unrelated) for people who do not speak a classifier language (Saalbach and Imai 2007; 2012). Thus, once regularity arises in language, words sharing the regular form (e.g., morphological marking) are

likely to be construed as 'similar', and subsequently, the sense of iconicity emerges.[1]

However, this type of iconicity should be distinguished from the first type. The former is characterized as 'primary iconicity', which reflects biologically grounded inherent iconicity and hence is expected to be universally accessible regardless of speakers' native language. In contrast, the latter, characterized as 'emergent iconicity', emerges from or is 'acquired' through linguistic knowledge and experience. Emergent iconicity is likely to be largely language-specific as it arises from the regularity of the lexical and grammatical systems of individual languages. The distinction between primary and emergent iconicity is closely related to Sonesson's (1997) 'primary' vs. 'secondary iconicity' (see also Ahlner and Zlatev 2010). Sonesson defines secondary iconicity as "an iconic relation between an expression and a content, which can only be perceived once the sign function, and a particular variety of it, is known to obtain" (1997: 741; see also Dingemanse [2011] for a similar notion termed 'coerced iconicity'). We use the term 'emergent iconicity' to refer more specifically to the type of iconicity that is motivated and constrained by linguistic structure. Thus, in the iconicity ring model, human language lexicons start as a set of primary icons, which gradually gains arbitrariness and finally develops back into iconic systems via linguistic regularity. Children appear to acquire primary icons first but once they learn emergent icons, they presumably become unaware of the difference between the two types of icons, finding both 'natural'.

We maintain that specifying the nature of emergent iconicity and clarifying factors and forces to create emergent iconicity provide important clues for understanding language evolution. Furthermore, these simultaneously offer insight into the symbol grounding problem in lexical acquisition (Harnad 1990; Imai 2017): how do children go beyond iconicity to acquire abstract linguistic systems, without falling into the symbol-to-symbol merry-go-round?[2]

In this chapter, we first outline the criteria for primary and emergent iconicity in sound symbolism in Sections 2 and 3, respectively. In Section 4, we explore how the shift from primary to emergent iconicity may take place, with special reference to Japanese ideophones. Section 5 concludes this chapter by discussing the possible role the iconicity ring may play in language evolution.

[1] Perry et al. (2015) report that English and Spanish speakers' subjective iconicity ratings for interjections, such as *uh oh*, *ouch*, and *yum*, were second highest following onomatopoeia. Given that interjections are primarily indexical, rather than iconic, of the speaker's internal state or emotional attitude, the perceived resemblance between their phonological forms and meanings should be considered of secondary status. This (mis-)perception appears to come from the conceptual closeness between our psychological state and voice (for a related discussion, see Little and Sulik 2018).

[2] In contrast to the Peircean terminology, 'symbol' is here synonymous with 'linguistic sign'.

2. Primary iconicity in sound symbolism

Candidates for primary icons in spoken language include ideophones for human voice, such as *ha-ha* for laughter. Voice ideophones are clearly iconic because they mimic vocal sounds by means of vocal sounds. However, even these ideophones may not be identical to their referents as they consist of conventional linguistic sounds in individual languages, and each language is likely to have multiple candidate sounds that would match the properties of the referent (for the dialectal/diachronic variation of Japonic ideophones, see Yamaguchi [2019]; McLean [2021]). In other words, even though it is supported by biologically grounded iconicity, the meaning-sound mapping is a one-to-multiple mapping rather than a one-to-one mapping. A good example of this is the *maluma-takete* case. There are multiple candidate sounds that are judged to match round shapes (e.g., *maluma*, *moma*, *bouba*, *boba*, *yuya*, *rura*) and spiky shapes (e.g., *takete*, *goga*, *gogu*, *kiki*, *kipi*, *zozi*) (see Köhler 1929; Ramachandran and Hubbard 2001; Imai et al., under review). Hence, iconicity involved in spoken language lexicons is not comparable to visual icons, such as realistic pictures and handling signs in sign language (e.g., a drinking gesture for DRINK in Japanese Sign Language). However, it is still possible to argue that some spoken words such as laughter ideophones involve more primary iconicity than others. In this section, we describe the two major characteristics of primary icons in spoken language—universality and early acquisition—and summarize related linguistic and psychological findings.

2.1 Universality

Primary icons are found in similar forms across languages. Ideophones for human voice and animal cries are the clearest cases. Laughter ideophones in numerous languages contain the glottal fricative /h/ and an open vowel: *ha-ha* (English), *ahaha* (Japanese), *hā-hā* (Chinese), *haha* (French, Turkish, Hungarian, etc.). These sounds are acoustically similar to laughter even though they are linguistically segmented and assimilated to the phonological structure of each language. Likewise, ideophones for a cat's cry often contain a nasal consonant: *meow* (English), *nyaa* (Japanese), *miāo* (Chinese), *miaou* (French), *yaong* (Korean). This crosslinguistic similarity is attributed to the likely fact that nasals are the closest equivalents of a meow among the cardinal consonants in human languages. Thus, primary icons consist of form and meaning that

are objectively similar to each other, and the linguistic form may serve as an unambiguous clue to the meaning (for an acoustic analysis of onomatopoeia, see Masuda 2007; Assaneo et al. 2011).

The crosslinguistic quantification of the sound-meaning correspondences in basic words suggests that primary iconicity is not limited to ideophones (Blasi et al. 2016, etc.). For example, Johansson et al.'s (2020) investigation of the words for 344 near-universal concepts in 245 language families has found that /r/-like consonants are found in words for TURN in significantly many languages (e.g., /mawaɾɯ/ in Japanese, /ɾuutininiu/ in Arabela, /keerama/ in Estonian, /ɟeɾe/ in Guaraní, /murundi/ in Mandinka, /ʈiɾumbu/ in Tamil). The crosslinguistic prevalence of this sound-meaning mapping is attributed to the flapping/tapping movement or curved shape of the tongue that these consonants involve. Similarly, in many languages and nonword-based experiments, close/front and open/back vowels are associated with smallness and largeness, respectively, as illustrated by the Hungarian adjectives *kicsi* /ˈkitʃi/ 'small' and *nagy* /nɒɟ/ 'large' and Sapir's (1929) experimental stimuli *mil* and *mal* (see also Fónagy 2001). The crosslinguistic prevalence of vowel-size symbolism is presumably attributed to its articulatory basis: the oral cavity is smaller in close/front vowels than in their corresponding open/back vowels. According to the Frequency Code Hypothesis (Ohala 1984), vocalic symbolism also has an acoustic basis: higher F_2 is associated with smaller size. Smaller vocalizers make higher-frequency sounds in the animal world (compare mice and elephants, for example), and this zoological correspondence can give an additional innate or experiential basis to the sound-symbolic mapping. Although these types of sound-meaning correspondences are not as widespread as ideophonic sound symbolism, they often have a clear embodied basis and, therefore, are considered examples of primary iconicity.

The *maluma/takete* effect may also involve primary iconicity, as it is sensed broadly across speakers of diverse languages, although it concerns one-to-multiple mappings at the word-form level. *Maluma* and other round-sounding words contain sonorants (e.g., /m, n, l, ɹ, j, w/), bilabials (e.g., /m, b/), or rounded back vowels (e.g., /u/, /o/), whereas *takete* and other spiky-sounding words contain non-bilabial obstruents (/t, k, s, d, g, z/) or unrounded front vowels (e.g., /i, e/). The association between rounded vowels and round shapes has an obvious articulatory basis. Moreover, it appears that smooth, gradual amplitude changes involved in sonorant consonants are associated with rounded shapes, while the abrupt or irregular amplitude changes involved in obstruents evoke spiky images (D'Onofrio 2014).

2.2 Early acquisition

Children acquire primary icons earlier than emergent icons. Relevant reports come from Japanese. Several corpus and experimental studies (Akita 2009; among others) have shown that children acquire onomatopoeic ideophones (ideophones for sound; e.g., *batabata* 'fluttering', *patipati* 'clapping') earlier than non-onomatopoeic ideophones (those for visual, motor, tactile, or inner perceptions; e.g., *sarasara* 'smooth and dry', *mukamuka* 'disgusted'). The same contrast has been observed for non-Japanese-speaking adults, who were more successful in guessing the meanings of onomatopoeic than non-onomatopoeic ideophones in Japanese (Iwasaki et al. 2007; Dingemanse et al. 2016). Assuming that onomatopoeic ideophones involve primary iconicity, these findings suggest that primary iconicity gives particularly useful cues to language learners.

Some sound-meaning mappings are detected even by infants who have not started active word learning (Peña et al. 2011; Asano et al. 2015; Pzturk 2013; Pejovic and Molnar 2017). Asano and colleagues (2015) demonstrated that 11-month-old infants showed different brain responses when word sounds and shapes matched (round-*moma*; spiky-*kipi*) and mismatched (round-*kipi*; spiky-*moma*). These results support the idea that the *maluma-takete* sound symbolism involves primary iconicity and that the mapping is biologically grounded. However, it remains ambiguous whether even sound symbolism based on primary iconicity can be detected without linguistic experience. Fort et al.'s (2018) meta-analysis of research on infants' sensitivity to sound symbolism revealed that prelinguistic children are sensitive to the roundedness of *bouba* but not to the spikiness of *kiki*.[3] As already mentioned, the roundedness of *bouba* is clearly motivated by its rounded vowel /u/ and also presumably by its bilabial /b/. On the other hand, given that children acquire /k/ later than /b/ (Brett et al. 1987), the detection of *kiki*'s spikiness may require more linguistic experience. Thus, it is possible that the results of the meta-analysis suggest that *bouba* involves more primary iconicity, and *kiki* involves more emergent iconicity. However, Fort et al.'s results could have arisen from methodological issues since most studies subjected to their meta-analysis used looking time measures, which require infants to show fairly long-lasting duration time to demonstrate the sound-symbolic effect. Asano et al. (2015) indeed showed observed sound-symbolic effects in both *moma*- and *kipi*-sound trials, and the effect size for *kipi* was reported to be larger than that for *moma* by Fort et

[3] In contrast, Imai et al. (under review) suggest that the roundedness of *bouba* as such may not be as clear as the spikiness of *kiki* and may be highlighted when contrasted with *kiki*.

al. (2018) in their meta-analysis. Considering these results, further research is required to draw a conclusion concerning whether sensitivity to primary iconicity based sound symbolism should require linguistic experiences.

Despite the crosslinguistic availability of onomatopoeia and size and shape sound symbolism, when viewed globally, primary iconicity in the lexicon seems to be limited. In the next section, we discuss emergent iconicity in detail.

3. Emergent iconicity in sound symbolism

3.1 Language-specificity

Emergent iconicity accounts for a large part of sound symbolism in adults' sound-symbolic lexicon. By our definition given in Section 1, emergent iconicity is largely language-specific, and its phonetic basis is often unclear. We can find examples in non-onomatopoeic ideophones. For example, ideophones for slipping movement sound diverse across languages: *turun* (Japanese), *pʊrʊʊtʃ* (Didinga; de Jong 2001: 122), *shelele* (Zulu; Msimang and Poulos 2001: 246), *giɽir-ken* (Mundari; Badenoch and Osada 2019: 268). In extreme cases, language-specific sound-meaning mappings run counter to major sound-symbolic mappings. A well-known example is the vocalic symbolism in Korean ideophones in which /a/ is mapped onto smallness and /i/ onto largeness (e.g., *chi.ləŋ* vs. *cha.laŋ* 'dropping of a longer vs. shorter object'; Kwon 2018: 4; see also Diffloth 1994). These exceptional vowel-size correspondences are attributed to the traditional *yin/yang* 'dark/light' distinction in Korean culture, which is symbolized in the flag of South Korea. The Korean vowel system is divided into 'light' and 'dark vowels', which are linked with small and large images, respectively; /i/ is dark, and /a/ is light, hence the counter-universal mappings in size symbolism (Cho 1994).

Language-specific sound symbolism has also been attested in production experiments. Saji et al. (2019) showed short video clips of human motion to Japanese and English speakers. Participants were asked to create novel words that match different manners of motion and to rate each manner of motion on five Likert scales (e.g., heavy-light, fast-slow, energetic-non-energetic). Speakers of the two languages used quite different sets of form-meaning associations in their new word creation. For example, only Japanese speakers used nasality to express slow motion, perhaps by analogy with existing ideophones such as *noronoro* 'moving sluggishly', *nossinossi* 'walking heavily', *nosonoso* 'moving slowly', and *nonbiri* 'in a leisurely

way'.

Emergent iconicity may not have a clear phonetic basis. However, the unclarity of bodily motivations does not prevent native speakers from perceiving iconicity because they find their native language most 'natural'. A related anecdote is cited here from Hinton et al. (1994; see also Haiman 2018: 1).

> Children feel this [= the 'naturalness' of names] especially strongly, as illustrated once by Stephanie, the stepdaughter of one of the authors: she said, "English is the one true language, isn't it?" When asked what she meant, she replied, "Well, when [our Mexican friend] Lupe says 'agua', what she *means* is 'water'. But when *I* say 'water', I don't mean 'agua', I really mean 'water'!"
>
> (Hinton et al. 1994: 5)

The 'naturalness' of our native language receives support from iconicity rating studies. Occhino et al. (2017) asked American Sign Language (ASL) signers and German Sign Language (DGS) signers to rate the iconicity of both ASL and DGS signs and found that both groups of signers rated their own language more iconic. Moreover, Thompson et al. (2020) show that Japanese speakers rated native lexemes (e.g., *toki* 'time', *karada* 'body') more iconic than loanwords from Chinese (e.g., *ziko* 'accident', *tosi* 'city') or European languages (e.g., *deeta* 'data', *hoteru* 'hotel'). These findings confirm that the sense of naturalness emerges from our long-time frequent exposure to native words. This acquired naturalness can be equated with emergent iconicity. Given this equation, phonesthemes, such as *gl-* in English vision-related words (e.g., *glisten*, *glitter*, *glance*, *glow*), which are sometimes cited as examples of non-iconic systematicity (Thompson and Do 2019), may also be included in the category of icons (Hinton et al. 1994).

3.2 Later acquisition

Children tend to acquire emergent icons later than primary icons as emergent iconicity is often based on unclear phonetic motivations within a language-specific phonological system. Imai et al. (2008) report that Japanese 2-year-olds successfully mapped most novel ideophonic verbs created from existing Japanese ideophones, such as *nosunosu-su(ru)* 'walk heavily' and *to-kutoku-su(ru)* 'walk with short steps', to the 'correct' (in the sense of adult Japanese speakers) referent videos but failed to map *hyaihyai-su(ru)* 'walk semi-swiftly with light, playful steps'. The nasal /n/ has a low frequency, and Ohala's Frequency Code would successfully predict its

association with heavy movement. The voiceless stops /t, k/ involve momentary contact between the tongue and the upper surface of the oral tract and have a relatively high frequency. These phonetic properties are readily associated with swift movement in short steps. In contrast, the meaning of /hy/ [ç] is not fully motivated by its articulatory and acoustic properties. The high frequency of this consonant may evoke a light object's swift movement through the Frequency Code, but the 'playful' part of the meaning of the novel ideophone is too abstract to be captured by the universal code or other embodied principles. Therefore, this abstract sound symbolism should be considered an example of emergent iconicity.

The meanings of emergent icons are also hard for nonnative speakers to guess. Iwasaki et al. (2007) report that English speakers without prior knowledge of Japanese failed to guess the evaluative aspect of the meanings of Japanese laughter ideophones, such as the vulgarity of *hehehe*. The negative feeling for /e/ in *hehehe* is ascribed to the marked infrequency of this vowel in the Japanese lexicon. This system-motivated sound symbolism is clearly an example of emergent iconicity. Similarly, in Imai et al. (2008), English-speaking adults failed to map *hyaihyai* onto playful semi-swift motion. The Japanese-specific association between /hy/ and this manner of motion makes this ideophone inaccessible to nonnative speakers without prior knowledge of Japanese.

4. Shift from primary to emergent iconicity

In Sections 2 and 3, we outlined the basic characteristics of primary and emergent iconicity in the context of sound symbolism. In this section, we illustrate the gradual shift from primary to emergent iconicity via arbitrariness and systematicity, with special attention to Japanese ideophones. Section 4.1 examines the morphosyntactic integration of ideophones that may cause iconicity loss. Section 4.2 focuses on ideophonization as an example of the emergence of iconicity.

4.1 From primary iconicity to arbitrariness

Not all ideophones are equally iconic. The iconicity of ideophones decreases as a function of morphosyntactic integration. Morphosyntactic integration refers to the integration of ideophones into sentence structure, especially the predicate. Predicative ideophones, which are

often headed by light verbs (e.g., *irrist egin* 'slide (IDPH make)' in Basque (Ibarretxe-Antuñano [2017: 203])), are more morphosyntactically integrated than non-predicative ideophones since they constitute the essential part of a sentence. While non-predicative ideophones, especially holophrastic ones, tend to be prominent in terms of prosody and morphophonology and be accompanied by iconic gesture, predicative ideophones tend to be plain and behave like non-ideophonic words. Ideophones in some languages, such as Dagaare and Semai, abound with non-predicative ideophones that stand out clearly, whereas ideophones in other languages, such as Somali, only have predicative ideophones that are low in expressivity (Dingemanse 2017).

Japanese has both predicative and non-predicative ideophones, and they illustrate the integration-iconicity trade-off within a single linguistic system (Dingemanse and Akita 2017). A representative pair of examples that show this trade-off are cited in (1). The upward arrows represent intonational prominence.[4]

(1) a. *Hune ga ↑guruguruut↑ to mawat-te, don to*
ship NOM IDPH QUOT turn-GER IDPH QUOT
butukat-te, hane-kaes-are-te ku-ru. (non-predicative)
hit-GER bounce-make.return-PASS-GER come-NPST
Lit. 'The ship turns around and around, hits [something] with a bang, and comes bouncing back'.

b. *De, zyup-pun gurai, ie no mawari o*
and 10-min about house GEN neighborhood ACC
guruguru-si-te tara, de-te ki-ta-n-des-u kedo,
IDPH-do-GER when exit-GER come-PST-NMLZ-POL-NPST but
sorede, obaatyan mukae ni ki-ta yo tte. (predicative)
then grandma pick.up DAT come-PST FP QUOT
Lit. 'And when [I] was wandering around the house for about 10 minutes, [she] came out, and then, [I shouted] "Grandma, [I] am here to pick you up!"'
(*NHK Great East Japan Earthquake Archive*)

The quotative-marked ideophone *gurut to* 'turning around' in (1a) is made particularly iconic by partial reduplication and vowel lengthening. This ideophone is intonationally foregrounded and is synchronized with a hand gesture for a circling motion. In contrast, the reduplicated ideophonic verb *guruguru-su(ru)* 'wander around' in (1b) does not have any of these features. As these examples show, the expressive power typical of ideophones tends to be weakened when they are deeply integrated into sentence structure.

Crucially, it is likely that the morphosyntactic integration of Japanese ideophones also took place diachronically (Akita [2021]; for a related discussion, see Haiman [2018]). Adverbial

[4] Abbreviations: ACC = accusative, DAT = dative, FP = final particle, GEN = genitive, GER = gerundive, IDPH = ideophone, NMLZ = nominalizer, NOM = nominative, NPST = nonpast, PASS = passive, POL = polite, PST = past, QUOT = quotative.

ideophones have arguably been the most common since Old Japanese (the 8th century, e.g., *soyo ni* 'blowing softly', *uraura to* 'with a mild sunlight'). However, in Early Middle Japanese (until the late 12th century), many ideophone roots were combined with the suffix *-mek* to form verbs (e.g., *huta-mek(u)* 'scurry', *gara-mek(u)* 'rattle', *soso-mek(u)* 'whisper'). This suffix became less productive in Late Middle Japanese (until the 16th century) and now has very limited productivity in the central dialects of Japanese. In Early Modern Japanese (until the mid-19th century), while the so-called quotative marker *to* started dropping from adverbial (and verbal) ideophones (e.g., *gatagata* 'with a shiver', *yotayota* 'staggering'), the predicative construction headed by *su(ru)* 'do' (e.g., *nikoniko-su(ru)* 'smile', *burabura-su(ru)* 'sway') became increasingly common. In summary, morphosyntactically integrated ideophones increased due to *mek*-verbalization in Early Middle Japanese, decreased in Late Middle Japanese, and increased again due to the quotative drop and the rise of 'do'-verbs in Early Modern Japanese. Thus, the evolution of iconic words in light of morphosyntactic integration is not a linear change from separated to integrated. Nevertheless, a bird's-eye view may allow us to conclude that the overall degree of morphosyntactic integration has increased in the history of Japanese ideophones (cf. Yamaguchi 2019).

The current discussion supports the primary-iconicity-to-arbitrariness part of the iconicity ring. Ideophones lose their iconicity when they are integrated into the core of sentence structure and language-particular grammatical constructions. This phenomenon appears to mirror language evolution in which primitive language evolved out of holophrastic vocal mimicry and highly iconic ideophone-gesture combinations (Haiman 2018).

The morphosyntactic typology of ideophones may provide further insights with respect to Talmy's (2000) 'framing typology'. This cognitive-semantic typology focuses on how languages tend to encode complex events, especially spatial motion events. Languages are divided into two types according to where in the sentence Path of motion is expressed. Languages in which Path is typically encoded in the main verb, such as Romance and Semitic languages, are called 'verb-framed languages'; languages in which Path is typically expressed outside the main verb, such as Germanic, Slavic, and Uralic languages, are called 'satellite-framed languages'. As illustrated in (2), Japanese is verb-framed, whereas English is satellite-framed.

(2) a. Verb-framed:
Otoko ga toori o [Manner *tobotobo arui-te*] [Path *yokogit-ta*]
man NOM street ACC IDPH walk-GER cross-PST
Lit. 'A man crossed the street walking ploddingly'.

b. Satellite-framed:
A man [Manner *plodded*] [Path *across*] *the street.*

Given that ideophones primarily express Manner, the verb- and the satellite-framed patterns correspond to non-predicative and predicative constructions for ideophones, respectively. Ideophones are more prominent and iconic in verb-framed constructions than in satellite-framed ones because in the former ideophones are typically realized as adverbials, which retain their iconicity. In fact, some researchers argue that ideophones are localized to verb-framed languages (Wienold 1995; Akita 2021).

Moreover, Chinese is said to have changed from verb-framing to satellite-framing (or so-called 'equipollently framing'). For example, the Old Chinese clauses in (3a) are headed by the path verbs *rū* 'enter' and *chū* 'exit', whereas their Modern Chinese translations in (3b) use serial-verb constructions containing the manner verbs *zǒu* 'run' and *liū* 'slip'.

(3) a. Old Chinese (end of 5th century; verb-framed):
Jiāng ***rū*** *yú shì, yǔ Cuīzǐ zì cèhù* ***chū***.
Jiāng enter to room with Cuīzǐ from side.door exit
'Jiāng entered the room and exited the side door with Cuīzǐ'.

b. Modern translation (satellite/equipollently-framed):
Jiāngshì [...] ***zǒu*** ***jìn*** *nèishì hé Cuī Zhù cóng*
Jiāngshì [...] run enter inner.room with Cuī Zhù from
biānmên ***liū*** ***chū*** ***qū***.
side.door slip exit go
'Jiāngshì [...] ran into the inner room and slipped out of the side door with Cuīzhù'.

(adapted from Shi and Wu 2014: 1246-1247)

The shift from verb- to satellite-framing is consistent with the grammaticalization of Path expressions, in which verbs (e.g., *passed*) become particles or adpositions (e.g., *past*). Thus, it might be that the diachronic integration of ideophones at the expense of iconicity is correlated with the overall constructional development of the language.

4.2 From arbitrariness to systematicity to emergent iconicity

The shift from arbitrariness to emergent iconicity via systematicity is clearly observed in ideophonization. It is fairly common in ideophone-rich languages that verbs, nouns, and adjectives become ideophones through certain morphological or prosodic operations, such as reduplication and lengthening (Awoyale 1983; Dingemanse 2017; among others). Japanese examples

are given in (4).[5]

(4) a. Deverbal ideophones:
hiyahiya 'thrilled' (< *hiyas(u)* 'make cool'), *hukihuki* 'wiping repeatedly' (< *huk(u)* 'wipe'), *hurihuri* 'wagging' (< *hur(u)* 'swing'), *ikiiki* 'lively' (< *iki(ru)* 'live'), *konekone* 'kneading' (< *kone(ru)* 'knead'), *mazemaze* 'mixing' (< *maze(ru)* 'mix'), *momimomi* 'massaging' (< *mom(u)* 'massage'), *musimusi* 'sultry' (< *mus(u)* 'steam'), *nobinobi* 'free and easy' (< *nobi(ru)* 'stretch'), *surisuri* 'rubbing' (< *sur(u)* 'rub'), *ukiuki* 'buoyant' (< *uk(u)* 'float'), *wakuwaku* 'excited' (< *wak(u)* 'spring up')

b. Denominal ideophones:
iraira 'irritated' (< *ira* 'thorn'), *motimoti* 'chewy' (< *moti* 'rice cake'), *siwasiwa* 'wrinkled' (< *siwa* 'wrinkle'), *raburabu* 'lovey-dovey' (< *rabu* 'love'), *togetoge* 'prickly' (< *toge* 'prickle')

All these words originated in non-ideophonic words but have acquired ideophonic sound-symbolic effects through reduplication (Akita 2009). Specifically, the intervocalic voiceless velar stop /k/ in *ikiiki* 'lively', *ukiuki* 'buoyant', and *wakuwaku* 'excited' appears to evoke a light upward movement that is associated with happy feelings. Importantly, the non-ideophonic originals (i.e., *iki(ru)* 'live', *uk(u)* 'float', *wak(u)* 'gush out') do not clearly have the same effects (for a related experimental report, see Akita 2011). The low sound-symbolic relevance of /k/ in these verbs receives support from the fact that it can disappear in conjugation. For example, the past and gerundive forms of the verbs *uk(u)* 'float' and *wak(u)* 'gush out' involve euphonic consonant-vowel alternations: *u**i**-ta* 'floated', *u**i**-te* 'floating', *wa**i**-ta* 'gushed out', *wa**i**-te* 'springing up'. These examples indicate that iconicity can emerge from morphological ideophonization.

Ideophonization is also common in Japanese manga. Manga is full of ideophonic sound effects that are often located on the background. カニカニカニカニ *kanikanikanikani* utilizes this technique for ideophonizing the non-ideophonic noun *kani* 'crab' to vividly decorate the small river crab's special manner of walking.

[5] There is no consensus among Japanese linguists on whether these words should really be called ideophones. We assume that Japanese ideophones constitute a prototype category with a fuzzy boundary in which one member is a better exemplar than another. Even marginal members share some formal or functional properties with central members.

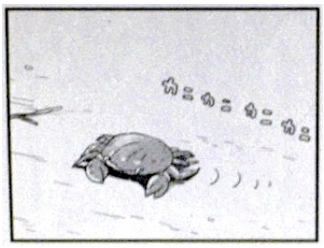

Figure 2. Ideophonization in manga (Choborau 2011)

All these examples are crucially based on the ideophone systems of individual languages or media, and they do not necessarily have a universal bodily basis. This system-specificity contributes to the emergent iconicity of ideophonic sound symbolism and to its creativity.

5. Implications for the symbol grounding problem

In this chapter, we proposed a distinction between primary and emergent iconicity and a four-step model of the development of sound symbolism from primary to emergent iconicity via arbitrariness and systematicity. Primary iconicity, illustrated by size and shape sound symbolism and onomatopoeic ideophones, has a clear bodily basis and is available across languages. Emergent iconicity, illustrated by a large part of ideophonic sound symbolism, arises from linguistic knowledge and experience and, therefore, is often language-specific. The morphosyntactic and functional diversity of ideophones supports the first shift in the iconicity ring (i.e., from primary iconicity to arbitrariness), and ideophonization illustrates the second and third shifts (i.e., from arbitrariness to systematicity to emergent iconicity).

Critically, we hypothesized that native speakers of ideophone-rich languages, such as Japanese, do not distinguish primary and second iconicity, considering both 'iconic'. The feeling that primary iconicity and emergent iconicity are equally iconic and natural offers insight into how a language may (chronologically) develop a lexicon. Although human language may have evolved from vocal and gestural imitations (i.e., primary icons), it is arbitrariness that enhanced the flexibility of their form-meaning mappings (Monaghan et al. 2014). System-internal regularity creates emergent iconicity out of these arbitrary mappings and mingles it with primary iconicity.

The scenario explored in this chapter also provides implications for the symbol grounding problem (Harnad 1990) noted in Section 1. We may reinterpret and recapitulate this cognitive-scientific problem as follows: how do children ground the first set of linguistic signs to things in the world and build up a language-specific lexicon, eventually embodying it to the extent

that they feel that the lexical system they have constructed is the only way of naturally categorizing the world? The development of the sound-symbolic lexicons of ideophone-rich languages may be considered as a miniature model for this process.

Acknowledgements

MI presented earlier versions of this chapter at the 'Types of Iconicity in Language Use, Development and Processing' workshop (Max Planck Institute for Psycholinguistics, 2017) and the Twelfth International Symposium on Iconicity in Language and Literature (Lund University, 2019). We especially thank Jordan Zlatev for his valuable comments. Any remaining inadequacies are ours. This study was supported by MEXT/JSPS KAKENHI Grants (no. 20H00014, 16H01928, 18H05084) to MI.

References

Ahlner, F. and Zlatev, J. 2010. Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. *Sign Systems Studies* 38: 298–348.

Akita, K. 2009. A Grammar of Sound-Symbolic Words in Japanese: Theoretical Approaches to Iconic and Lexical Properties of Mimetics. Ph.D. dissertation, Kobe University.

Akita, K. 2011. Toward a phonosemantic definition of iconic words. In *Semblance and Signification*, P. Michelucci, O. Fischer and C. Ljungberg (eds), 3–18. Amsterdam and Philadelphia: John Benjamins.

Akita, K. 2021. Nihongo no onomatope to gengo ruikeiron [Japanese ideophones and linguistic typology]. In *Nihongo kenkyū to gengo riron kara mita gengo ruikeiron* [Linguistic typology from the perspectives of Japanese linguistics and linguistic theories], H. Kubozono, H. Noda, P. Pardeshi and Y. Matsumoto (eds), 49–73. Tokyo: Kaitakusha.

Ameel, E., Malt, B. and Storms, G. 2008. Object naming and later lexical development: From baby bottle to bear bottle. *Journal of Memory and Language* 58: 262–285.

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H. and Thierry, G. 2015. Sound symbolism scaffolds language development in preverbal infants. *Cortex* 63: 196–205.

Assaneo, M. F., Nichols, J. I. and Trevisan, M. A. 2011. The anatomy of onomatopoeia. *PLoS ONE* 6: e28317.

Awoyale, Y. 1983–1984. On the semantic fields of Yoruba ideophones. *Journal of the Linguistic Association of Nigeria* 2: 11–22.

Badenoch, N. and Osada, T. 2019. *A Dictionary of Mundari Expressives*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F. and Christiansen, M. H. 2016. Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences* 113: 10818−10823.

Brett, L., Chiat, S. and Pilcher, C. 1987. Stages and units in output processing: Some evidence

from voicing and fronting processes in children. *Language and Cognitive Processes* 2: 165–177.
Cho, M.-H. 1994. Vowel Harmony in Korean: A Grounded Phonology Approach. Ph.D. dissertation, Indiana University.
Choborau, N. 2011. *Fushigi na somerachan* [Mysterious Somera], Vol. 2. Tokyo: Ichijinsha.
Christiansen, M. H. and Monaghan, P. 2016. Division of labor in vocabulary structure: Insights from corpus analyses. *Topics in Cognitive Science* 8: 610–624.
de Jong, N. 2001. The ideophone in Didinga. In *Ideophones*, F. K. E. Voeltz and C. Kilian-Hatz (eds), 121–138. Amsterdam and Philadelphia: John Benjamins.
de Saussure, F. 1916. *Cours de linguistique générale*. Paris: Payot.
Diffloth, G. 1994. *i*: big, *a*: small. In *Sound Symbolism*, L. Hinton, J. Nichols and J. J. Ohala (eds), 107–114. Cambridge: Cambridge University Press.
Dingemanse, M. 2011. The Meaning and Use of Ideophones in Siwu. Ph.D. dissertation, Radboud University Nijmegen.
Dingemanse, M. 2017. Expressiveness and system integration: On the typology of ideophones, with special reference to Siwu. *STUF – Language Typology and Universals* 70: 363–384.
Dingemanse, M. and Akita, K. 2017. An inverse relation between expressiveness and grammatical integration: On the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics* 53: 501–532.
Dingemanse, M., Blasi, D. E., Monaghan, P., Lupyan, G. and Christiansen, M. H. 2015. Arbitrariness, iconicity and systematicity in language: Empirical and theoretical advances. *Trends in Cognitive Science* 19: 603–615.
Dingemanse, M., Perlman, M. and Perniss, P. 2020. Construals of iconicity: Experimental approaches to form-meaning resemblances in language. *Language and Cognition* 12: 1–14.
Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S. and Mitterer, H. 2016. What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language* 92: e117–133.
D'Onofrio, A. 2014. Phonetic detail and dimensionality in sound-shape correspondences: Refining the *bouba-kiki* paradigm. *Language and Speech* 57: 367–393.
Fónagy, I. 2001. *Languages within Language: An Evolutive Approach*. Amsterdam and Philadelphia: John Benjamins.
Fort, M., Lammertink, I., Peperkamp, S., Guevara-Rukoz, A., Fikkert, P. and Tsuji, S. 2018. Symbouki: A meta-analysis on the emergence of sound symbolism in early language acquisition. *Developmental Science* 21: e12659.
Gasser, M. 2004. The origin of arbitrariness in language. In *Proceedings of the 26th Annual Cognitive Science Society Conference*, K. Forbus, D. Gentner and T. Regier (eds), 434–439. Mahwah, NJ: Lawrence Erlbaum Associates.
Haiman, J. 2018. *Ideophones and the Evolution of Language*. Cambridge: Cambridge University Press.
Harnad, S. 1990. The symbol grounding problem. *Physica D: Nonlinear Phenomena* 42: 335–346.
Hinton, L., Nichols, J. and Ohala, J. J. 1994. Introduction: Sound-symbolic processes. In *Sound Symbolism*, L. Hinton, J. Nichols and J. J. Ohala (eds), 1–12. Cambridge: Cambridge University Press.
Ibarretxe-Antuñano, I. 2017. Basque ideophones from a typological perspective. *Canadian Journal of Linguistics* 62: 196–220.
Imai, M. 2017. The "symbol grounding problem" reinterpreted from the perspective of language acquisition. In *Meaning, Mind and Communication Explorations in Cognitive*

Semiotics, J. Zlatev, G. Sonesson and P. Konderak (eds), 145–159. Frankfurt am Main: Peter Lang.
Imai, M. and Kita, S. 2014. The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B* 369: 20130298.
Imai, M., Kita, S., Nagumo, M. and Okada, H. 2008. Sound symbolism facilitates early verb learning. *Cognition* 109: 54–65.
Imai, M., Miyazaki, M., Yeung, H., Hidaka, S., Kantartzis, K., Okada, H. and Kita, S. 2015. Sound symbolism facilitates word learning in 14-month-olds. *PLoS ONE* 10: e0116494.
Iwasaki, N., Vinson, D. P. and Vigliocco, G. 2007. What do English speakers know about *gera-gera* and *yota-yota*?: A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese-Language Education around the Globe* 17: 53–78.
Johansson, N. E., Anikin, A., Carling, G. and Holmer, A. 2020. The typology of sound symbolism: Defining macro-concepts via their semantic and phonetic features. *Linguistic Typology* 24: 253–310.
Kantartzis, K., Imai, M. and Kita, S. 2011. Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science* 35: 575–586.
Kirby, S., Cornish, H. and Smith, K. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences* 105: 10681–10685.
Köhler, W. 1929 [1947]. *Gestalt Psychology*. New York: Horace Liveright.
Kwon, N. 2018. Iconicity correlated with vowel harmony in Korean ideophones. *Laboratory Phonology* 9: 1. 1–18.
Masuda, K. 2007. The physical basis for phonological iconicity. In *Insistent Images*, E. Tabakowska, C. Ljungberg and O. Fischer (eds), 57–71. Amsterdam and Philadelphia: John Benjamins.
McLean, B. 2021. Revising an implicational hierarchy for the meanings of ideophones, with special reference to Japonic. *Linguistic Typology* 25: 507–549.
Monaghan, P., Shillcock, R. C., Christiansen, M. H. and Kirby, S. 2014. How arbitrary is language? *Philosophical Transactions of the Royal Society B* 369: 20130299.
Msimang, C. T. and Poulos, G. 2001. The ideophone in Zulu: A re-examination of conceptual and descriptive notions. In *Ideophones*, F. K. E. Voeltz and C. Kilian-Hatz (eds), 235–250. Amsterdam and Philadelphia: John Benjamins.
NHK Great East Japan Earthquake Archive. Tokyo: NHK (Japan Broadcasting Corporation). https://www9.nhk.or.jp/archives/311shogen/ (last accessed: 21 March 2021)
Nielsen, A. K. S. and Dingemanse, M. 2021. Iconicity in word learning and beyond: A critical review. *Language and Speech* 64: 52–72.
Occhino-Kehoe, C., Anible, B., Wilkinson, E. and Morford, J. R. 2017. Iconicity is in the eye of the beholder: How language experience affects perceived iconicity. *Gesture* 16: 99–125.
Ohala, J. J. 1984. An ethological perspective on common cross-language utilization of F_0 of voice. *Phonetica* 41: 1–16.
Ortega, G. 2017. Iconicity and sign lexical acquisition: A review. *Frontiers in Psychology* 8: 1280.
Ozturk, O., Krehm, M. and Vouloumanos, A. 2013. Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology* 114: 173–186.
Pejovic, J. and Molnar, M. 2017. The development of spontaneous sound-shape matching in monolingual and bilingual infants during the first year. *Developmental Psychology* 53: 581–586.

Peña M., Mehler J. and Nespor M. 2011. The role of audiovisual processing in early conceptual development. *Psychological Science* 22: 1419–1421.
Perniss, P., Lu, J. C., Morgan, G. and Vigliocco, G. 2017. Mapping language to the world: The role of iconicity in the sign language input. *Developmental Science* 21: e12551.
Perniss, P., Thompson, R. L. and Vigliocco, G. 2010. Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology* 1: 227.
Perry, L. K., Perlman, M. and Lupyan, G. 2015. Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PLoS ONE* 10: e0137147.
Ramachandran, V. S. and Hubbard, E. M. 2001. Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies* 8: 3–34.
Saalbach, H. and Imai, M. 2007. Scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General* 136: 405–501.
Saalbach, H. and Imai, M. 2012. The relation between linguistic categories and cognition: The case of numeral classifiers. *Language and Cognitive Processes* 27: 381–428.
Saji, N., Akita, K., Kantartzis, K., Kita, S. and Imai, M. 2019. Cross-linguistically shared and language-specific sound symbolism in novel words elicited by locomotion videos in Japanese and English. *PLoS ONE* 14: e0218707.
Saji, N., Asano, M. and Imai, M. 2020. Acquisition of the meaning of the word orange requires understanding of the meanings of red, pink and purple: Constructing a lexicon as a connected system. *Cognitive Science* 44: e12813.
Saji, N., Imai, M., Saalbach, H., Zhang, Y., Shu, H. and Okada, H. 2011. Word learning does not end at fast-mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition* 118: 45–61.
Sapir, E. 1929. A study in phonetic symbolism. *Journal of Experimental Psychology* 12: 225–239.
Shi, W. and Wu, Y. 2014. Which way to move: The evolution of motion expressions in Chinese. *Linguistics* 52: 1237–1292.
Sonesson, G. 1997. The ecological foundations of iconicity. In *Semiotics around the World: Synthesis in Diversity: Proceedings of the Fifth International Congress of the IASS*, I. Rauch and G. F. Carr (eds), 739–742. Berlin: Mouton de Gruyter.
Talmy, L. 2000. *Toward a Cognitive Semantics, Volume 2: Typology and Process in Concept Structuring*. Cambridge, MA: MIT Press.
Thompson, A. L., Akita, K. and Do, Y. 2020. Iconicity ratings across the Japanese lexicon: A comparative study with English. *Linguistics Vanguard* 6: 20190088.
Thompson, A. L. and Do, Y. 2019. Defining iconicity: An articulation-based methodology for explaining the phonological structure of ideophones. *Glossa: A Journal of General Linguistics* 4: 72. 1–40.
Wienold, G. 1995. Lexical and conceptual structures in expressions for movement and space: With reference to Japanese, Korean, Thai, and Indonesian as compared to English and German. In *Lexical Knowledge in the Organisation of Language*, U. Egli, P. E. Pause, C. Schwarze, A. Stechow and G. Wienold (eds), 301–340. Amsterdam and Philadelphia: John Benjamins.
Yamaguchi, N. 2019. *Onomatope no rekishi 1: Sono shujusō to shiteki suii; "obenchara" nado no goshi.* [The history of onomatopoeia 1: Their various characteristics and historical changes; the history of words like 'obenchara']. Tokyo: Kazama Shobo.