View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Warwick Research Archives Portal Repository

THE UNIVERSITY OF WARWICK

Original citation:

Imai, M. and Kita, Sotaro, 1963-. (2014) The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. Philosophical Transactions of the Royal Society B: Biological Sciences, Volume 369 (Number 1651). Article number 20130298. ISSN 0962-8436

Permanent WRAP url:

http://wrap.warwick.ac.uk/65926

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Link to published version: http://dx.doi.org/10.1098/rstb.2013.0298

A note on versions:

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP url' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: publications@warwick.ac.uk

warwickpublicationswrap

highlight your research

http://wrap.warwick.ac.uk

Accepted version of

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651). http://dx.doi.org/10.1098/rstb.2013.0298

(Figures at the end of the document).

The sound symbolism bootstrapping hypothesis for language acquisition and language evolution

Mutsumi Imai & Sotaro Kita

Keio University, University of Warwick

Abstract

Sound symbolism is a non-arbitrary relationship between speech sounds and meaning. We review evidence that, contrary to the traditional view in linguistics, sound symbolism is an important design feature of language, which affects on-line processing of language, and most importantly, language acquisition. We propose the *sound symbolism bootstrapping hypothesis*, claiming that (1) pre-verbal infants are sensitive to sound symbolism, due to a biologically-endowed ability to map and integrate multi-modal input; (2) sound symbolism helps infants gain *referential insight* for speech sounds; (3) sound symbolism helps infants and toddlers associate speech sounds with their referents to establish a lexical representation; (4) sound symbolism helps toddlers learn words by allowing them to focus on referents

embedded in a complex scene, alleviating Quine's problem [1]. We further explore the possibility that sound symbolism is deeply related to language evolution, drawing the parallel between historical development of language across generations and ontogenetic development within individuals. Finally, we suggest that sound symbolism bootstrapping is a part of a more general phenomenon of bootstrapping by means of iconic representations, drawing on similarities and close behavioural links between sound symbolism and speech-accompanying iconic gesture.

Sound symbolism is not peripheral in modern language

Since de Saussure's [2] highly influential work, the arbitrary relationship between form and meaning in words has been considered to be one of the "design features" of language [3] in traditional linguistics. This is supported by the fact that different languages assign different sounds to the same concept (e.g., "tree" in English vs. "arbre" in French) [2]. Though de Saussure acknowledged that onomatopoeias (words which imitate sounds, e.g., "bowwow" for dogs' bark) are exceptions to the arbitrariness principle, they were considered to be a marginal phenomenon in language. This view has been inherited in more recent writing [4]. For example, Newmeyer [5] writes that "the number of pictorial, imitative, or onomatopoetic nonderived words in any language is vanishingly small" (p.758).

Indeed, for most words in the lexicons of all languages, mapping between sound and meaning may seem arbitrary. However, sound symbolic words—those that have an inherent non-arbitrary link between sound and meaning—are more abundant than typically assumed, as we will review below.

Sound symbolism can be seen as a form of iconicity. Perniss and Vigliocco [6] define iconicity as resemblance between properties of a linguistic form and the sensori-motor and/or affective properties of referents. At least some types of sound symbolism show clear resemblance between properties of speech sounds and properties of their referents. For example, reduplication in Japanese sound symbolic words indicates repetition in the referent events (e.g., "goron" a heavy object rolling once, "gorogoro" a heavy object rolling repeatedly).

Many languages of the world have a large grammatically defined class of sound symbolic words (called "ideophones", "expressives", or "mimetics") in which the iconic relation between sounds and meaning is apparently felt by native speakers of the language and sometimes even by people who do not speak that language. Such a specialized word class exists in most of the East Asian languages (Japanese: [7, 8]; Korean: [9]; Cantonese: [10]), many of the Southeast Asian languages [11, 12], most of the sub-Saharan African languages [13, 14], some of the Australian Aboriginal languages [15,16], some of the South American languages [17, 18], and some non-Indo-European languages of Europe (Finnish and Estonian: [19], Basque: [20]). Sound symbolic word classes may contain thousands of words; for example, one dictionary of Japanese mimetics lists 4500 entries [21].

Sound symbolism can also be observed in conventionalized words that are not considered as “specialized sound symbolic words” (e.g., mimetics, expressives and ideophones). Berlin [22] noted that there is sound symbolism in names of two species of birds, rails and tinamous, in 17 languages spoken in indigenous tribes in South America. As seen in Figure 1, the shape of rails (Fig. 1a) is angular and sharp, whereas that of the tinamous (Fig. 1b) is round. In these languages, the words naming rails include a stop consonant [t] or [k], whereas the words naming tinamous birds tend to have nasals, which connote slowness, roundness, fatness, softness and heaviness.

-- insert Figure 1 about here--

Systematic sound-meaning correspondence exists in English, which does not have a special class of sound symbolic words. Some clusters of words in English with similar meanings have the same sounds at the beginning or the end (called phonoesthemes [25, 26]). For example, several words beginning with "gl" have meanings related to light: "glitter", "glare", "glow", and "glistening".

The psychological reality of sound symbolism has been well established, not only for speakers of languages with a grammatically defined class of sound symbolic words but also for those without such a word class. For example, Köhler [37] noted that certain sound- shape correspondences were judged to be good match: when novel words "maluma" and "takete" are presented as labels for a rounded vs. a spikey object, speakers of different languages judged "maluma,” to be more appropriate for the rounded object and "takeke" for the spikey object (see also [22], [37-41]). The size symbolism pointed out by Sapir [42] -- "mil" is judged to be more appropriate for the small object and "mal" for the large object-- has been empirically established as well across speakers of different languages ([22], [43]). It has also been shown that English speakers’ automatic lexical processing is affected by sound symbolism. For example, in a lexical decision task, English-speaking adults were faster to reject non-words when they sound-symbolically matched the shape of the frame in which the words appeared in (e.g., "kide" in a spikey shape, i.e., Köhler's shape sound symbolism [37]) than when they did not match [44].

Furthermore, in a large-scale corpus analysis, Monaghan et al. (this issue, [27]) demonstrated that sound-meaning mappings in the English lexicon are more systematic than would be expected by chance. That is, subtle sound symbolism, which people do not consciously detect, may exist throughout the conventional (i.e., non-mimetic) lexicon [27, 28][1]. Thus, form-meaning correspondence is not entirely arbitrary even in a language like English

Monaghan et al.’s [27] finding further suggests that sound symbolism is continuous rather than dichotomous, as some sound symbolic sound-meaning relations may be so subtle that people do not consciously detect under usual circumstances, while others can be more apparently detectable (e.g., sound symbolism in words like “thump” and “bump”, [26]). The degree of iconicity varies within specialized sound symbolic words as well. For example, Japanese mimetic words for sound (i.e., "wan-wan" for dog barking) are more iconic than mimetic words for perceptible motions (e.g., doshi-doshi) and object properties (e.g., tsuru-tsuru), which are, in turn, more iconic than those for mental states and emotions (e.g., uki-uki), according to Akita’s analysis [29].

The prevalence of sound symbolism may vary across different conceptual categories of words, reflecting the pressure from two directions—one toward expanding the vocabulary to accommodate needs to make finer contrasts, and the other toward maintaining iconicity (cf. [27,30])—as well as reflecting how easy it is to represent the referents by sound features . While numerous mimetics in Japanese make fine-grained distinctions in manners of actions, manners of physical sensations, and certain properties of objects (e.g., texture), few mimetics denote objects. This tendency has also been observed in other languages with specialized sound symbolic

---

[1] However, it should be noted that this systematicity found in Monaghan et al.'s study and in phonoesthemes may not always be a case of iconicity, as pointed out by Perniss and Vigliocco [6].

vocabulary [11, 31]. One exception to this tendency is the mimetics used in Japanese Infant Directed Speech (IDS), where caretakers extensively use onomatopoeias as object names (e.g., "wan-wan", dogs' barking, to refer to dogs) [32] in the early stage of language development, during which children's vocabularies are still small [33].

Object names are underrepresented among mimetics, in comparison to action, relation or property names. This may be because the need for fine discrimination among similar concepts is greater for object names than for actions, relations and object properties. It has shown that sound symbolism is useful for making contrasts among large conceptual categories such as grammatical word-class categories, it could impede word learning when there is a need to make fine discriminations among similar concepts [34]. One reason why fine discrimination is necessary for object names may be that nouns are more open to new entries, compared to predicate terms such as verbs and adjectives (e.g., [35]).

Object names may well have been iconic in their historical origin. For example, names of birds such as “karasu” (crow) and “gan”(goose) that are conceived as non-mimetic conventional lexical words in modern Japanese have onomatopoeic origins (e.g., mimicry of bird vocalizations) [36]. As the object name lexicon has grown over historical time, however, the original iconicity may have become obscure because preserving iconicity became disadvantageous as the words denoting similar objects were continuously added to the lexicon.

Importantly, however, Monaghan et al. [27] reported that the degree of sound-meaning systematicity was not any greater for nouns than for verbs in English. Thus, the distribution of sound symbolic words may be different for languages with specialized sound symbolic vocabulary like Japanese and for those without such vocabulary like English. How languages stand on the balance between iconicity and arbitrariness is an extremely important issue for future research.

In any case, recent findings from cognitive psychology, cognitive neuroscience, developmental psychology, cognitive and anthropological linguistics converge on the view that iconicity plays a core role for philogenesis and ontogenesis of language as well as for on-line language processing, as pointed out by Perniss and Vigliocco [6]. As such, sound symbolism is not a marginal phenomenon, but is an integral part of language.

The primary goal of this article is to discuss why the role of sound symbolism is important for language, with a special focus on the role of sound symbolism in language development and propose the *sound symbolism bootstrapping hypothesis*, which claims that sound symbolism provides a scaffolding mechanism for children in various stages of language development. In what follows in this article, we extensively discuss the details of the *sound symbolism bootstrapping hypothesis*. We then explore how sound symbolism could have played a role in our ancestors' proto-language.

The induction problem in word learning and sound symbolism bootstrapping hypothesis

During the first year of life, infants start to map speech sounds onto meaning, and in subsequent years they acquire a vast number of words that build a lexicon. How children acquire language—its ontogenesis and the developmental path thereafter— is still not entirely understood. How do they come to know, for example, that the sounds people make with their mouths are "words," and that they are names for objects, actions or properties? How do they learn the meanings of words? We will argue below that sound symbolism can help children overcome these challenges.

When a child hears a word in an everyday situation, the visual information they receive from the world is very rich, as well as unsegmented. Imagine that a child hears his mother say “Oh, look at the dog walking over there! ” when they pass by a dog with a leash walking with its owner. How would he identify the referent of “dog” or “walking” out of the extremely rich perceptual information that is changing from moment to moment? How can he be prevented from thinking that “dog” means a thing with four legs moving together with a person?

Even when a novel word is explicitly associated with an object in an unambiguous manner, and the child successfully connects the word to the referent in that scene, this success still does not allow children to use the word in new situations. For example, to be able to use the word “dog,” the correct visual identification of a particular dog in the observed scene is not sufficient for a child to judge what other things can also be called a “dog.” The child somehow has to have a visual as well as conceptual representation of “dog” to be able to determine whether other objects that carry similarity to the original referent (e.g., other dogs of a different kind, other small four legged animals, etc.) can also be considered as a referent of ”dog” (cf. [45]). Likewise, to be able to use the verb “to walk,” whose meaning is usually considered to be very concrete, one needs to know that it can be applied to a wide range of motion events, including toddlers tottering, a woman sashaying, an athlete walking very fast for speed walking, or a horse walking with four legs, but not to visually similar events with the same agents such as a human running or a horse galloping.

In other words, to be able to use a word—be it a noun, verb, or adjective—, children need to find the invariance in the contexts in which the word has appeared, working at first from a small number of exemplars. This induction is logically not possible when a child encounters a new word for the first time because there are too many possible ways of generalization from a single exemplar [1]; however, this is exactly what children face when they learn words [46-48].

A large body of research has addressed how children get around this problem. Young children recruit constellations of cues—conceptual, social, pragmatic and distributional—to constrain the inference of word meanings [47-50]. For example, children know that, in order to identify the referent, the speaker's eye-gaze or other social cue is useful [51, 52]. They also know that words appearing in different positions in a sentence and in different forms are mapped to different kinds of concepts such as objects, substances, actions or properties [53, 54]. In generalizing object names, for example, children know that words (count nouns) are extended on the basis of shape [55-57].

Not all cues are available from the earliest stages of lexical development. For example, word leaning biases such as the shape bias and the mutual exclusivity bias are likely to emerge through experience of word learning [56, 58]. For verb learning, it is not known whether children at the initial stage of lexical development can exploit the knowledge of the relation between the argument structure and the verb meaning (i.e., the syntax-semantics mapping) when inferring the meaning of a novel word. Furthermore, although this cue is helpful for mapping the verb to a rough, macro level concept (e.g., whether it should be mapped to a caused motion or a spontaneous motion [59-61]), it would not help children to find the differences among words that appear in the same argument structure (e.g., walking vs. running vs. hopping).

Thus, finding the meaning of a word is challenging for children, especially for infants who have few words in their vocabulary and cannot yet take advantage of cues that have to be learned through experience. What other cues are available for such infants? Here we propose that a biologically-endowed ability to detect sound symbolism provides one such cue.

We argue that sound symbolism helps children learn the meaning of words at different stages of early lexical development. In particular, this *sound symbolism bootstrapping hypothesis* consists of several related claims.

1. Children, even pre-verbal infants, are sensitive to sound symbolism, due to a biologically endowed ability to map and integrate multi-modal input, as suggested by Ramachandran and Hubbard [41] and Spector and Maurer [62].
2. Young children are sensitive to a wider range of possible sound symbolic correspondences than adults, but this sensitivity gets pruned and reorganised as they learn more words in their native language.
3. Sound symbolism helps infants who have just started word learning to gain the insight that speech sounds refer to entities in the world (i.e., the *referential insight* for speech sounds).
4. Sound symbolism helps infants associate speech sounds and their referents, and establish a lexical representation.
5. Sound symbolism helps toddlers identify referents embedded in a complex scene, alleviating Quine's problem [1].

Here, we review evidence for these claims. We further explore how the ontogenesis of lexical development might mirror the evolution of language in our ancestors and how sound symbolism relates to iconicity in sign language and speech-accompanying gesture.

Children's sensitivity to sound symbolism

In order for sound symbolism to bootstrap lexical development, children have to be able to detect it. Cross-linguistially recognised sound symbolism may have an especially natural correspondence between sounds and meanings. Thus, such sound symbolism may be the best place to explore whether children have sensitivity to sound symbolism at the start of, or even prior to, lexical development.

Toddlers' sensitivity to universal sound symbolism has been demonstrated by Maurer and colleagues in two-way forced choice tasks [40], using Köhler's shape sound symbolism [37]. Canadian toddlers were presented with a novel word (e.g., "kay-kee" or "boo-baa") and two line drawings, one of which had a rounded shape and the other had a spikey/jagged shape. They were able to pick the shape and the matching novel word at levels above chance [40].

Toddlers' sensitivity to cross-linguistically recognisable sound symbolism was also shown with Japanese 25 month olds [63] in the domain of manner of motion. The children were presented with a novel sound symbolic words and two video clips, each showing a person walking in a specific manner. They were asked to select the manner of walking to which the word referred. Only one of the videos sound symbolically matched the word, after being established through the results of prior experiments with Japanese speaking and English speaking adult participants (see Figure 2 for an example). The children were able to select the correct video at levels above chance.

Japanese 3-year-olds also use their sensitivity to cross-linguistic sound symbolism in generating novel sound symbolic words. When describing events involving rolling and jumping, the toddlers produced novel mimetics, along with conventional ones. English speaking adults with no knowledge of Japanese were able to guess which novel mimetics were used for which type of event (rolling or jumping) at above chance levels of accuracy [64]. This indicates that the novel mimetics produced by Japanese toddlers included cross-linguistically recognisable sound symbolism.

-- Insert Figure 2 about here--

Young infants can also detect cross-linguistically shared sound symbolism. Spanish-reared 3-month-olds are sensitive to the sound symbolism of vowels and size [65], that is, the association of high/mid-frontal vowels (/i/, /e/) and low/mid-posterior vowels (/o/, /a/) with small objects and large objects, respectively [42]. In a two-way preferential looking paradigm, infants were presented with a syllable (e.g., "di" vs. "do", or "de" vs. "da") with two geometric objects (e.g., ovals), which differed only in size and were presented side by side (thus one object was a sound-symbolic match, and the other was a mismatch). The infants looked at the sound-symbolically matching object (e.g., for "di", the smaller object; for "do", the larger object) longer than the mismatching object.

The results are mixed for very young infant's sensitivity to detect Kölher's sound symbolism [37] for shapes. In a study by Ozturk and colleagues [66], American 4-month-olds showed sensitivity to this sound symbolism. Using an infant-controlled sequential preferential looking paradigm, infants were presented with one of two words ("bubu" and "kiki") and one of two shapes (a rounded shape and a spikey shape). Results showed that infants looked longer at the sound symbolically mismatching shape than the matching one. In contrast, Forte and colleagues [67] conducted a series of experiments with French-learning 5- and 6-month-olds, who failed to show sensitivity to Köhler's sound symbolism [37]. Here, in a simultaneous preferential looking paradigm, infants were presented with a novel word (e.g., "buba" or "kite") along with two shapes side by side, one of which was a rounded one and the other a spikey one. There was no significant difference in looking times for the two objects. It is not clear why Ozturk et al.'s [66] and Fort et al.'s [67] results differed, but the discrepant results suggest that the effect of this type of sound symbolism in this age group may be fragile.

Development of language-specific vs. universal sound symbolism.

There has been an assumption in the literature that sound symbolism is universal; if a certain sound-meaning correspondence is identified by speakers of one language, this should be generalisable to speakers of any other languages. This assumption has been supported by the fact that speakers of many different languages sense Köhler's shape sound symbolism [37] in the same way, as reviewed earlier (English: [43]; Japanese: [68, 69]; Himba: [38]; Kitwonge-Swahili bilinguals: [70]). Furthermore, some aspects of sound symbolism in words in a given language can successfully be decoded by speakers of another language (Japanese sound symbolic words for laughing/smiling and for pain by English speaker [71, 72], see also Imai et al.'s result [63] for motion sound symbolism). The finding that people can correctly match a pair of antonyms in a foreign language to the corresponding pair of words in their native language [73-75] also endorses this view.

However, some sound symbolic words in a given language are opaque to adult speakers of other languages. Iwasaki et al. [71, 72] found that adult English speakers' judgments of conventional Japanese mimetic words for laughing and walking tended to converge with those of Japanese speakers on semantic dimensions concerning the magnitude (of size, sound), while they were quite different on evaluative dimensions (e.g., beauty, pleasantness), supporting the idea that some aspects of sound symbolism are universal, while others are language-specific.

Details of cross-linguistically common and language-specific aspects of sound symbolism were explored in a study in which English- and Japanese-speaking adults generated novel words for various manners of motion. Saji et al. presented various locomotion videos to Japanese and English speakers and asked them to generate a word that would sound-symbolically match each action, then rate that action on five semantic dimensions (size, speed, weight, energeticity, jerkiness) [76]. Results showed that certain sound-meaning links were common across the two languages.

For example, English- and Japanese-speakers shared voicing-speed and nasality-speed mappings (i.e., voiced = slow, voiceless = fast; nasal = slow, non-nasal = fast). This may be accounted for by the longer duration of vocal cord vibration in voiced and nasal stops (see [8, 77] for similar discussion). As voiced consonants have a shorter voice onset time, vocal cord vibration starts earlier and voicing is sustained for a longer duration than their voiceless counterparts. Nasal consonants can be prolonged without a change in quality. The longer voicing duration involved in these consonants and their non-turbulent nature appear to be readily mapped to the long duration of slow motion. Thus the sound-meaning associations shared by Japanese and English may be accounted for by our common bodily experience with articulation (see also Shinohara and Kawahara [78]) or audition.

These similarities do not mean that English and Japanese speakers always mapped sounds and meanings in the same way, however. In fact, even though the two languages used the same sound properties on the same semantic dimensions, the directionality of the sound-meaning mapping was sometimes reversed (see [75] for a relevant finding). For example, the affricate manner of articulation (e.g., the palato-alveolar affricate [tʃ]) was associated with light motions in Japanese, but with heavy motions in English. These disagreements may be explained by the cross-linguistic differences in the phonological status of these sounds. For example, in Japanese, the phone [tʃ] often appears secondarily, as a result of the palatalization process (in a context such as /ty/), whereas this is not the case in English. In any case, this result implies that language-specific sound symbolism exists, and that cross-linguistically shared and language-specific parts of sound symbolism are intricately intertwined within each language.

How, then, does the sensitivity to language-specific and universal sound symbolism develop in children? One possibility is that young children first detect only sound symbolism that are shared universally, and later learn language specific sound symbolism through learning of their native language. An alternative possibility, however, is that young children are sensitive to all possible sound symbolic correspondences that could appear in any language of the world, but only a subset of these correspondences are compatible with the phonological inventory and the existing words in the language the children are learning. As they grow up, the sensitivity to the incompatible correspondences wanes, and adults maintain only the sensitivity to the compatible correspondences. Thus, each language draws from a universal inventory of possible sound symbolic correspondences. Some sound symbolic correspondences (e.g., Köhler's shape sound symbolism and Sapir’s size sound symbolism) appear in many languages perhaps because they are strongly supported by iconic relationships between articulatory and/or acoustic features of speech sounds and the referents.

Evidence for the latter possibility is found in a study of sound symbolic sensitivity in English- and Greek-speaking adults and 3-year-olds [79]. In a pretest, adult speakers of Japanese, English and Greek rated the degree of sound-symbolic match between novel words and various manners of walking, similar to the ones in Figure 2. Based on the ratings, three types of items were selected: universal items (rated as a good sound symbolic match by the speakers of all three languages), English-specific items (rated as a good sound symbolic match by English speakers, but not by Greek and Japanese speakers), and distractor items (rated as a poor sound symbolic match by the speakers of all three languages). Then, adult and child speakers of English and Greek (adult speakers were different from the ones in the pretest) were presented with a novel word and a pair of manners of walking. In the universal condition, one manner of walking was a *universal* item (target) and the other was a distractor item. In the English-specific condition, one,manner of walking was *English-specific* and the other was a distractor.

Not surprisingly, adult and child English-speakers correctly chose the target in both the universal and English-specific conditions. Adult Greek speakers correctly chose the target at levels above chance only in the universal condition. Crucially, Greek-speaking children could correctly choose the target video in both universal and English-specific conditions. That is, Greek-speaking children were sensitive to a wider range of sound symbolic correspondences than Greek-speaking adults. This suggests that Greek-speaking adults had lost sensitivity to sound symbolic correspondences supported by properties of English but not by properties of Greek.

This narrowing of sound symbolic sensitivity during development may be analogous to the way language specific phonemic contrasts are acquired. Up to 10-12 months, infants are sensitive to phonemic contrasts in a foreign language that their caregivers have long before lost sensitivity to [80]. For example, infants growing up in an English-speaking environment can distinguish Hindi contrasts (i.e., dental and alveolar stops) that are not contrastive in English and thus English-speaking adults cannot distinguish.

How do children become tuned into the system of sound symbolism in their native language? Here, it is relevant to note that sound symbolic words in a specialized word class, such as Japanese mimetics, are very difficult for adult second language learners to master [71, 72]. To acquire a native speaker's sensitivity and productive competence— that is, to be able to comprehend and use conventional and novel mimetics productively and creatively—may require extensive exposure to mimetics used in real contexts. This suggests that, even though some aspects of sound symbolism may be biologically grounded, it is crucial to have intensive exposure to a specific language in early stages of development. Through statistical learning [81], young children may be able to extract patterns of form-meaning co-occurrences in the words they have learned [27] and abstract out language-specific aspects of sound symbolism. Such learning experiences may result in sound-symbolic words that have both universal and language-specific components.

The studies reviewed so far provide evidence for sound- symbolic sensitivity, but these studies do not provide evidence that sound symbolism is directly related to the ontogenesis of language. The following infant EEG study addresses this issue.

Neural response to sound symbolism in 11-month-old infants

Sound symbolism may arise from the sense of similarity between speech sounds and other types of information through naturally occurring cross-modal mapping. Due to dense connectivity across different sensory brain regions, infants may spontaneously map perceptual experiences across different modalities onto speech sounds (for a review, see [82]). Human infants can already map information in different modalities in the way adults would. For example, they can map size and numerosity [83], or acoustic properties of speech and non-sounds onto properties of visually presented objects [84, 85]. This cross-modal mapping ability may not be limited to humans, as chimpanzees can map auditory pitch and luminance [86]. This perceptual ability may later develop into a more abstract system of symbols embodied in language [87, 88] both ontogenetically and phylogenetically.

Is infants' ability for cross-modal mapping directly linked to language processing, and if so, how? Asano et al. [68] (see also [89]) investigated this question with Japanese-reared 11-month-olds in an EEG study. In each trial, infants were presented with a picture of a shape (spikey or rounded) followed by a novel word ("kipi" or "moma"). The word-shape pairs were either sound symbolically matching or mismtching (Köhler's shape sound symbolism [37]).

The recorded EEG was submitted to averaging (ERP, "event related potentials")) and large-scale phase synchronization analyses to explore (a) whether 11-month-old infants detect sound symbolism and treat sound-symbolically mismatching words as semantically unexpected; and (b) how different regions of the infant brain communicate while sound-symbolically matching and mismatching words are processed. Growing evidence indicates that large-scale synchronous neural oscillations play an important role in dynamically linking multiple brain regions in adults, something which presumably reveals functional communication among these regions [90-97].

Concerning the ERP pattern, infants responded differently to the sound symbolically matching and mismatching word-shape pairs. The timing and topography were similar to the typical N400 effect, with a stronger negative deflection for the mismatching pairs at about 400 ms after the stimulus onset [98], an index of semantic integration difficulty [98-101]. Second, phase synchronization of neural oscillations (phase locking value, *PLV*) increased (as compared to the baseline period) significantly more in the mismatch condition than in the match condition. This effect was observed in the beta band and most pronounced over left-hemisphere electrodes during the time window of the N400 (301-600ms). The time course of large scale synchronization suggests that cross-modal binding was achieved quickly in the match condition, but sustained effort into the time range of the N400 effect was required in the mismatch condition. An additional brain oscillation analysis showed an increase of early (< 200 ms latency) gamma-band oscillations in the match condition compared to the mismatch condition. A number of adult studies have revealed that early gamma-band activity is related to multisensory integration (see Senkowski, Schneider, Foxe, & Engel, 2008 for a review).

In a different study [102], when adult participants were presented with real words and non-words in isolation, real words elicited strong EEG coherence in the beta-band in the left hemisphere, in comparison to the resting state, but non-words did not do so. The stronger inter-regional communication in the left hemisphere in Asano et al.'s infants [68] thus may indicate that the sound-shape pairings were processed in the language-processing network (in the left-hemisphere) in 11-month-old infants.

Taken together, the results from ERP and phase synchronization analyses together suggest that 11-month-olds could clearly detect Köhler's shape sound symbolism [37], and further suggests that sound symbolic associations fosters multi-sensory integration and semantic processing. When infants at this age are presented with a spoken word and a visual stimulus, they already attempt to integrate the two stimuli and establish a lexical representation. This process requires substantial effort when the word and visual stimuli do not sound symbolically match. However, when the two stimuli matched sound symbolically, the sound-vision integration comes naturally due to the iconicity between the sound and visual stimuli, leading to a nascent representation of the word meaning without effort. This may help infant realize that words stand for concepts; that is, it may provide infants with referential insight for speech sounds.

It is difficult to know whether the 11-month-old infants in Asano et al.' study had already gained this insight in some form. Given that they showed N400 response, it could be that at 11-months, these infants already assumed that the two speech sounds *referred* to the objects, and thought that the round shapes and the angular shapes were anomalous as referents of the word "kipi" and "moma", respectively. However, even if 11-month-olds already had the referential insight in a nascent form through previous word learning experiences, they would need much further experience of word learning to solidify it. If they are able to identify the referents of a word due to sound symbolism and infer the meaning of words without using the spontaneous multi-modal binding ability, this should provide important bootstrapping experiences for infants who scarcely have any words in their vocabulary.

Sound symbolism scaffolds acquisition of word meaning

Establishing word-referent associations.

One of the key claims of the sound symbolism bootstrapping hypothesis is that sound symbolism facilitates children's word learning. Previous research has shown that at 11 months, the language processing network in the infant's brain responded to sound symbolism [68], but it is not clear whether they were able to establish the word-referent associations and retain it in memory.

We demonstrated that sound symbolism facilitates word learning in 14-month-old Japanese-speaking infants [69]. The infants were tested in a word learning task which combined habituation and preferential looking. They were repeatedly presented with two word-shape pairs. For a half of the infants, word and shape sound symbolically matched ("moma" for a round shape and "kipi" for a spiky shape); for the other half, they mismatched ("moma" for a spiky shape and "kipi" for a round shape) [37]. After infants had been habituated, they heard either "kipi" or "moma" and saw the two shapes side by side. Infants looked at the shape that had been associated with the word during habituation faster and longer in the sound symbolically matching condition than in the mismatch condition. This suggests that sound symbolism helps children to learn word-referent associations at 14 months of age.

Helping children find the invariance for generalization

As discussed earlier, the establishment of word sound-referent associations is not sufficient for word meaning acquisition. For children to be able to use a word in new situations, they need to extract invariance across referents (i.e., to create a word meaning representation). However, this is extremely difficult to achieve from a single or a limited number of exposures, because what the child sees is very rich and contains a great deal of information that is not part of the meaning of the word.

This problem is particularly serious for verb learning as compared to the learning of object names. Unlike objects, actions are ephemeral and difficult to individuate [35, 103], so it is not obvious when the action referred to by a given word starts and when it ends. Finding a referent in a spatio-temporally changing scene itself is not easy, and indeed, young children become able to associate a verb to the referent (i.e., the action) later than for nouns in experimental settings [104, 105]. To be able to use the verb in different situations, children further need to understand which specific aspect of the action is invariant for the verb and which aspects can vary across different situations in which the verb is used [103, 106-108].

As reviewed earlier, it is difficult for preschool age children to extend a novel verb to a new scene in which the action is the same, but the agent (or the theme object, or the instrument of the action) in the original scene is replaced with a new one [103, 106, 109-111]. Imai et al. [63] and Kantartzis et al. [112] tested whether sound symbolism would help Japanese- and English-speaking 3-year-olds find the invariance for a newly taught verb in action events for successful generalization. Children were assigned to one of three conditions and were taught novel verbs while observing a person walking in different manners. In the experimental condition, a novel verb, which had been created by modifying an existing Japanese mimetic word (i.e., a sound symbolic word) was paired with a manner of walking that sound symbolically matched. For example, for a fast walk with small steps, the novel mimetic *choka-choka* was created from the existing Japanese mimetic *choko-choko*, and presented as a verb ("choka-choka-shiteru"in Japanese and "doing choka-choka" in English) (see also Figure 2 for another example). In the first control condition, the mimetic-based nonsense word was paired with a different motion that did not sound-symbolically match. In the second control condition, a nonsense word that resembled a typical monosyllabic verb in Japanese and English (e.g., *neke-tteiru* or *fepping*, respectively) was paired with the same motion from the experimental condition; as a non-mimetic word it did not sound-symbolically match.

Replicating results from previous studies [103, 106, 110, 111], in the two control conditions with novel words that did not sound symbolically match their referent actions, both Japanese and English 3-year-olds failed to generalize the newly taught verb to the identical action performed by a different actor. However, when the novel verb sound-symbolically matched the action, not only Japanese 3-year-olds but also English-reared 3-year-olds (who were not familiar with the sound symbolic system of Japanese mimetics) were able to use this cue to generalize the verb to a new event (see also Yoshida' s study [113] for similar findings).

Thus, regardless of the language they were acquiring, sound symbolism helped the children to find the relevant invariance in the scene for the verbs. Here, as noted earlier, young children are sensitive to a broader range of sound symbolism [40, 79], including sound symbolism that adults speaking the same language might not detect. Thus, they may be more likely than adults to take advantage of sound symbolism in word learning.

Sound symbolic words in Child-Directed-Speech

Another support for the sound symbolism bootstrapping hypothesis comes from studies investigating sound symbolic words in Child Directed Speech (CDS). Caretakers often use sound symbolic words with CDS for young children in a way that is appropriate for the children's language comprehension ability. In a classic study, Fernald and Morikawa [114] noted that Japanese mothers used sound symbolic words such as onomatopoeia/mimetics frequently when talking to Japanese infants.

Saji and Imai [115] further studied how Japanese caretakers of 2- and 3-year-old children use mimetics when describing action events in semi-experimental settings. A mother and her child were presented with animated videos of everyday actions that could be expressed either by a conventional verb or a mimetic (e.g., clapping hands, cutting a piece of paper with scissors), and the mother was asked to describe the video to her child. After that, the mother was also asked to describe the video to the experimenter. The mother used mimetics more often for younger listeners, that is, most often to 2-year-olds, the second most often to 3-year olds, and least often to the adult experimenter.

The caretakers also adjusted the way they used the mimetics according to their child's comprehension ability. For example, the mothers in Saji and Imai's study [115] used mimetics in different syntactic frames depending on their listener's age. When they were talking to 2-year-olds, they used a mimetic word by itself without incorporating it in a sentence, exactly when the motion started in the animation, often with accompanying gestures depicting the motion. In contrast, when they were talking to 3-year-olds, they tended to incorporate mimetics into a sentence; they either attached the light verb –suru ("to do") to the mimetic, or they used it adverbially to modify a conventional non-mimetic verb. When describing the action video to the experimenter, they mostly used conventional non-mimetic verbs.

The same pattern was also found in a longitudinal corpus study [116] of a father's input to a boy between the age of 8 months and 36 months. In addition to the syntactic frame change seen in Saji and Imai's study [115], the father adjusted the choice of words for action reference along with the child's development. As the child developed from no use of mimetics to productive use of mimetics, the father's input shifted from mostly mimetics alone, to mimetics plus semantically similar conventional verbs, and finally to mostly conventional verbs alone. These studies suggest that Japanese caretakers adjust their input closely according to their child's level of lexical development, and they used mimetics as a tool for the adjustment.

In parallel to caretakers' more frequent use of mimetics with younger children, younger children produce more mimetics than older children and adults. When Japanese 3-year-olds, 5-year-olds and adults described motion events, the manner of motion was described either by a conventional verb or with a mimetic. The proportion of descriptions using mimetics was higher for younger participants [64].

In the production-elicitation studies by Saji and Imai [115], when caretakers used onomatopoeic/mimetic words without embedding them in a sentence to 2-year-olds, the onomatopoeic/mimetic words were often uttered exactly when the referent action took place. Previous studies demonstrated that parents use various devices to help children connect the word and the referent. For example, parents often point to the object [117,118], and children tend to learn the name better when the referent object was pointed at in the past [119]. Temporal co-occurrence of the referent and labelling is also important [120, 121].

However, although pointing and temporal occurrence can help children notice the here-and-now word-referent relation, they do not necessarily help them to use the word beyond that particular context. Sound symbolism, in contrast, can directly link the word form and the word meaning, and thus it can help children extract invariance of the word meaning, which is crucial for using the word in new situations. In this sense, sound symbolism would be a more powerful cue for word meaning than pointing and temporal co-occurrence of a word and its referent.

Caretakers also use different types of sound symbolic words, depending on the child's stage of language development [115]. As noted earlier, some mimetic words (e.g., onomatopoeias) are considered to be more iconic than others (e.g., mimetics denoting object properties, manners of action, emotion) (cf. [29]). When their child have a very small number of words in their vocabulary, caretakers mostly use onomatopoeias—direct mimicry of sounds (e.g., "chirin- chirin", mimicry of a bicycle bell when referring to a bicycle) [114]. When children become more experienced in word learning, caretakers no longer use such onomatopoeic words for labelling objects and instead use conventional object names. However, caretakers continue to rely on sound symbolism when their child still shows difficulty in extracting invariance of word meaning. When children become advanced word learners, they no longer use mimetic words in the contexts in which adults would use non-mimetic, conventional words [116].

Does sound symbolism always facilitate word learning?

Monaghan and colleagues [34, 122] argue that sound symbolism can inhibit one-to-one mappings when each mapping competes with a similar mapping, but can help make class-to-class associations (e.g., words with certain phonological features refer to actions, and words with other phonological features refer to objects); hence, sound symbolism is beneficial only for grouping words into a large cluster (e.g., identifying grammatical category of individual words) and it impedes learning of individual words. However, many of infants' first words are basic-level terms (e.g., "bird") that refer only to category prototypes (robins, but not penguins). Thus, young infants are not likely to face the kind of situations in which sound symbolism is detrimental to word learning, that is, situations in which phonologically similar words have to be mapped onto similar shape categories, as suggested by Monaghan et al [27]. Sound symbolism may instead help infants identify the particular part of an ambiguous scene as the referent: When hearing *kipi*, infants may selectively attend to pointiness in visual shapes, which may help to guide them to the correct referent.

We agree that sound symbolism does not always help or sometimes even impedes word leaning. Learning the meaning of a new word may be impeded if another similar sounding word with similar meaning is activated in children's mind. Also, if sound symbolism is too powerful and children always map a word onto only the element in the scene that is most salient due to sound symbolism, then it would be difficult to build a large lexicon. Eventually children need to learn to pay attention to other cues as well, and thus the relative influence of sound symbolism may decrease as their lexical development proceeds and children become able to access many other cues for the inference of word meanings, as discussed in the previous section.

However, the novel verb learning studies [63,112, 113] indicates that sound symbolism could be helpful for word learning even for preschool age children. The facilitative role of sound symbolism for individual word learning may differ across different classes of words and across different developmental stages. That is, sound symbolism could impede learning of nouns when the noun vocabulary is sufficiently large, but it could still facilitate learning of verbs. This is because verbs do not compete with each other as much because verbs tend to carve up the world in less fine-grained way than nouns, and consequently, languages typically have fewer verbs than nouns. The same facilitative effect may also extend to names of properties (adjectives and adverbs), as learning a property name involves mapping a word to a single property of an object out of the multitude of properties (size, texture, color, weight, speed, abruptness, etc.), which also poses a challenge for young children [123].

An important unanswered question is how children learn words that are *not* sound-symbolic and yet make use of sound symbolism when, and only when, the word carries sound-symbolic properties. It is possible that, through experience in language learning, children quickly learn that words are not always sound-symbolic and are willing to form word-referent associations even when they do not detect sound symbolism between the word and the referent, especially for object names. However, when children do detect sound symbolism in learning a novel word, they take advantage of it, and this additional cue is especially helpful for the learning of names for actions and properties, which are especially challenging for children.

Implications for theories of language evolution

Researchers have often discussed the possibility that the process of language development in modern-day children mirrors how language was started by our distant ancestors and evolved through history. Some have even speculated that sound symbolism in a modern-day language may be a vestige of protolanguage that was mostly sound symbolic [124, 125] and hypothesized that symbolic use of cross-modal mapping is one of the key steps in language evolution.

The emergence of sound symbolic words and how they began to be used as symbols by our ancestors has been debated by researchers. One possibility is that the motor system played a key role. It has been argued that a critical foundation for language evolution was humans' ability to mimic the external world [126]. In the course of evolution, sound symbolism may have arisen as mimicking of events and object properties in the external world with movements of lips and the tongue [22, 62]. For example, the size sound symbolism for vowel heights may be based on the oral cavity size mimicking the referent object size [42]. In contrast to these motor-based accounts of sound symbolism, some argued that cross-modal mapping between audition and other modalities is the key to sound symbolism. For example, the size sound symbolism of vowel quality (higher-front = small, lower-back = big) [42] and consonants (e.g., voiceless = small, voiced = big) could be explained based on acoustic frequency of speech (more acoustic energy in higher frequency = small, in lower frequency = big) [127]. These accounts of sound symbolism do not have to be mutually exclusive; instead, they could be thought of reflecting different forms of iconicity [6]. Be it motor-based or audition-based, sound symbolism may have set the foundation for using speech sound to systematically refer to concepts.

How might sound symbolism have played a role in the bootstrapping process for emergence of a modern-day voluminous lexicon in language evolution? First, the very idea that speech sound could have been used to refer to objects and events in the world may have arisen due to intrinsic and biologically endowed association between speech and information in other modalities (e.g., vision). The awareness could have further brought “referential insight” to our ancestors – that is, the insight that oral sounds can be used to symbolize things that are not the oral sounds themselves, including things that are not present at the speech event (i.e., "displacement" in Parniss and Vigliocco [6]). Second, sound symbolic associations may have helped our ancestors quickly build a shared lexicon that can be intuitively understood by members of the community, which would have promoted the use of “words” (speech sounds) for their primary medium of communication [124].

Sound symbolism may have played a further role in the emergence of combinatoric structure in language [124]. Because phonetic features and other units smaller than a word can carry sound symbolic meanings, a word can have a complex meaning that combines the meanings of parts of the word. To illustrate this in existing Japanese sound symbolic words, the word "gorogoro" refers to a heavy object rolling repeatedly. This word contrasts with the following related words: "goron" = a heavy object rolling once, "korokoro" = a light object rolling repeatedly, "koron" = a light object rolling once, "guruguru" = a heavy object rotating around an axis repeatedly, "gurun" = a heavy object rotating around an axis once, "kurukuru" = a light object rotating around an axis repeatedly, "kurun" = a light object rotating around an axis once. Here, the sequence of a velar stop plus /r/ indicates rotational movement. The voicing of the initial consonant indicates a heavy object. The word final "n" indicates that the event is completed after a single rotation, and has been analyzed to symbolize reverberation [77]. The reduplication indicates a repetitive event. That is, the meanings of these words are a combination of component meanings. This compositional nature of sound symbolic words may have facilitated a transition from a "holophrase", a single unanalyzable (monomorphemic) word with complex meaning [128], to a complex word with a morphological structure and combinatoric semantics. This principle of combinatoric semantics can subsequently be extended to words in a sentence [129]. This development would further expand the expressive power of protolanguage.

As discussed earlier, language is shaped by two competing forces: one towards arbitrariness and the other towards iconicity [30]. When the size of the lexicon becomes large and different words are used to make fine grained contrasts for similar concepts, it will become difficult for language to maintain sound symbolism for all words and it could impede growth of the lexicon, as pointed out by Monaghan et al [34]. Thus, in modern languages, sound symbolism may often be a very subtle tendency in the lexicon, which may not be consciously detectable even by naive speakers, unless a language has a clearly defined word class dedicated to sound symbolic words (e.g., mimetics, expressives, ideophones in various languages of the world).

Relation to the iconicity in gestures and sign language

The idea that motivated form-meaning relations facilitate lexical development can be extended to gesture and sign language. In British Sign Language, signs that are judged as iconic by adult raters (the sign form resembles the meaning, e.g., bringing a

C-shaped hand closer to the mouth for the sign "drink") were learned earlier than non-iconic arbitrary ones [130]. Speech-accompanying iconic gestures can also guide word learning in hearing children. When 3-year old English-reared children were presented with a novel verb and a complex action scene, along with an iconic gesture, children interpreted the verb's referent to be the part of the scene depicted by the iconic gesture [131]. That is, iconic gestures guided children to pick out a particular part of a complex scene as the referent of a novel verb.

Sound symbolism has a direct link to iconic gestures. When Japanese speakers produce mimetics during description of motion and action, they tend to produce a co-expressive iconic gesture at the same time [7, 132] (see also [133-135] for further discussions). This tendency is stronger in children (3 year olds) than in adults [64]. Such a link suggests that a common imagistic representation underlies both sound symbolic words and iconic gestures [7].

The close tie between sound symbolic words and iconic gestures has implications for theories of language evolution. The common-underlying-representation view suggests that sound symbolic words and iconic gestures emerged together in the course of language evolution [64]. As discussed earlier, the ability to use cross-modal non-arbitrary mappings to create symbols may have given rise to a communication system, which was consisted mainly of tightly linked sound symbolic words and iconic gestures. This contrasts with a gestural origin theory of language evolution, which states that protolanguage based on iconic gestures (without speech) preceded spoken protolanguage [136-138]. The gesture-first-theory has a difficulty in explaining the close tie between sound symbolic words and iconic gestures in modern humans.

The use of motivated form-meaning mapping may be an important foundation for the human symbolic ability. Be it sound symbolism or other types of iconicity in gestures and signs, children are equipped with an ability to readily take full advantage of it to crack the code of language.

Conclusion

Contrary to the traditional view that sound symbolism is a peripheral phenomenon in language, sound symbolism is widely observed in languages in the world. We have argued that sound symbolism facilitates lexical development in children. We reviewed the following lines of evidence for the *sound symbolism bootstrapping hypothesis*. Preverbal infants detect sound symbolism in unfamiliar words and process them as if they were real words, which may lead them to (or solidify) the realization that speech sounds have meanings. Sound symbolism further scaffolds word learning from infancy to early childhood, helping children to establish word-referent associations and also to extract the word meaning invariance from rich and unsegmented perceptual information children observe when they hear a word.

The impact of sound symbolism in children's language development may be surprising, given that most words are not apparently sound symbolic. We suggested that sound symbolism is a vestige of a protolanguage that was mostly sound symbolic. Sound symbolism may have helped our ancestors to develop their lexicon and also combinatoric nature of language. Furthermore, because of the tight link between sound symbolic words and co-speech iconic gestures, we also suggested that sound symbolism and iconic gestures, both of which involve non-arbitrary cross-modal mapping, evolved together. Children today still maintain the ability to take advantage of sound symbolism in word learning. Once children break into the system of linguistic symbols and start building lexicons with the help of iconic relationship between sound and meaning, they come to realize that many words do not have apparent form-meaning resemblance. When their vocabulary becomes substantially large, they may no longer expect sound symbolism in words even though there may be covert sound symbolism all across the lexicon [27].

To summarise, sound symbolism provides key insights into how language develops in children and how language evolved in human history. It should no longer be considered to be a peripheral phenomenon in language.

Acknowledgements

This research was supported by MEXT KAKENHI (#15300088, #22243043, Grant-in-Aid for Scientific Research on Innovative Areas #23120003) to Imai and BBSRC Research Development Fellowship (BB/G023069/1) to Kita. We would like to thank Gabriella Vigliocco, Stephanie Archer, Kimi Akita and anonymous reviews for her valuable comments on the manuscript.

References

1. Quine, W. V. O. 1960 Word and object. Cambridge, MA: The MIT Press.
2. de Saussure, F. 1983 Course in general linguistics [R. Harris, trans]. La Salle, IL: Open Court.
3. Hockett, C. 1960 The origin of speech. Scientific American 203, 89-97.
4. Pinker, S. 1999 Words and rules: The ingredients of language. New York: Norton.
5. Newmeyer, F. J. 1993 Iconicity and generative grammar. Language 68, 756-796.

6. Perniss, P. & Vigliocco, G. 2014 The bridge of iconicity: From a world of experience to the experience of language. Philosophical Transactions of the Royal Society B Biological Sciences, XX.
7. Kita, S. 1997 Two-dimensional semantic analysis of Japanese mimetics. Linguistics 35(2), 379-415. (DOI 10.1515/ling.1997.35.2.379)
8. Hamano, S. 1998 The sound-symbolic system of Japanese. Stanford, CA & Tokyo: CSLI & Kuroshio.
9. Kim, K. O. 1977 Sound symbolism in Korean. Journal of Linguistics 13, 67-75.
10. Bodomo, A. 2006 The structure of ideophones in African and Asian languages: The case of Dagaare and Cantonese. In Selected proceedings of the 35th annual conference on African linguistics (eds. J. Mugane, J. P. Hutchison & D. A. Worman), pp. 203-213. Somerville, MA: Cascadilla Proceedings Project.
11. Watson, R. L. 2001 A comparison of some Southeast Asian ideophones with some African Ideophones. In Ideophones (eds. F. K. E. Voeltz & C. Kilian-Hatz), pp. 385-405. Amsterdam: John Benjamins.
12. Diffloth, G. 1976 Expressives in Semai. In Austroasiatic Studies Vol. Part I (eds. P. N. Jenner, L. E. Thompson & S. Storasta), pp. 249-264. Hawaii: University of Hawaii Press.
13. Childs, G. T. 1994 African ideophones. In Sound symbolism (eds. L. Hinton, J. Nichols & J. J. Ohala), pp. 178-206. Cambridge: Cambridge University Press.
14. Samarin, W. J. 1971 Survey of Bantu ideophones. African Language Studies 7, 130-168.
15. Alpher, B. 1994 Yir-Yoront ideophones. In Sound symbolism (eds. L. Hinton, J. Nichols & J. J. Ohala), pp. 161-177. Cambridge: Cambridge University Press.
16. Schultze-Berndt, E. 2001 Ideophone-like characteristics of uninflected predicates in Jaminjung (Australia). In Ideophones (eds. F. K. E. Voeltz & C. Kilian-Hatz), pp. 355-374. Amsterdam: John Benjamins.
17. Derbyshire, D. C., & Pullum, G. (eds.) 1986 Handbook of Amazonian Languages, Vol. 1. Berlin: Mouton de Gruyter.
18. Nuckolls, J. B. 2004 To be or not to be ideophonically impoverished. Texas Linguistic Forum 47, 131-142.
19. Mikone, E. 2001 Ideophones in the Balto-Finnic languages. In Ideophones (eds. F. K. E. Voeltz & C. Kilian-Hatz), pp. 223-233. Amsterdam: John Benjamins.

20. Ibarretxe-Antuñano, I. 2006 Sound symbolism and motion in Basque. München: Lincom.
21. Ono, M. (ed). 2007 Nihongo onomatope jiten [Japanese Mimetics Dictionary]. Tokyo: Shogakkan.
22. Berlin, B. 2006 The first congress of ethnozoological nomenclature. Journal of the Royal Anthropological Institute 12(s1), S23-S44.
23. Sclater, P. L. & Salvin, O. 1878 Descriptions of three new species of birds from Ecuador. In Proceedings of the Zoological Society of London 46(1), 438-440. Blackwell Publishing Ltd.
24. Salvadori, T. 1895 Catalogue of the birds in the British Museum 27, London.
25. Bergen, B. K. 2004 The psychological reality of phonaethemes. Language 80(2), 290-311.
26. Firth, J. R. 1935 The use and distribution of certain English sounds. English Studies 17(1), 2-12.
27. Monaghan, P., Shillcock, R. C., Christiansen, M. H. & Kirby, S. 2014 How arbitrary is language? Philosophical Transactions of the Royal Society B Biological Sciences, XX.
28. Farmer, T. A., Christiansen, M. H. & Monaghan, P. 2006 Phonological typicality influences on-line sentence comprehension. Proceedings of the National Academy of Sciences 103(32), 12203-12208. (DOI 10.1073/pnas.0602173103)
29. Akita, K. 2009 Gradient integration of sound symbolism in language. Japanese and Korean Linguistics 17, 221-230.
30. Pernis, P., Thompson, T., & Vigliocco, G. 2010 Iconicity as a general property of language: Evidence from spoken and signed languages. Frontiers in Psychology 1(227), 1-15. (DOI 10.3389/fpsyg.2010.00227)
31. Bartens, A. 2000 Ideophones and sound symbolism in Atlantic creoles. Helsinki: Finnish Academy of Sciences and Letters.
32. 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.
33. Ogura, T. 2006 How the use of ‘non-adult words’ varies as a function of context and children’s linguistic development. In Studies in language science (5): Papers from the fifth annual conference of the Japanese society for language science

(eds. M. Nakayama, M. Minami, H. Morikawa, K. Nakamura, & H. Sirai), pp. 103-120. Tokyo: Kurosio Publishers.

34. Monaghan, P., Mattock, K. & Walker, P. 2012 The role of sound symbolism in word learning. Journal of Experimental Psychology: Learning, Memory and Cognition 38, 1152-1164.
35. Gentner, D. 1982 Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In Language development: Vol. 2. Language, thought, and culture (ed. S.A. Kuczaj), pp.301-334. Hillsdale, NJ: Erlbaum.
36. Yamaguchi, N. 1989 Chinchin chidori no naku koe wa: Nihonjin ga kiita tori no koe [How Japanese people heard and mimiced bird vocalizations]. Tokyo: Taishukan Shoten.
37. Köhler, W. 1947 Gestalt psychology, 2nd edn. New York: Liveright Publishing Corporation.
38. Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J. & Spence, C. 2013 “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. Cognition 126(2), 165-172. (DOI 10.1016/j.cognition.2012.09.007)
39. Kovic, V., Plunkett, K. & Westermann, G. 2010 The shape of words in the brain. Cognition 114(1), 19-28.
40. Maurer, D., Pathman, T., & Mondloch, C. J. 2006 The shape of boubas: sound-shape correspondences in toddlers and adults. Developmental Science 9(3), 316-322.
41. Ramachandran, V. S. & Hubbard, E. M. 2001 Synaesthesia--a window into perception, thought and language. Journal of Consciousness Studies 8(12), 3-34. (DOI 10.1111/1468-0068.00363)
42. Sapir, E. 1929 A study in phonetic symbolism. Journal of Experimental Psychology 12, 225-239.
43. Thompson, P. D. & Estes, Z. 2011 Sound symbolic naming of novel objects is a graded function. The Quarterly Journal of Experimental Psychology 64(12), 2392-2404.
44. Westbury, C. 2004 Implicit sound symbolism in lexical access: Evidence from an interference task. Brain and Language 93, 10-19.

45. Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. 2010 The effects of feature-label-order and their implications for symbolic learning. Cognitive Science 34, 909-957.
46. Markman, E. M. 1989 Categorization and naming in children: Problems of induction. Cambridge, MA: The MIT Press.
47. Imai, M. & Gentner, D. 1997 A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. Cognition 62(2), 169-200. (DOI 10.1016/S0010-0277(96)00784-6)
48. Hollich, G., Golinkoff, R. M. & Hirsh-Pasek, K. 2007 Young children associate novel words with complex objects rather than salient parts. Developmental Psychology 43(5), 1051-61. (DOI 10.1037/0012-1649.43.5.1051)
49. Haryu, E. & Imai, M. 2002 Reorganizing the lexicon by learning a new word: Japanese children's interpretation of the meaning of a new word for a familiar artifact. Child Development 73(5), 1378-1391. (DOI 10.1111/1467-8624.00478)
50. Imai, M. & Haryu, E. 2001 Learning proper nouns and common nouns without clues from syntax. Child Development 72(3), 787-802. (DOI 10.1111/1467-8624.00315)
51. Akhtar, N. & Tomasello, M. 1996 Two-year-olds learn words for absent objects and actions. British Journal of Developmental Psychology 14(1), 79-93.(DOI 10.1111/j.2044-835X.1996.tb00695.x)
52. Baldwin, D. A. 1993 Infants' ability to consult the speaker for clues to word reference. Journal of Child Language 20(2), 395-418. (DOI 10.1017/S0305000900008345)
53. Bloom, P. 2000 How children learn the meanings of words. Cambridge, MA: MIT Press.
54. Gleitman, L. 1990 The structural sources of verb meanings. Language Acquisition 1(1), 3-55. (DOI 10.1207/s15327817la0101_2)
55. Imai, M., Gentner, D. & Uchida, N. 1994 Children's theories of word meaning: The role of shape similarity in early acquisition. Cognitive Development 9(1), 45-75. (DOI 10.1016/0885-2014(94)90019-1)
56. Smith, L. B. 2000 Learning how to learn words: An associative crane. In Breaking the word learning barrier: What does it take? (eds. R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, A. Woodward, N. Akhtar, M. Tomasello & G. Hollich), pp. 51-58. New York: Oxford University Press.

57. Landau, B., Smith, L. B. & Jones, S. S. 1988 The importance of shape in early lexical learning. Cognitive Development 3(3), 299-321. (DOI 10.1016/0885-2014(88)90014-7)
58. Spiegel, C. & Halberda, J. 2011 Rapid fast-mapping abilities in 2-year-olds. Journal of Experimental Child Psychology 109(1), 132-140.
59. Fisher, C. 1996 Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. Cognitive psychology 31(1), 41-81.
60. Lidz, J., Gleitman, H. & Gleitman, L. 2003 Understanding how input matters: Verb learning and the footprint of universal grammar. Cognition 87(3), 151-178.
61. Naigles, L. 1990 Children use syntax to learn verb meanings. Journal of Child Language 17(2), 357-374.
62. Spector, F. & Maurer, D. 2009 Synesthesia: a new approach to understanding the development of perception. Developmental Psychology 45(1), 175-189.
63. Imai, M., Kita, S., Nagumo, M., & Okada, H. 2008 Sound symbolism facilitates early verb learning. Cognition 109(1), 54-65. (DOI 10.1016/j.cognition.2008.07.015)
64. Kita, S., Özyürek, A., Allen, S., & Ishizuka, T. 2010 Early links between iconic gestures and sound symbolic words: Evidence for multimodal protolanguage. In The Evolution of Language: Proceedings of the 8th International Conference, pp. 429-430. Singapore: World Scientific.
65. Peña, M., Mehler, J., & Nespor, M. 2011 The role of audiovisual processing in early conceptual development. Psychological Science 22(11), 1419-1421.
66. Ozturk, O., Krehm, M., & Vouloumanos, A. 2012 Sound symbolism in infancy: Evidence for sound-shape cross-modal correspondences in 4-month-olds. Journal of Experimental Child Psychology 114(2), 173–186.
67. Forte, M., Weiß, A., Martin, A., & Peperkamp, S. 2013 Looking for the bouba-kiki effect in prelexical infants. In The 12th International Conference on Auditory-Visual Speech Processing (eds. S. Ouni, F. Berthomier & A. Jesse), pp. 71-76. Le Chesnay Cedex, France: Inria.
68. Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. under review Sound Symbolism Scaffolds Language Development in Preverbal Infants.
69. Miyazaki, M., Hidaka, S., Imai, M., Yeung, H. H., Kantartzis, K., Okada, H., & Kita, S. 2013 The facilitatory role of sound symbolism in infant word learning. In Proceedings of the 35th Annual Conference of the Cognitive Science Society

(eds. M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth), pp. 3080-3085. Austin, Texas: Cognitive Science Society.

70. Davis, R. 1961 The fitness of names to drawings: A cross-cultural study in Tanganyika. British Journal of Psychology 52, 259-268.
71. Iwasaki, N., Vinson, D. P., & Vigliocco, G. 2007 How does it hurt, "kiri-kiri" or "siku-siku"? Japanese mimetic words of pain perceived by Japanese speakers and English speakers. In Applying theory and research to learning Japanese as a foreign language (ed. M. Minami), pp. 2-19. New Castle upon Tyne: Cambridge Scholars Publishing.
72. Iwasaki, N., Vinson, D. P., & Vigliocco, G. 2007 What do English speakers know about gera-gera and yota-yota?: A cross-lnguistic investigation of mimetic words for laughing and walking. Sekai no nihongo kyoiku [Japanese Language Education Around the Globe] 17, 53-78.
73. Brackbill, Y., & Little, K. 1957 Factor determining the guessing of meanings of foreign words. Journal of Abnormal and Social Psychology 54, 312-318.
74. Brown, R. W., Black, A. H., & Horowitz, A. E. 1955 Phonetic symbolism in natural languages. Journal of Abnormal and Social Psychology 50, 388-393.
75. Nygaard, L. C., Cook, A. E., & Namy, L. L. 2009 Sound to meaning correspondences facilitate word learning. Cognition 112(1), 181-186.
76. Saji, N., Akita, K., Imai, M., Kantartzis, K., & Kita, S. 2013 Cross-Linguistically Shared and Language-Specific Sound Symbolism for Motion: An Exploratory Data Mining Approach. In Proceedings of the 35th Annual Conference of the Cognitive Science Society (eds. M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth), pp. 1253-1259. Austin, Texas: Cognitive Science Society.
77. Tamori, I., & Schourup, L. 1999 Onomatope: Keitai-to imi [Onomatopoeia: Form and Meaning]. Tokyo: Kuroshio.
78. Shinohara, K. & Kawahara, S. 2012 A cross-linguistic study of sound symbolism: The images of size. In Proceedings of the 36th Annual Meeting of the Berkeley Linguistics Society.
79. Kantartzis, K., Imai, M., & Kita, S. under review Toddlers are sensitive to a wider range of sound symbolic links between form and meaning of words than adults.

80. Werker, J. F., & Tees, R. C. 1984 Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant behavior and development 7(1), 49-63.
81. Saffran, J. R., Aslin, R. N., & Newport, E. L. 1996 Statistical learning by 8-month-old infants. Science 274(5294), 1926-1928.
82. Maurer, D., & Mondloch, C. 2006 The infant as synaesthete. In Attention and Performance XXI: Processes of Change in Brain and Cognitive Development (eds. M. H. Johnson & Y. Munakata), pp.449-472. Oxford: Oxford University Press.
83. Lourenco, S. F., & Longo, M. R. 2010 General magnitude representation in human infants. Psychological Science 21(6), 873-881.
84. Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. 2010 Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. Psychological Science 21(1), 21-25.
85. Yeung, H. H., & Werker, J. F. 2009 Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. Cognition 113(2), 234-243.
86. Ludwig, V. U., Adachi, I., & Matsuzawa, T. 2011 Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. Proceedings of the National Academy of Sciences 108(51), 20661-20665.
87. Cytowic, R. E., & Eagleman, D. 2009 Wednesday is indigo blue: Discovering the brain of synesthesia. The MIT Press.
88. Marks, L. E., Hammeal, R. J., Bornstein, M. H., & Smith, L. B. 1987 Perceiving similarity and comprehending metaphor. Monographs of the Society for Research in Child Development 52(1), 1-100.
89. Imai,M., Asano,M., Miyazaki,M., Okada,H., Yeung,H., Kitajo,K., Thierry,G., & Kita, S. 2012 Sound symbolism helps infants' word learning. In The Evolution of Language: Proceedings of the 9th International Conference (eds. T. Scott-Phillips, M. Tamariz, E. A. Cartmill & J. R. Hurford), pp.456-457. Singapore: World Scientific.
90. Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., & Varela, F. J. 1999 Perception's shadow: long-distance synchronization of human brain activity. Nature 397(6718), 430-433.

91. Lachaux, J. P., Rodriguez, E., Le van Quyen, M., Lutz, A., Martinerie, J., & Varela, F. J. 2000 Studying single-trials of phase synchronous activity in the brain. International Journal of Bifurcation and Chaos 10(10), 2429-2439.
92. Engel, A. K., & Singer, W. 2001 Temporal binding and the neural correlates of sensory awareness. Trends in Cognitive Sciences 5(1), 16-25.
93. Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. 2001 The brainweb: phase synchronization and large-scale integration. Nature Reviews Neuroscience 2(4), 229-239.
94. Ward, L. M. 2003 Synchronous neural oscillations and cognitive processes. Trends in Cognitive Sciences 7(12), 553-559.
95. Fries, P. 2005 A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends in Cognitive Sciences 9(10), 474-480.
96. Kitajo, K., Doesburg, S. M., Yamanaka, K., Nozaki, D., Ward, L. M., & Yamamoto, Y. 2007 Noise-induced large-scale phase synchronization of human-brain activity associated with behavioural stochastic resonance. Europhysics Letters 80(40009), 1-6.
97. Kawasaki, M., Kitajo, K., & Yamaguchi, Y. 2010 Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory. European Journal of Neuroscience 31(9), 1683-1689.
98. Kutas, M., & Federmeier, K. D. 2011 Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annual review of psychology 62, 621-647.
99. Friedrich, M., & Friederici, A. D. 2005 Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. Neuroreport 16(6), 653-656.
100. Friedrich, M., & Friederici, A. D. 2011 Word learning in 6-month-olds: fast encoding–weak retention. Journal of Cognitive Neuroscience 23(11), 3228-3240.
101. Parise, E., & Csibra, G. 2012 Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. Psychological Science 23(7), 728-733.
102. von Stein, A., Rappelsberger, P., Sarnthein, J., & Petsche, H. 1999 Synchronization between temporal and parietal cortex during multimodal object processing in man. Cerebral Cortex 9(2), 137-150.

103. Imai, M., Haryu, E., & Okada, H. 2005 Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? Child Development 76, 340-355. (DOI 10.1111/j.1467-8624.2005.00849_a.x)
104. Casasola, M. & Cohen, L .B. 2000 Infants' association of linguistic labels with causal actions. Developmental Psychology 36(2), 155-168. (DOI 10.1037/0012-1649.36.2.155)
105. Werker, J.F., Cohen, L. B., Lloyd, V.L., Casasola, M., Stager, C. L. 1998 Acquisition of word-object associations by 14-month-old infants. Developmental Psychology 34(6), 1289-1309.
106. Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. & Shigematsu, J. 2008 Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. Child Development 79, 979-1000. (DOI 10.1111/j.1467-8624.2008.01171.x)
107. Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., Maguire, M. J., & Hennon, E. A. 2002 Young children can extend motion verb labels to point-light displays. Developmental Psychology 38, 604-614.
108. Tomasello, M. 2000 The item based nature of children's early syntactic development. Trends in Cognitive Sciences 4, 156-163. (DOI 10.1016/S1364-6613(00)01462-5)
109. Forbes, J. & Farrar, M. 1995 Learning to represent word meaning: What initial training events reveal about children's developing action verb concepts. Cognitive Development 10(1), 1-20.
110. Kersten, A. W. & Smith, L. B. 2002 Attention to Novel Objects during Verb Learning. Child Development 73(1), 93–109. (DOI 10.1111/1467-8624.00394)
111. Maguire, M. J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M., Slutzky, C. B., & Sootsman, J. 2002 Mapping words to actions and events: How do 18-month-olds learn a verb. In Proceedings of the 27th Annual Boston University Conference on Language (eds. B. Skarabela, S. Fish, & A. H. Do), pp. 371-382. Somerville, MA: Cascadilla Press.
112. Kantartzis, K., Imai, M., & Kita, S. 2011 Japanese Sound-Symbolism Facilitates Word Learning in English-Speaking Children. Cognitive Science 35(3), 575–586.

113. Yoshida, H. 2012 A cross-linguistic study of sound symbolism in children's verb learning. Journal of Cognition and Development 13(2), 232-265. (DOI 10.1080/15248372.2011.573515)
114. Fernald, A. & Morikawa, H. 1993 Common Themes and Cultural Variations in Japanese and American Mothers' Speech to Infants. Child Development 64(3), 637–656. (DOI 10.1111/j.1467-8624.1993.tb02933.x)
115. Saji, N. & Imai, M. 2013. Goishutoku ni okeru ruizousei no kouka no kentou [The Role of Iconicity in Lexical Development]. In Onomatope kenkyu no shatei –chikadzuku oto to imi [Sound Symbolism and Mimetics] (eds. K. Shinohara & R. Uno), pp. 151-166. Tokyo: Hituji Syobo.
116. Suzuki, Y. 2013 Intarakushon no naka de tsukawareru “onomatope + suru” doushi [The Mimetic + Light Verb (suru) Construction in the Father-child Interaction].In Onomatope kenkyu no shatei –chikadzuku oto to imi [Sound Symbolism and Mimetics] (eds. K. Shinohara & R. Uno), pp. 167-181. Tokyo: Hituji Syobo.
117. Butterworth, G. 2003 Pointing is the royal road to language for babies. In Pointing: Where language, culture, and cognition meet. (eds. Kita S), pp. 9-33. Mahwah, NJ: Erlbaum.
118. Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. 2006 12-and 18-month-olds point to provide information for others. Journal of Cognition and Development, 7(2), 173-187.
119. Goldin-Meadow, S. 2007 Pointing sets the stage for learning language--and creating language. Child Development, 78, 741-745.
120. Gogate, L. J., Bolzani, L.H., & Betancourt, E. A. 2006 Attention to maternal multimodal naming by 6- to 8-month-old infants and learning of word-object relations. Infancy, 9(3), 259-288.
121. Yu, C. & Smith, L. 2012 Embodied attention and word learning by toddlers. Cognition, 125, 244-262.
122. Monaghan, P., Christiansen, M.H., & Fitneva, S.A. 2011 The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. Journal of Experimental Psychology: General 140, 325-347. (DOI: 10.1037/a0022924)
123. Waxman, S. R. & Klibanoff, R. 2000 The role of comparison in the extension of novel adjectives. Developmental Psychology 36, 571-581.

124. Kita, S. 2008 World-view of protolanguage speakers as inferred from semantics of sound symbolic words: A case of Japanese mimetics. In Origins of language (eds. N. Masataka), pp. 25-38. Tokyo: Springer.
125. Kita, S., Kantartzis, K., & Imai, M. 2010 Children learn sound symbolic words better: Evolutionary vestige of sound symbolic protolanguage. In The evolution of language: Proceedings of the 8th International Conference (eds. A. D. M. Smith, M. Schouwstra, B. de Boer & K. Smith), pp. 206-213. Singapore: World Scientific.
126. Donald, M. 1997 Précis of Origins of the modern mind: three stages in the evolution of culture and cognition. Behavioral and Brain Sciences, 16, 737-791.
127. Ohala, J. J. 1994 The frequency code underlies the sound-symbolic use of voice pitch. In Sound symbolism (eds. L. Hinton, J. Nichols & J. J. Ohala), pp. 325-347. Cambridge: Cambridge University Press.
128. Wray, A. 2000 Holistic utterances in protolanguage: The link from primates to humans. In Evolutionary Emergence of Language : Social Function and the Origins of Linguistic Form (eds. C. Knight, M. Studdert-Kennedy & J. Hurford), pp. 285-302. West Nyack, NY: Cambridge University Press.
129. Senghas, A., Kita, S., & Özyürek, A. 2004 Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. Science 305, 1779-1782. (DOI 10.1126/science.1100199)
130. Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. 2012 The road to language learning is iconic: Evidence from British Sign Language. Psychological Science 23(12), 1443-1448.
131. Mumford, K. H., & Kita, S. 2014 Children Use Gesture to Interpret Novel Verb Meanings. Child Development XX.
132. Kita, S. 2001 Semantic schism and interpretive integration in Japanese sentences with a mimetic: A reply to Tsujimura. Linguistics 39, 419-436.
133. Dingemanse, M. 2011 The meaning and use of ideophones in Siwu. Doctoral dissertation, Radboud University Nijmegen, Nijmegen.
134. Dingemanse, M. 2013. Ideophones and gesture in everyday speech. Gesture*,* 13, 143-165.
135. Mihas, E. 2013. Composite ideophone-gesture utterances in the Ashéninka Perené ‘community of practice’, an Amazonian Arawak society from Central-Eastern Peru. Gesture*,* 13, 28-62.

136. Arbib, M. A. 2005 Interweaving protosign and protospeech: Further developments beyond the mirror. Interaction Studies 6(2), 145-171.
137. Corballis, M. C. 2002 From hand to mouth: The origins of language. Princeton University Press.
138. Hewes, G. W. 1973 An Explicit Formulation of the Relationship Between Tool-Using, Tool-Making, and the Emergence of Language. Visible Language 7(2), 102-127.

Figure Captions

Figure 1. Wood Rail (a), Great Tinamow (b). Drawn by Joseph Smit [23, 24], which are allied species in Berlin [22].

Figure 2. An example of manner of walking, used in Imai et al. [63], which sound symbolically matches the novel word, "nosunosu". In a pretest, Japanese and English speaking adults judged the novel word "nosunosu" to sound symbolically match this heavy and slow manner of walking.

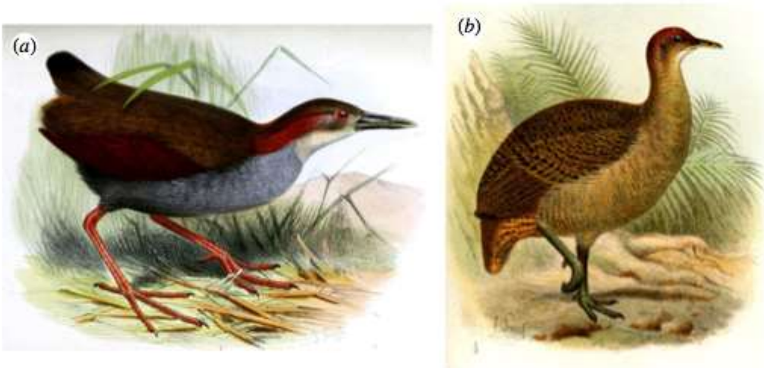

Figure 1. Wood Rail (*a*) and Great Tinamow (*b*). Drawn by Joseph Smit [22,23], which are allied species in Berlin [21]. (Online version in colour.)

Figure 2. An example of manner of walking, used by Imai *et al.* [63], which sound symbolically matches the novel word, 'nosunosu'. In a pretest, Japanese- and English-speaking adults judged the novel word 'nosunosu' to sound symbolically match this heavy and slow manner of walking. (Online version in colour.)