音声研究　第 22 巻第 2 号
2018（平成 30）年 8 月
122–130 頁

Journal of the Phonetic Society
of Japan, Vol. 22 No. 2
August 2018, pp. 122–130

特集論文

Acquisition of Sound Symbolic Values of Vowels and Voiced Obstruents by Japanese Children: Using a Pokémonastics Paradigm

Shigeto Kawahara*, Miwa Isobe**, Yukino Kobayashi***,
Tomoko Monou*, ****, ***** and Reiko Okabe******

有声性と母音の音象徴の獲得
—ポケモン名付けタスクを使った実験—

SUMMARY: Recent studies on sound symbolism have demonstrated that the number of voiced obstruents in Japanese Pokémon names is positively correlated with their evolution levels. This correlation is likely to be rooted in the sound symbolic relationship between voiced obstruents and largeness/heaviness/strength. This study shows that when Japanese children are presented with two non-existing names and a pair of pre-evolution and post-evolution Pokémon characters, they are more likely to associate names having voiced obstruents with post-evolution Pokémon characters. The experiment also shows that Japanese children associate post-evolution characters more with [a] than with [i], which shows that they are also sensitive to vocalic sound symbolism.

Key words: sound symbolism, voiced obstruents, the Frequency Code Hypothesis, vowels, language acquisition

1. Introduction

Modern thinking about languages generally assumes that the relationships between sounds and meanings are arbitrary. This dictum was made explicit by Saussure (1916), and was reiterated by Hockett (1959) as an important property of human languages, which has had a substantial influence on modern linguistic theories. On the other hand, there has also been the observation, which dates back to the time of Plato (the Dialogue *Cratylus*), that certain sounds can be systematically associated with certain meanings, even though these associations may be stochastic. For example, in the early 20th century, Jespersen (1922) and Sapir (1929) both argued that the vowel [i] invokes images of “smallness,” especially when compared to vowels like [a] and [o]. In Japanese, voiced obstruents are often associated with images of largeness, heaviness or dirtiness (e.g. Hamano 1986, Kawahara 2015, 2017, Kawahara et al. 2008, Kubozono 1999, Shinohara and Kawahara 2016, Suzuki 1962), an observation that is present in English as well (Newman 1933, Shinohara and Kawahara 2009, 2016). These systematic associations between sounds and meanings are referred to as “sound symbolism,” and are now actively studied in phonetics and psycholinguistics (see Dingemanse et al. 2015, Lockwood and Dingemanse 2015, Sidhu and Pexman 2017 for recent reviews).

Sound symbolism is an interesting topic for phonetic inquiry, because most if not all sound symbolic patterns make phonetic sense; i.e. the images of the sounds seem to derive from—or at least be compatible with—phonetic properties of these sounds (see Kawahara 2017 for recent extended exemplification). For example, [i] can be associated with image of being small, and this image may arise from the small aperture of the oral cavity for the articulation of [i] (Sapir 1929), or its high F2, which results from the small size of the oral cavity in front of the tongue (Jespersen 1922, Ohala 1984, 1994). As a general principle that dictates sound symbolic patterns in natural languages, Ohala

* Keio University（慶應義塾大学）
** Tokyo University of the Arts（東京藝術大学）
*** Tsukuba University of Technology（筑波技術大学）
**** Mejiro University（目白大学）
***** International Christian University（国際基督教大学）
****** Senshu University（専修大学）

(1984, 1994) proposed a very general hypothesis—now referred to as the Frequency Code Hypothesis—which contends that sounds with low frequency energy are associated with images of largeness, whereas sounds with high frequency energy are associated with images of smallness (see also Bauer 1987, Gussenhoven 2004, 2016). These associations reflect the physical law of vibration; everything else being equal, larger objects emit lower frequency sounds[1].

The Frequency Code Hypothesis explains why voiced obstruents are associated with images of largeness, as voiced obstruents are characterized by low frequency energy: during constriction, voiced obstruents often exhibit low frequency energy (a.k.a. "voice bar") as a result of vocal fold vibration (Stevens and Blumstein 1981). Furthermore, voiced obstruents also show lower f0 and F1 in the surrounding vowels, compared to voiceless obstruents and sonorants (see Hombert et al. 1979, House and Fairbanks 1953, Kingston and Diehl 1994, Lehiste and Peterson 1961, Lisker 1986, Stevens and Blumstein 1981 among many others). Kingston and his colleagues (Kingston and Diehl 1994, 1995, Kingston et al. 2008a, 2008b) have argued that the [+voice] feature should be characterized by a perceptually integrated property of low frequency energy, to which closure voicing, low f0 and low F1 contribute. In addition, a recent study by Chodroff and Wilson (2014) demonstrates that voiced stops characteristically have bursts with lower frequency energy compared to voiceless stops. The low frequency energy can be the basis of the images of largeness of voiced obstruents[2].

While there is an extensive body of literature on sound symbolism that has uncovered many patterns of sound-meaning relationships, one important question that remains within the research of sound symbolism is its acquisition. There has been some research on this issue. For example, Maurer et al. (2006) studied the so-called "bouba-kiki" effect (Ramachandran and Hubbard 2001) in which nonce words like *bouba* are more likely to be associated with a round object, whereas nonce words like *kiki* are more likely to be associated with an angular object. Maurer et al. (2006) demonstrated that 2.5-year-old children are sensitive to sound-shape associations, just like adult speakers. Imai et al. (2008) found that nonce verbs following sound symbolic principles are more easily learned by Japanese children than nonce verbs that do not. Kantartzis et al. (2011) showed that a similar result is obtained even if Japanese-sounding nonce words are used for English-speaking children. Through a looking-time experiment, Ozturk et al. (2013) demonstrated that 4-month old infants look at congruent sound-shape pairs longer than incongruent sound-shape pairs. Asano et al. (2015) demonstrate through an EEG experiment that 11-month-old infants may be sensitive to sound symbolic associations. Building on these results, Imai and Kita (2014) raised the possibility that sound symbolism may guide the language acquisition process to some nonnegligible extent (though see also Monaghan et al. 2012)[3]. Studying how children acquire sound symbolic patterns is thus an important topic in general linguistic inquiry. However, this topic is still understudied, despite these recent illuminating results.

To put the questions raised above in more concrete terms, do children have knowledge of sound symbolic associations just like adults? If so, are they able to use that knowledge to name new objects? This paper reports a case study which examined whether six-year-old Japanese children possess sound symbolic knowledge about vowels and voiced obstruents, and whether they are able to make use of that knowledge when they are asked to name new objects.

We made use of a research paradigm "Pokémonastics," initiated by Kawahara et al. (2018), later followed up on by various scholars (Kawahara and Kumagai 2019, Kumagai and Kawahara to appear, Shih et al. 2018, Suzuki 2017). This paradigm explores the sound symbolic nature of Pokémon characters. In Pokémon games, fictional creatures, who are called "Pokémon," evolve into related characters. When they undergo evolution, they are called by a different name (e.g. *himbasu* → *mirukarosu*); they also generally get bigger, larger, and stronger. It has been found that there is a positive correlation between the number of voiced obstruents in the existing Pokémons' names on the one hand and their evolution levels on the other (Kawahara et al. 2018). In other words, Japanese adults—or Pokémon designers, to be more specific—seem to associate voiced obstruents with images of largeness, heaviness, and strength, and use these sound symbolic associations when naming Pokémon characters. Later studies (Kawahara and Kumagai 2019, Kumagai and Kawahara to appear) show that these sound symbolic patterns are productive in that adult Japanese speakers, including those who are not familiar with Pokémon, reproduce these sound symbolic patterns in experimental settings, when they are asked to name new Pokémon characters. The current paper is a direct follow-up of these two experimental studies, but with a new focus on Japanese children.

2. Method

In the current study, we presented Japanese children with pairs of (non-existing, nonce) Pokémon characters, where one character was the pre-evolution version and the other one was the post-evolution version. The participants were also presented with two possible names and asked to choose which name was more appropriate for the pre-evolution version, and which name was more appropriate for the post-evolution version.

2.1 Stimuli

This experiment had five conditions, following Kumagai and Kawahara (to appear):

(1) Experimental conditions
- (a) [i] vs. [a]
- (b) [u] vs. [a]
- (c) The effects of voiced obstruents
- (d) Combination of (a) and (c)
- (e) Combination of (b) and (c)

The first condition tested whether [a] would be more likely to be associated with the post-evolution version of Pokémon characters than [i]. This would be the case if Japanese children possess sound symbolic knowledge in which [a] is bigger than [i] (Jespersen 1922, Sapir 1929, Shinohara and Kawahara 2016, Ultan 1978), and are able to apply that knowledge when naming new characters. Whether [u] is associated with image of size smaller than [a] is less clear from the previous studies of sound symbolism, although there have been observations that high vowels are generally considered to be smaller than low vowels (Newman 1933, Shinohara and Kawahara 2016, Yoshioka 2004)[4]. The third condition tested the sound symbolic values of voiced obstruents in Japanese children; as discussed in the introduction, the Frequency Code Hypothesis (Ohala 1984, 1994) predicts that, due to their low frequency energy, voiced obstruents should be associated with images of largeness and heaviness. Recall also that in the existing Pokémon names, there is a positive correlation between the number of voiced obstruents in their name and their evolution levels. The last two conditions (d) and (e) tested the combined effects of the vocalic sound symbolism and the consonantal sound symbolism.

The actual stimuli used in the experiment are provided in Table 1. Within each pair, two nonce words always have the same prosodic structure (mora-wise and syllable-wise). In conditions (a) and (b), the phonemic status of consonants was controlled, although affrication and palatalization in front of [i] and [u] were unavoidable, due to phonotactic constraints in Japanese phonology (Vance 2008). For condition (c), we used an online nonce word generator that combines Japanese syllables randomly[5]. We used this generator to avoid any bias in selecting the stimuli for experiments on sound symbolism (see Westbury 2005 for a cautionary remark about choosing stimuli for experiments on sound symbolism).

2.2 Task

Within each trial, the participants were visually presented with a pair of pre-evolution and post-evolution Pokémon characters. To make clear to the participants that post-evolution Pokémon characters are generally larger, they were about 1.5 times larger than the pre-

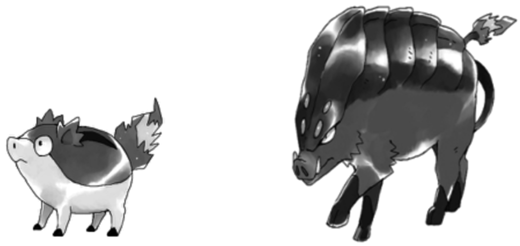

Figure 1 An example of visual stimuli. Left, pre-evolution character; right, post-evolution character.

Table 1 The stimuli (also used by Kumagai and Kawahara to appear).

(a) [i] vs. [a]	(b) [u] vs. [a]	(c) Voiced obstruents	(d) (a) + (c)	(e) (b) + (c)
[kiiki] vs. [kaaka]	[tsuutsu] vs. [taata]	[jasaha] vs. [gebiki]	[piipiɴ] vs. [baabaɴ]	[pumpuu] vs. [bambaa]
[ɕiiɕi] vs. [saasa]	[nuunu] vs. [naana]	[mesonu] vs. [dadera]	[kiikiɴ] vs. [gaagaɴ]	[sunsuu] vs. [zanzaa]
[miimi] vs. [maama]	[kuuku] vs. [kaaka]	[kejajo] vs. [zedotɕi]	[ɕiiɕiɴ] vs. [zaazaɴ]	[tsuntsuu] vs. [dandaa]
		[tsusoki] vs. [zoʑike]		
		[munere] vs. [zadoja]		
		[ɸureju] vs. [ʑiboru]		

evolution versions. An example pair of the visual stimuli is given in Figure 1. These visual stimuli were drawn by a digital artist, *toto-mame*, whose Pokémon pictures are judged to be very authentic by Pokémon practitioners[6)].

The participants were also provided with a pair of two nonce names in Japanese (those in Table 1). The names were read by a female experimenter, and the participants were asked to choose which name suited the pre-evolution character better and which name suited the post-evolution character better.

2.3 Participants

The participants were 24 native speakers of Japanese (ages 6;1–6;11, average 6;7), all from Tokyo or surrounding area. They participated in the experiment in March 2018 before starting elementary school. The participants were all female, since this experiment was conducted as a part of a larger project that required female participants only.

2.4 Procedure

The trials were created using SurveyMonkey. A female experimenter engaged with each participant through the experiment. The experimenter first asked each participant if she knew *Pikachu*, and all the participants responded affirmatively. Then the experimenter asked if they knew more about Pokémon, and if their response was negative, then the experimenter explained what Pokémon is, including the fact that when Pokémon characters evolve, they generally become bigger, larger, and stronger. Next, the experimenter showed participants a website with many Pokémon pictures, and asked how many Pokémon characters they could name. This response was coded using a 7-point scale where one end was “they can name almost all characters” and the other end was “they only know *Pikachu*.” The participants were also asked if they watched the Pokémon anime show *Sun & Moon*, which was broadcast on the television at the time of the experiment.

The participants completed all the trials with an experimenter. Although the two name choices were written in Japanese *katakana* orthography on SurveyMonkey, these orthographic prompts were not shown to the participants. Instead, the prompts were read aloud by the experimenter, so that the participants would base their judgments on auditory information rather than on orthography[7)]. The order of the two choices, as well as the order of the trials, were randomized per participant by SurveyMonkey.

3. Result and Discussion

Table 2 shows the ratios of expected responses; i.e., responses in which [a] is associated with the post-evolution characters and/or in which voiced obstruents are associated with the post-evolution characters. Since the responses were binary, 95% Confidence Intervals (CIs) were calculated based on binomial distributions. If these CIs do not overlap with 0.50, it implies that responses are skewed in such a way that they are higher than chance level.

We observe that in all conditions except for (b), the responses are above chance level. The result for condition (a) suggests that Japanese children consider [a] to be a better match for post-evolution characters than [i]. This result is likely to be rooted in the sound symbolic relation, identified by Sapir (1929), that [a] is perceived to be larger than [i].

The result for condition (b) did not significantly deviate from chance level—the choice between [a] and [u] was at chance level. It may be the case that, as Ohala (1984, 1994) hypothesizes, images of size are largely dictated by F2, at least for Japanese children; since Japanese [a] and [u] do not differ much in terms of F2 ([a]=1,383 Hz; [u]=1,419 Hz, according to the measurement reported in Keating and Huffman 1984), they did not differ in terms of images of size, hence the current results[8)]. With this said, a similar study by Kumagai and Kawahara (to appear) found that Japanese adults tend to associate names containing [a] more often with the post-evolution characters than with names containing [u]. Furthermore, previous studies on sound symbolism with adult participants also show that Japanese speakers find [a] to be larger than [u] (Shinohara and Kawahara 2016, Yoshioka 2004). It may be possible that children attend to acoustic properties only,

Table 2 Expected response ratios. Averages and 95% confidence intervals.

	(a)	(b)	(c)	(d)	(e)
average	0.653	0.467	0.847	0.707	0.787
95% CIs	0.598–0.708	0.409–0.524	0.817–0.876	0.654–0.759	0.739–0.834

whereas adults attend to articulatory properties as well so that they take degrees of oral aperture into account and judge [a] to be larger than [u]. This is admittedly a post-hoc speculative hypothesis, which needs to be tested against other sound symbolic patterns in Japanese and other languages. The simplest follow-up study may be to ask Japanese children to judge the largeness of [a] and [u] using a Lickert scale (see Shinohara and Kawahara 2016, Yoshioka 2004 for this sort of experiments using adult speakers).

The expected responses in condition (c) are highest among all the conditions tested. Recall that this condition consisted of pairs of three light syllables, one name containing no voiced obstruents, and the other name containing voiced obstruents in the first two syllables. The expected response ratio was 84.7%, which indicates that Japanese children very robustly associate voiced obstruents with post-evolution characters.

The results for conditions (d) and (e) are also higher than chance, although they are lower than condition (c). The fact that conditions (d) and (e)—which tested the combined sound symbolic effects of vowels and voiced obstruents—showed lower expected response ratios than condition (c) is somewhat surprising. At this point, we can only speculate; however, it may be the case that condition (c) involved two types of voiced obstruents (except for [dadera]), whereas (d) and (e) involved two occurrences of one voiced obstruent. In other words, repeating the same type of voiced obstruents twice may not have as much impact as having two different types of voiced obstruents[9]. This speculation is worth testing in future studies of sound symbolism.

One final question that needs to be addressed is whether the positive results obtained in this experiment came from exposure to the actual Pokémon names, rather than from abstract sound symbolic knowledge[10]. To examine this possibility, Figure 2 shows the correlation between familiarity with Pokémon (which was asked before the experiment) and the expected response ratios, together with a linear regression line and its 95% confidence intervals. If the patterns of sound symbolism observed in the current experiment derived from the existing Pokémon names, there should be a positive correlation between the two. We observe that there is absolutely no correlation between the two measures (r=0.01, *n.s.*); i.e. it is unlikely that the current results arise from the knowledge of actual Pokémon names.

Figure 3 compares the distribution of expected response ratios between the two groups: (1) those children who watch the Pokémon anime (left), and (2) those who do not (right). There are no substantial differences between the two groups (Wilcoxson test, *n.s.*)

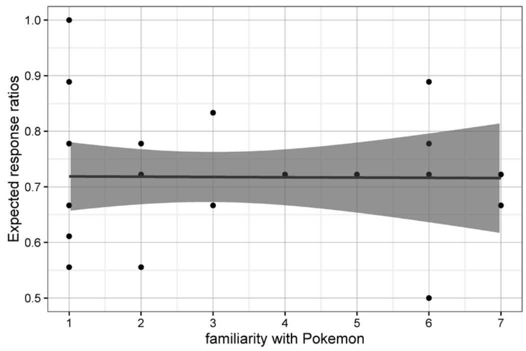

Figure 2 Correlation between familiarity and expected response ratios.

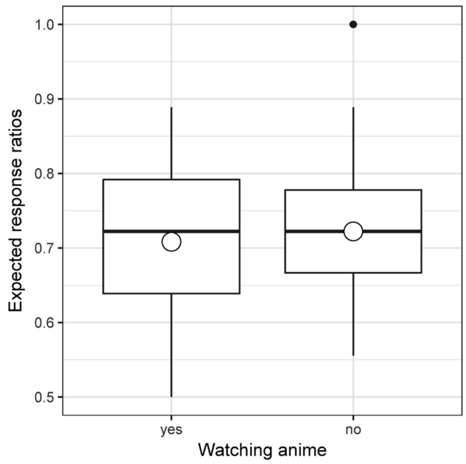

Figure 3 Comparison between those who watch Pokémon anime (*Sun & Moon*) and those who do not. White circles represent the means.

4. Conclusion

The contribution of the current paper may be modest; yet it is important in that we have shown that (i) Japanese children, before explicit school education on Japanese in elementary school, possess knowledge of certain sound symbolic patterns, and (ii) they can use that knowledge to name new objects (i.e. Pokémon characters). Since the study of acquisition of sound symbolism pattern is still limited in its empirical coverage, we believe that our results provide a non-trivial contribution. The difference between children and adults in the comparison between [a] and [u] is also intriguing—these results indicate that at least the current participants do not yet possess the sound symbolic knowledge that [a] is larger than [u] (or maybe they do but they fail to execute it), which Japanese adults demonstrably have. This observation raises the important question: what triggers the acquisition of sound symbolic patterns?

Another contribution of the current project is to have shown the lack of correlation between familiarity with Pokémon and observed effect sizes of the sound symbolic patterns, which implies that knowledge of sound symbolism is sufficiently abstract. We do not deny the possibility that sound symbolic knowledge can be learned from the lexicon. Indeed, the current participants may have learned the sound symbolic values of vowels and voiced obstruents from the Japanese lexicon, and this scenario is likely, given that not all sound symbolic patterns are universal (Blasi et al. 2016, Diffloth 1994, Kim 1977, Shih et al. 2018). Our conclusion is more nuanced—the participants could apply their sound symbolic knowledge in order to choose Pokémon's names, even if that knowledge itself may not come from exposure to Pokémon.

In general, the Pokémonastics paradigm offers an universe in which we can study sound symbolic patterns across languages. The Pokémon universe allows us to explore the universality and language-specificity of sound symbolic patterns, since in this universe, the set of denotations is fixed across different languages (Shih et al. 2018). In addition, this universe provides a forum to explore how sound symbolism may be influenced by phonological restrictions. As a concrete example, in the Japanese Pokémon names, the size and weight show positive correlation with mora counts (Kawahara et al. 2018), whereas in English, segment counts are a better predictor than mora or syllable counts (Shih et al. 2018). This observation is understandable from a phonotactic restriction in Japanese. Japanese speakers cannot simply add a consonant to express largeness or heaviness, because consonants do not generally stand alone in the language (Ito 1989)—Japanese speakers need to add a vowel as well. Therefore, English and Japanese differ in how to express largeness; the way in which they differ makes sense given the phonological characteristics of these two languages.

On this note, an anonymous reviewer pointed out one possibility for an interesting follow-up study. In Tohoku dialects of Japanese, many intervocalic voiceless obstruents become voiced; in some dialects at least, the difference between voiced and voiceless obstruents are neutralized (although in some dialects, underlying voiced obstruents are realized with pre-nasalization) (see e.g. Kanai 1982, Muraki 1970). It is predicted that speakers of these dialects would not show sound symbolic effects of intervocalic voiced obstruents. One complication is that since Pokémon is broadcast on TV using the Tokyo dialect, even children from the Tohoku area may interpret the Pokémon names using the phonology of the Tokyo dialect. To conduct this study, we would have to find pure speakers of Tohoku dialects, who have not watched the Pokémon TV series. Despite this challenge, studying a language/dialect where a voicing distinction is neutralized would be an interesting experiment.

We would like to end this paper with a final remark on experimental methodology. This Pokémonastic paradigm is fun to do for children (as well as adults—see Kawahara and Kumagai 2019). Our impression is that among all the experiments conducted at the same setting, the children enjoyed this experiment the most. Thus we hope that this Pokémonastic paradigm is used more widely in future acquisition studies.

Acknowledgements

This project is supported by the following JSPS grants: #17K02711, #17K13448, #26284059 and #26770198. We thank two anonymous reviewers for their helpful comments. We also received useful feedback from the audience at the 1st International Conference on Pokémonastics and the 156th meeting of the Linguistic Society of Japan. Remaining errors are ours.

Notes

1) The fundamental frequency (f_0) of an ideal string is:

$$f_0 = \frac{1}{2L}\sqrt{\frac{\sigma}{\rho}} \tag{1}$$

where L is length, σ is tension, and ρ is density. Crucially, f_0 and L are disproportional to each other.

2) An alternative, articulation-based, explanation of why voiced obstruents are associated with images of largeness is possible (Kawahara 2015, 2017). Due to aerodynamic conditions that need to be met for the production of voiced obstruents, speakers expand their oral cavity during the production of voiced obstruents (Ohala 1983, Ohala and Riordan 1979, Proctor et al. 2010). This expansion of the oral cavity may result in images of largeness.

3) See Nygaard et al. (2009) for the potential role of sound symbolism in second language learning.

4) In the corpus of existing Pokémon names, high vowels in initial syllables tend to be associated with lower evolution levels; however, the effect size is very small and not significant statistically (Kawahara et al. 2018). There are no substantial differences between the two high vowels either.

5) http://bit.ly/2iGaKko (last access, 08/31/2018)

6) These pictures were used with permission from the artist. Her website, where one can view other original Pokémon characters, can be found at https://t0t0mo.jimdo.com (last access, 8/31/2018)

7) In order to make the experiment more interactive and less boring for the 6-year-old participants, instead of using pre-recorded stimuli, the stimuli were read live to the children. Although the two stimuli within a pair were read in a naturalistic style, the amplitude of the stimuli was not perfectly controlled; we cannot entirely deny the possibility that in the experimenter's utterances, f0 and F1 differences due to voiced obstruents were unconsciously exaggerated. Although we suspect that the influence of these factors, if present at all, was minor, a follow-up experiment with pre-recorded stimuli can address this question in a more objective way.

8) This hypothesis can be further tested by comparing [o] and [e], since Japanese [o] has unambiguously lower F2 compared to [e] ([o]=1,136 Hz vs. [e]=1,720 Hz, Keating and Huffman 1984).

9) This speculation predicts that [dadera] should show lower expected responses than the other items that involve two voiced obstruents. This prediction is unfortunately not borne out—[dadera] was chosen as the post-evolution character 84% of the time, which is almost identical to the average of expected responses for the other five items. Another factor that may have contributed to the low expected response ratios for conditions (d) and (e) is that the stimuli involved (quasi-)reduplication. The participants thus may have perceived two voiced obstruents in the stimuli as two realizations of one form, which contributed to the judgment less substantially than two realizations of two distinct forms.

10) This question is reminiscent of a similar question in theoretical phonology: whether phonotactic knowledge is based on abstract phonological grammar or they can be learned from the statistics in the lexicon (see e.g. Berent et al. 2007 vs. Daland et al. 2011).

References

Asano, Michiko, Mutsumi Imai, Sotaro Kita, Keiichi Kitaji, Hiroyuki Okada and Guillaume Thierry (2015) "Sound symbolism scaffolds language development in preverbal infants." *Cortex* 63, 196–205.

Bauer, Harold R. (1987) "Frequency code: Orofacial correlates of fundamental frequency." *Phonetica* 44, 173–191.

Berent, Iris, Donca Steriade, Tracy Lennertz and Vered Vaknin (2007) "What we know about what we have never heard: Evidence from perceptual illusions." *Cognition* 104, 591–630.

Blasi, Damián, Søren Wichman, Harald Hammarström, Peter F. Stadler and Morten H. Christianson (2016) "Sound-meaning association biases evidenced across thousands of languages." *PNAS* 113(39), 10818–10823.

Chodroff, Eleanor and Colin Wilson (2014) "Burst spectrum as a cue for the stop voicing contrast in American English." *Journal of the Acoustical Society of America* 136(5), 2762–2772.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis and Ingrid Norrmann (2011) "Explaining sonority projection effects." *Phonology* 28(2), 197–234.

Diffloth, Gérard (1994) "i: *big*, a: *small*." In Leane Hinton, Johanna Nichols and John J. Ohala (eds.) *Sound symbolism*, 107–114. Cambridge: Cambridge University Press.

Dingemanse, Mark, Damián E. Blasi, Gary Lupyan, Morten H. Christiansen and Padraic Monaghan (2015) "Arbitrariness, iconicity and systematicity in language." *Trends in Cognitive Sciences* 19(10), 603–615.

Gussenhoven, Carlos (2004) *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos (2016) "Foundations of intonational meaning: Anatomical and physiological factors." *Topics in Cognitive Science* 8, 425–434.

Hamano, Shoko (1986) *The sound-symbolic system of Japanese.* Doctoral dissertation, University of Florida.

Hockett, Charles (1959) "Animal "languages" and human language." *Human Biology* 31, 32–39.

Hombert, Jean-Marie, John Ohala and William G. Ewan (1979) "Phonetic explanations for the development of tones." *Language* 55, 37–58.

House, Arthur and Grant Fairbanks (1953) "The influence of consonant environment upon the secondary acoustical characteristics of vowels." *Journal of the Acoustical Society of America* 25, 105–113.

Imai, Mutsumi and Sotaro Kita (2014) "The sound symbolism bootstrapping hypothesis for language acquisition and language evolution." *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1651), 20130298.

Imai, Mutsumi, Sotaro Kita, Miho Nagumo and Hiroyuki Okada (2008) "Sound symbolism facilitates early verb learning." *Cognition* 109, 54–65.

Ito, Junko (1989) "A prosodic theory of epenthesis." *Natural Language and Linguistic Theory* 7, 217–259.

Jespersen, Otto (1922) "Symbolic value of the vowel *i*." In Otto Jespersen (ed.) *Linguistica: Selected papers in English, French and German*, vol. 1, 283–303. Copenhagen: Levin and Munksgaard.

Kanai, Yoshimitsu (1982) "A case against the morphophonemic-allophonic principle." *Linguistic Inquiry* 13, 320–323.

Kantartzis, Katerina, Mutsumi Imai and Sotaro Kita (2011) "Japanese sound symbolism facilitates word learning in English-speaking children." *Cognitive Science* 35(3), 575–586.

Kawahara, Shigeto (2015) *Oto-to kotoba-no hushigi-na sekai*. Tokyo: Iwanami Shoten.

Kawahara, Shigeto (2017) *Introducing phonetics through sound symbolism*. Tokyo: Hituzi Syobo.

Kawahara, Shigeto and Gakuji Kumagai (2019) "Expressing evolution in Pokémon names: Experimental explorations." *Journal of Japanese Linguistics* .

Kawahara, Shigeto, Atsushi Noto and Gakuji Kumagai (2018) "Sound symbolic patterns in Pokémon names." *Phonetica* 75, 219–244.

Kawahara, Shigeto, Kazuko Shinohara and Yumi Uchimoto (2008) "A positional effect in sound symbolism: An experimental study." *Proceedings of the Japan Cognitive Linguistics Association* 8, 417–427.

Keating, Patricia A. and Marie Huffman (1984) "Vowel variation in Japanese." *Phonetica* 41, 191–207.

Kim, Kong-On (1977) "Sound symbolism in Korean." *Journal of Linguistics* 13, 67–75.

Kingston, John and Randy Diehl (1994) "Phonetic knowledge." *Language* 70, 419–454.

Kingston, John and Randy Diehl (1995) "Intermediate properties in the perception of distinctive feature values." In Bruce Connell and Amalia Arvaniti (eds.) *Papers in laboratory phonology IV: Phonology and phonetic evidence*, 7–27. Cambridge: Cambridge University Press.

Kingston, John, Randy Diehl, Cecelia Kirk and Wendy Castleman (2008a) "On the internal perceptual structure of distinctive features: The [voice] contrast." *Journal of Phonetics* 36, 28–54.

Kingston, John, Aditi Lahiri and Randy L. Diehl (2008b) *Voice.* Manuscript, University of Massachusetts, Amherst.

Kubozono, Haruo (1999) *Nihongo-no onsei: Gendai gengogaku nyuumon 2 [Japanese phonetics: An introduction to modern linguisitcs 2]*. Tokyo: Iwanami Shoten.

Kumagai, Gakuji and Shigeto Kawahara (to appear) "Pokémon-no nazuke-ni okeru boin-to yuuseisogaion-no kouka." *Gengo Kenkyu [Journal of the Linguistic Society of Japan]* 155.

Lehiste, I. and G.E. Peterson (1961) "Some basic considerations in the analysis of intonation." *Journal of the Acoustical Society of America* 33, 419–425.

Lisker, Leigh (1986) ""Voicing" in English: A catalog of acoustic features signaling /b/ versus /p/ in trochees." *Language and Speech* 29, 3–11.

Lockwood, Gwilym and Mark Dingemanse (2015) "Iconicity in the lab: A review of behavioral, developmental, and neuroimaging research into sound-symbolism." *Frontiers in Psychology*, doi: 10.3389/fpsyg.2015.01246.

Maurer, Daphne, Thanujeni Pathman and Catherine J. Mondloch (2006) "The shape of boubas: Sound-shape correspondences in toddlers and adults." *Developmental Science* 9, 316–322.

Monaghan, Padraic, Karen Mattok and Peter Walker (2012) "The role of sound symbolism in language learning." *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition* 38(5), 1152–1164.

Muraki, Masatake (1970) "A sound change in a Tohoku dialect." *Papers in Linguistics* 3, 341–347.

Newman, Stanley (1933) "Further experiments on phonetic symbolism." *American Journal of Psychology* 45, 53–75.

Nygaard, Lynne C., Alison E. Cook and Laura L. Namy (2009) "Sound to meaning correspondence facilitates word learning." *Cognition* 112, 181–186.

Ohala, John J. (1983) "The origin of sound patterns in vocal tract constraints." In Peter MacNeilage (ed.) *The production of speech*, 189–216. New York: Springer-Verlag.

Ohala, John J. (1984) "An ethological perspective on common cross-language utilization of f0 of voice." *Phonetica* 41, 1–16.

Ohala, John J. (1994) "The frequency code underlies the sound symbolic use of voice pitch." In Leane Hinton, Johanna Nichols and John J. Ohala (eds.) *Sound symbolism*, 325–347. Cambridge: Cambridge University Press.

Ohala, John J. and Carol J. Riordan (1979) "Passive vocal tract enlargement during voiced stops." In Jared J. Wolf and Dennis H. Klatt (eds.) *Speech communication papers*, 89–92. New York: Acoustical Society of America.

Ozturk, Ozge, Madelaine Krehm and Athena Vouloumanos (2013) "Shape symbolism in infacy: Evidence for sound-shapre cross-modal correspondences in 4-month-olds." *Journal of Experimental Child Psychology* 14(2), 173–186.

Proctor, Michael I., Christine H. Shadle and Khalil Iskarous (2010) "Pharyngeal articulation differences in voiced and voiceless fricatives." *Journal of the Acoustical Society of America* 127(3), 1507–1518.

Ramachandran, Vilayanur S and Edward M. Hubbard (2001) "Synesthesia–A window into perception, thought, and language." *Journal of Consciousness Studies* 8(12), 3–34.

Sapir, Edward (1929) "A study in phonetic symbolism." *Journal of Experimental Psychology* 12, 225–239.

Saussure, Ferdinand de (1916) *Cours de linguistique générale*. Paris: Payot.

Shih, Stephanie S., Jordan Ackerman, Noah Hermalin, Sharon Inkelas and Darya Kavitskaya (2018) "Pokémonikers: A study of sound symbolism and Pokémon names." *Proceedings of LSA*.

Shinohara, Kazuko and Shigeto Kawahara (2009) "A cross-linguistic study of sound symbolism: Images evoked by voicing." *Proceedings of the 26th Annual Meeting of Japanese Cognitive Science Society*.

Shinohara, Kazuko and Shigeto Kawahara (2016) "A cross-linguistic study of sound symbolism: The images of size." *Proceedings of the Thirty Sixth Annual Meeting of the Berkeley Linguistics Society*, 396–410.

Sidhu, David and Penny M. Pexman (2017) "Five mechanisms of sound symbolic association." *Psychonomic Bulletin & Review*, 1–25.

Stevens, Kenneth and Sheila Blumstein (1981) "The search for invariant acoustic correlates of phonetic features." In Peter Eimas and Joanne D. Miller (eds.) *Perspectives on the study of speech*, 1–38. New Jersey: Lawrence Erlbaum.

Suzuki, Michinori (2017) "The sound symbolic patterns in Pokémon move names." *Talk presented at Asia Junior Linguistics (AJL)*.

Suzuki, Takao (1962) "Oninkookan to igibunka no kankei ni tsuite–iwayuru seidakuon tairitsu-o chuushin toshite." *Gengo Kenkyu [Journal of the Linguistic Society of Japan]* 42, 23–30.

Ultan, Russell (1978) "Size-sound symbolism." In Joseph Greenberg (ed.) *Universals of human language II: Phonology*, 525–568. Stanford: Stanford University Press.

Vance, Timothy (2008) *The sounds of Japanese*. Cambridge: Cambridge University Press.

Westbury, Chris (2005) "Implicit sound symbolism in lexical access: Evidence from an interference task." *Brain and Language* 93, 10–19.

Yoshioka, Chisato (2004) "Nihongo-no boin ni okeru onshoochoo-no kenkyuu." *Gengobunka to Nihongo Kenkyuu*, 116–128.

(Received Apr. 10, 2018, Accepted Jun. 21, 2018)