The pluripotentiality of bilabial consonants: The images of softness and cuteness in Japanese and English

Abstract: The current study experimentally examined whether labial consonants were sound-symbolically associated with the images of softness and cuteness in Japanese and English. The results showed that all the bilabial consonants [p, b, m, φ, w] used in Japanese convey such images. In English, the consonants evoking the image of softness were bilabials but not labiodentals, and those linked to the image of cuteness were unaspirated, low-frequency bilabials. These results demonstrated the pluripotentiality of sound symbolism, meaning that a certain set of linguistic sounds evokes different meanings and images both within a single language and across languages (e.g. Winter et al. 2019; Kawahara and Kumagai to appear). Moreover, under the backcloth that the description of the glide /w/ concerning the place of articulation in Japanese and English is not uncontroversial, the current paper indicates – based on the current sound-symbolic experimental results – that the glide /w/ is phonologically labial in each language.

Keywords: sound symbolism, bilabial consonants, softness and cuteness

1 Introduction

1.1 Sound symbolism of bilabial consonants

In linguistics, it is assumed that an arbitrary relationship exists between sounds and meanings (de Saussure 1916). This enables people to refer to a certain item as one name in one language but to the same item as a different name in another language; for example, the animal that is referred to as cat in English is referred to as neko in Japanese and gato in Spanish. Natural languages also showcase precisely the opposite characterization – that a linguistic sound or feature evokes particular meanings and images, which is referred to as sound symbolism (Hinton et al. 1994/2006). A typical example of sound symbolism is the openness of vowels in articulation being sound-symbolically linked with the image of largeness (Newman 1933; Sapir 1929). Subsequent studies have confirmed that the sound-symbolic effect holds for numerous languages (e.g. Ultan 1978; Berlin 2006; Shinohara and Kawahara 2016). Sound symbolism has now become one of the featured research themes in linguistics, psychology, and cognitive science (for overview papers, see, e.g. Hinton et al. 1994/2006; Nuckolls 1999; Akita 2015; Lockwood and Dingemanse 2015; Sidhu and Pexman 2018; Kawahara 2020). These overview articles and edited books on sound symbolism show that a variety of research topics have been explored – such as basic vocabulary and interjections (e.g. Wichmann et al. 2010; Urban 2011; Blasi et al. 2016; Winter et al. 2019), human names (e.g. Cutler et al. 1990; Wright et al. 2005;
Pogacar et al. 2015; Bolts et al. 2016, fictional names of characters in TV animation and computer games (e.g., Pokémon character names by Kawahara et al. (2018) et seq.; Kumagai et al. (2020); Sidhu and Pexman (2019); Uno et al. (2020)), and brand names (e.g. Peterson and Ross 1972; Klink 2000; Yorkston and Menon 2004; Lowrey and Shrum 2007; Pogacar et al. 2015; Bolts et al. 2016).

A number of studies are available on the sound-symbolic effects of bilabial consonants, the sounds produced by touching the upper and lower lips together. In psychology, D’Onofrio (2014) used the matulama-take paradigm (Köhler 1929) to examine whether phonetic features affect the selection of shapes; the results suggested that labial consonants more likely match with the round shapes than with the spiky ones. Apart from such visual–auditory, sound-symbolic associations, recent studies have clarified that, even within a single language, bilabial consonants can be associated with multiple meanings and images such as babysness, cuteness, softness, innocence, and smallness. Kumagai and Kawahara (2017, 2020) suggest that Japanese speakers associate bilabial consonants [p, b, m, φ, w] with the image of babies.¹ Most baby diaper brands in Japan contain the bilabial consonants [p, m] (e.g. μunii, merizu, mamiipoko) (Kawahara 2017). This observation is interesting because bilabial consonants tend to appear in babbling in the earlier stages of language acquisition (Jakobson 1941/1968; MacNeilage et al. 1994). Kumagai and Kawahara (2017, 2020) conducted a forced-choice test that examined whether words that contain bilabial consonants (e.g. paraperu) are more appropriate for baby diaper brands than those that do not contain any bilabial consonants (e.g. tarateru). The study found that names that contain bilabial consonants are preferred as a baby diaper brand name. Kumagai and Kawahara (2017, 2020) also conducted a free elicitation task that asked native speakers of Japanese to come up with brand names for both baby diaper brands and adult cosmetics. The results showed that bilabial consonants were more likely to appear in baby diaper brand names than in cosmetics brand names.

Kumagai and Kawahara’s study has inspired researchers to analyse sound-symbolic effects of bilabial consonants; consequently, intriguing findings have been reported by several studies. Kumagai (2019) indicated a new type of alternation in which [h] becomes [p] in girls’ nicknames in Japanese (e.g. Haruka → Paruru, an ex member of AKB48, a young Japanese girl’s idol group). Since the new type occurs even in the word-initial position, it should be distinguished from the traditional [h] → [p] alternation observed word internally in Sino-Japanese words (e.g. kin “gold” + hatsu “hair” → kin-patsu “gold hair”). He also demonstrated via experiments that the bilabial [p] was the consonant that is most likely ([m] being the second most likely) to express cuteness.² Kawahara (2019) investigated the names of cute girls’ fighters who appeared in a Japanese TV animation named PreCure, in which normal girls transformed themselves into cute fighters called “PreCure” with new names that tended to possess bilabial consonants (e.g. Cure Peach, Cure Milky). Further, experiments using imaginary Pokémon characters demonstrated that, for Japanese speakers, names with bilabial consonants were more appropriate for fairy types, which look cuter than normal character types in Japanese (Kawahara and Kumagai 2019b) and that for English speakers, names with the bilabial [p] were also more appropriate for fairy-type characters (Kawahara et al. 2020a).³ Kumagai et al. (to appear) examined images of hardness and softness in the texture of snacks, thereby experimentally demonstrating that, among Japanese voiceless plosives [p, b, t, d, k, g], the bilabial [p] is more likely to evoke the image of softness than non-bilabial sounds like [k, g] (for similar results, see also Akita & McLean to appear). Hosokawa et al. (2018) and Uno et al. (2020) found that bilabial consonants were more likely to occur in the English names of non-villains that appear in Disney’s animated films (e.g. Pinocchio, Mickey) than in those of villains. Further, the number of bilabial consonants are negatively correlated with size in Japanese Pokémon names (Shih et al. 2018, 2019), and names that contain bilabial consonants are less

¹ Some scholars may point out that the glide [w] should be categorized into the bilabial consonant, as it is a labiovelar (IPA). See Section 1.2 for this issue.

² Apart from Japanese, there are a number of studies that explore sound-symbolic effects on the image of cuteness. For example, Jang (2018) reported that Korean exhibits tensification to express Aegyo, which is the way Korean speakers talk to pets and lovers.

³ Kawahara et al. (2018) and Kawahara & Kumagai (2019a) are pioneering sound-symbolic research papers on Pokémon names.
appropriate for post-evolution Pokémon characters than those that contain coronal consonants in experiments using imaginary Pokémon characters (Kawahara and Moore to appear). To summarize, what has been clarified in previous studies is that it is likely that [p] is associated with the image of cuteness in both Japanese and English and that bilabial consonants can evoke an image of innocence and smallness in English.

1.2 Labiality of [w]: Evidence from sound symbolism

As evident from the above discussion, although there has been an increased interest in sound-symbolic effects of bilabial consonants, a few studies utilize evidence from sound-symbolic studies to phonological argumentation (Kumagai and Kawahara 2017, 2020; Kawahara 2019; Kawahara and Kumagai 2019a, 2019b; see Kawahara 2020 for sound symbolism in phonology). In line with these studies, the current study addresses the issue of the place of articulation of glide /w/, based on sound-symbolic evidence. According to the International Phonetic Association, [w] is described as a labiovelar glide, pronounced with upper and lower lips touching together and with the back of the tongue moving upward to the velum. Since it involves the lips, it may not be a problem that [w] is included in the bilabial consonant group. However, there is no unanimous description of the glide concerning the place of articulation in introductory textbooks and papers on Japanese and English: the Japanese [w] is described as a labial (Shibatani 1990; Kubozono 2015), as a velar (Yamaguchi 2007; Tsujimura 2014; Pintér 2015), or as a labiovelar (Labrune 2012); the English glide /w/ is described as a labiovelar (or labial-velar) glide or approximant (Giegerich 1992; Hammond 1999; Cruttenden 2014; Yayas 2016; Collins et al. 2019; Carr 2020) or as a bilabial approximant (Roach 2009).

A number of scholars have addressed the issue of the glide /w/ concerning the place of articulation (e.g. Anderson 1976; Ohala and Lorentz 1977; Gick 2003; Nevins and Chitoran 2008), showing that the phonological behaviour of the glide /w/ varies across languages. Then is there evidence to determine the place-of-articulation feature of the glide /w/ in Japanese and English? Kumagai (2017) examined whether the Obligatory Contour Principle (OCP)-labial effect⁴ is generalized in a Japanese morphophonological phenomenon known as rendaku (see Vance 1987, 2015; Vance and Irwin 2016 for rendaku), suggesting that the glide /w/ is not specified with the place-of-articulation feature [labial] in phonology; this is because, unlike the other labial consonants /p, b, m, /, it does not participate in the OCP-labial effect. Meanwhile, it is observed that an English-speaking child replaces the cluster /sm, sw, sp/ with the labiodental [f] (e.g. smell = [f]ell; Smith = [f]ith; spoon = [f]oon; Tessier 2016: 176), which can be explained, provided that bilabials /m, w, p/ share the place-of-articulation feature [labial] with labiodental [f] and that the labiality is faithfully preserved between underlying and surface representations (i.e. input and output in optimality-theoretic terms; for Optimality Theory, see Prince and Smolensky 1993/2004; McCarthy 2002, 2008).⁷ In addition, as

---

4 Some native speakers of Japanese round their lips slightly when they produce it (Vance 1987).
5 For example, Karuk shows a /w/-/m/ alternation (e.g. /asíw/ → /ášim fák/ “to close one’s eyes;” cf. /ásíw/ “to sleep;” /st/ “to disappear;” Levi 2008), suggesting that the glide /w/ is specified with the place-of-articulation feature [labial]. Meanwhile, Luganda is an example in which the glide /w/ should be specified with [dorsal] (= velar); it shows a /w/-/g/ alternation (e.g. /wáŋa/ → /ggwáŋa “nation;” Clements 1986; Cole 1967; Kawahara 2007). For another example, there is a language called Fula showing both types of alternation (e.g. Nevins & Chitoran 2008; Paradis 1987).
6 OCP is an abbreviated form of Obligatory Contour Principle, which prevents identical elements from occurring in a certain domain (e.g. Bye 2011; Leben 1973; McCarthy 1986; Suzuki 1998; Yip 1988)
7 The labiality of the glide /w/ could be discussed in terms of English phonotactics. There are few English words with a sequence of homorganic consonants, such as /pw/, /bw/, /fw/, and /sw/ in the onset position (Giegerich 1992; Harris 1994). Similarly, the final consonant b in the sequence mb of the English words – for example, bomb, lamb, and lamb – is not pronounced. This might be accounted for by the rule that two sequential consonants with the place-of-articulation feature [labial] cannot occur in onset and coda clusters, but not all of the phonotactic restrictions can be explained by the homorganicity of sequential consonants. For example, there are also a few English words with a sequence of dw and gw (e.g. dwell,
1.3 Purposes of the current study

The first purpose of the current study is to experimentally examine whether labial consonants are sound-symbolically associated with the images of softness and cuteness in Japanese and English (yawarakai “soft” and kawaii “cute” in Japanese)\(^9\) for the following reasons. First, it remains unclear whether all labial consonants are associated with the image of softness in Japanese. The bilabial \(p\) was the most likely consonant to evoke the image of softness in Japanese, as lips are softer than other articulatory organs such as the alveolar ridge and the velum in the oral cavity (Kumagai et al. to appear). If the sound-symbolic association of bilabial \(p\) results from the sensation that lips are tied with an image of softness, it is predicted that all bilabial consonants are associated with the image of softness not only in a single language but also in others, such as English.\(^10\) Second, although bilabial \(p, m\) are likely to express cuteness in Japanese (Kumagai 2019), it is underexplored whether all bilabial consonants are associated with the image of cuteness in that language. Where does the image of cuteness originate from? One possibility is that it results from a (cute) pouting gesture that appears when humans articulate bilabial consonants. The pouting gesture is known as part of the photographic pose called a duck face. A study on duck faces for Chinese speakers (Qiu et al. 2015) showed that people making duck faces are more likely to be neurotic; however, on the other hand, online articles showed that duck faces appear to be sexually enticing,\(^11\) and that the Japanese duck face, called ahiru-guti “a duck mouth” or tyun-gao “face like a bird,” is also likely to be an appealing expression.\(^12,13\) If the pouting gesture evokes the image of cuteness, then it is not difficult to imagine that all bilabial consonants evoke such an image in Japanese, let alone other languages. Third, even if it is found that bilabial consonants are associated with the images of softness and cuteness in Japanese, it remains to be determined whether the sound-symbolic associations are conveyed by labial

dwarf, dwindle, Gwen, guano; Hammond 1999). It is impossible to account for the \(dw\) case in terms of the homorganic articulation of sequential consonants because \(d\) is [coronal], while \(w\) is not. It is also impossible to offer an account, in terms of the place of articulation, for why there are few words with the consonant cluster /gw/, while English words with a sequence of /kw/ are much more in number (e.g. quote, queen, quality, quantity, question), even though \(k\) and \(g\) belong to the same category [dorsal].

8 It should be noted that not all data available to researchers become evidence for the phonology of language speakers (see de Lacy 2009 for a discussion). In this sense, sound-symbolic associations do not always become firm evidence of phonological representations (see Kawahara 2020 for a related discussion), but it is of great importance to collect linguistic data related to the current issue and to continue the discussion.

9 As mentioned in Section 2, the current study asked participants to judge whether given names sound cute or soft, without providing any definition of these words soft/yawarakai and cute/kawaii in English and Japanese.

10 The words for “lips” are more likely to contain labial sounds in many languages (Urban 2011), and, of course, there is a labial sound in the Japanese word kutiburu for “lips.”


13 There is a book written in Japanese that discusses why the Japanese duck face is appealing (Nomura 2010).
consonants, including labiodental consonants and bilabial consonants, because of the absence of any non-bilabial consonants such as labiodentals in the Japanese language. Relatedly, although the p-initial names were more appropriate for the fairy-type (cute) Pokémon characters (Kawahara et al. 2020a), it remains to be seen whether the sound-symbolic association involves aspiration, as the study conducted an experiment using only nonce names that begin with p, but not those that begin with sp (when the following vowel is stressed, the p-initial is usually aspirated but the sp-initial ([s] + consonant clusters) is unaspirated). Thus, in order to resolve these issues, the current study examines whether the English labiodental fricative [f] has the same sound-symbolic effect as other bilabial consonants and whether aspirated and unaspirated bilabial consonants sound soft and cute to English speakers, respectively.

The second purpose of the current study is, based on the current sound-symbolic experimental results for Japanese and English, the discussion on whether the place-of-articulation feature [labial] is specified for the glide /w/ in each language. If the experiment shows that the glide [w] – in addition to bilabial consonants such as [p, b, m] – evokes the images of softness and cuteness in Japanese and English, then it will be concluded that it shares the feature [labial] with other bilabial consonants. Meanwhile, if the glide does not show the same tendency as other bilabial consonants, then it is unlikely to possess the feature [labial]. The current article suggests that the glide /w/ is phonologically labial in Japanese and English, although the English results were less clear.

2 Experiment

2.1 Task and stimuli

The task of the current experiment was that, given nonce words, participants were asked which of them sounded cuter or softer (kawaii and yawarakai in Japanese, respectively).¹⁴ The targeted languages were Japanese and English, each of which has different labial consonants and phonotactic restrictions, and thus different stimuli were prepared so as not to deviate from the phonotactic rules of each language. There were five to six conditions, each of which had three pairs of nonce words with and without labial consonants, as shown in Tables 1 and 2. The labial consonants in the experimental group were replaced with the alveolar counterparts.

The Japanese language has five labial consonants: the voiceless bilabial plosive [p], the voiced bilabial plosive [b], the bilabial nasal [m], the voiceless bilabial fricative [ɸ], and the labiovelar glide (approximant) [w] (e.g. Labrune 2012). All these consonants can occur in the onset position. The first three [p, b, m] can precede any vowel [i, e, a, o, u]. There are a number of phonotactic restrictions on [ɸ] and [w]. Although the [ɸ] can precede any vowel, and the [w] can precede [i, e, a, u] in loanwords (e.g. Pintér 2015), [ɸ] is known as an allophone of /h/ before [u], and [w] is followed by [a] in native Japanese words (e.g. Tsujimura 2014). Thus, in the stimuli, the vowels after [ɸ] and [w] were restricted to [u, a], respectively. All the stimuli consisted of four morae, and the targeted consonants were placed at the onset position in the first and third morae.

The English language has five labial consonants: the voiceless bilabial plosive [p], the voiced bilabial plosive [b], the bilabial nasal [m], the voiceless labiodental fricative [f], and the labiovelar glide (approximant) [w]. The only difference between Japanese and English is that while the Japanese fricative is a bilabial [ɸ], the English counterpart is a labiodental [f]. The current experiment did not target the voiced labiodental fricative [v] because, as the Japanese language had no voiced fricative, such as [β], it was difficult to compare the results of both languages. It is well known that voiceless plosives exhibit different

¹⁴ In the current experiment, no definition of cuteness (kawaii) and softness (yawarakai) was provided to participants of each language, which may have affected the participants’ responses (see Section 3.2 for a discussion).
realizations according to phonetic environments: /p/ is phonetically realized as an aspirated consonant [pʰ] when it is placed in the onset position of a syllable and when the syllable to which it belongs is assigned stress. However, it is not produced with aspiration in [s] + consonant clusters (e.g. /sp/). The current experiment distinguished these two different environments and set up a total of six conditions [pʰ, p, b, m, f, w]. All the stimuli consisted of two syllables, and the targeted consonants were syllable-initial in each of the stimuli.

All the pairs were randomized per participant and per session. Following the previous experimental studies on bilabial consonants, as detailed in Section 1.1 (Kumagai and Kawahara 2017, 2020; Kawahara and Kumagai 2019a, 2019b; Kumagai 2019; Kawahara and Moore to appear; Kumagai et al. to appear), the current study adopted experiments using non-auditory stimuli; the stimuli were provided in katakana orthography for Japanese speakers and in the Latin alphabet for English speakers.

<table>
<thead>
<tr>
<th>Table 1: Pairs of stimuli (Japanese)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Cond 1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 4</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 5</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Pairs of stimuli (English)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Cond 1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 4</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 5</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cond 6</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
2.2 Procedure

The present study used an online questionnaire through SurveyMonkey. The experiment consisted of two sessions: in the first session, participants were first provided an explanation that they would be presented with two names per question and that their task was to choose which they felt was a cuter name (kawaii namae “cute name” in Japanese); thereafter, they were asked to judge all of the pairs of the stimuli. As soon as the first session ended, the second session began, where the participants were provided an explanation that they would be presented with two names per question and that their task was to choose which name they felt had an image of “soft” (yawarakai imeeji “soft image” in Japanese); thereafter, they were asked to judge all the pairs of the stimuli. After the second session, the participants were asked several questions regarding their age, gender, and whether they were native speakers of Japanese/English. Moreover, the participants in the Japanese experiment were asked whether they had studied sound symbolism, and participants in the English experiment were asked whether they were familiar with the term sound symbolism.¹

2.3 Participants

For the Japanese participants, the author of the current article advertised the experiment in university classes, and a total of 83 native-speaking students participated. They did not belong to the university’s linguistics department or related fields, and they reported having never studied sound symbolism. One of the participants did not complete the task, and, thus, data for the remaining 82 were analysed. All the participants were female and their age range was 18–30 years.

For the English participants, the weblink for the experiment was distributed through SurveyMonkey, and a total of 106 native speakers were recruited. The survey asked whether they were familiar with the term sound symbolism and 60 participants answered in the affirmative; thus, they were excluded from the analysis, and the remaining 46 speakers were targeted. There was a total of 36 female speakers; 17 of the speakers were aged under 30 years.

2.4 Results

The current section presents the results of each experiment. The figures display the rate at which the nonce words containing labial consonants were chosen. Error bars represent 95% confident intervals.

2.4.1 Softness in Japanese

Figure 1 presents the rate at which Japanese speakers judged the stimuli with bilabial consonants to be soft: [p] = 0.614, [b] = 0.711, [m] = 0.695, [f] (represented as f) = 0.911, [w] = 0.797. All the bilabial consonants were above chance level. Following the previous studies that conducted a forced-choice task (e.g. Kumagai and Kawahara 2017, 2020), the author made a generalized linear mixed-effects logistic regression (glmer) analysis, with participants and items being the random effects. Consequently, the five consonants were significantly judged as soft names (β = 2.471, SE = 0.323, z = 7.646, p < 0.001). In order to compare the number of observed outcomes to that of expected outcomes, the author also ran binomial tests; the results

¹ For the Japanese experiment, the author of the current article advertised the experiment in his classroom, where he had already talked about what sound symbolism was. Thus, he asked a different question for Japanese speakers.
revealed that all the five consonants showed the significance of the observed outcomes ($|p| = 155/246 = 82$ participants $\times$ 3 pairs), $p < 0.001$; $|b| = 175/246$, $p < 0.001$; $|m| = 171/246$, $p < 0.001$; $|\phi| = 224/246$, $p < 0.001$; $|w| = 196/246$, $p < 0.001$).

2.4.2 Cuteness in Japanese

Figure 2 depicts the results for the image of cuteness in Japanese. The rates of each response were $|p| = 0.724$, $|b| = 0.707$, $|m| = 0.642$, $|\phi|$ (represented as $f$) = 0.776, and $|w| = 0.756$. Similar to the results of softness, all the bilabial consonants were above chance level. A glmer analysis showed that these consonants were significantly judged as cute names ($\beta = 2.036$, SE = 0.228, $z = 8.917$, $p < 0.001$). Moreover, a binomial test showed significance in the observed outcomes of all consonants ($|p| = 178/246$, $p < 0.001$; $|b| = 174/246$, $p < 0.001$; $|m| = 158/246$, $p < 0.001$; $|\phi| = 191/246$, $p < 0.001$; $|w| = 186/246$, $p < 0.001$).

2.4.3 Softness in English

Figure 3 depicts the extent to which English speakers judged the stimuli containing labial consonants as a soft name. The following rates are presented in each graph: $|p^h|$ (represented as $ph$) = 0.754, $|p|$ (represented as $sp$) = 0.746, $|b| = 0.775$, $|f| = 0.384$, $|m| = 0.703$, $|w| = 0.601$. All the bar graphs except the fourth
one were above chance level. A binomial test indicated the significance of the expected results for \([p^b, p, b, m, w]\) \((p^b) = 104/138 \ (= 46 \text{ participants} \times 3 \text{ pairs}), p < 0.001; [p] = 103/138, p < 0.001; [b] = 107/138, p < 0.001; [m] = 97/138, p < 0.001; [w] = 83/138, p < 0.05) but showed significance of the unexpected result for \([f]\) \((f) = 53/138, p < 0.01). This implies that it is evident that the labiodental \([f]\) cannot evoke the image of softness. Then the author conducted a \textit{glmer} analysis that excluded the data of responses to the labiodental \([f]\); the results revealed that the names that contain bilabial consonants \([p^b, p, b, m, w]\) were judged to be soft names \((\beta = 1.188, SE = 0.151, z = 7.844, p < 0.001)\).

### 2.4.4 Cuteness in English

Figure 4 depicts the responses by English speakers in which the stimuli with labial consonants sounded cuter than those without them: \([p^h]\) (represented as \(ph\)) = 0.523, \([p]\) (represented as \(sp\)) = 0.616, \([b]\) = 0.775, \([f]\) = 0.196, \([m]\) = 0.659, and \([w]\) = 0.37. The expected responses were obtained in the second, third, and fifth graphs (i.e. \([p, b, m]\)) but not in the others (i.e. \([p^h, f, w]\))). A binomial test yielded significant results for \([p, b, m]\) \((p) = 85/138, p < 0.01\ (= 46 \text{ participants} \times 3 \text{ pairs}); [b] = 107/138, p < 0.001; [m] = 91/138, p < 0.001) but showed no significant result for \([p^h]\) \((p^h) = 72/138, p = 0.671) and significance of the unexpected result for \([f, w]\) \((f) = 27/138, p < 0.001; [w] = 51/138 p < 0.01). Then a \textit{glmer} analysis with the three consonants \([p^h, f, w]\) excluded indicated that the responses to three consonants \([p, b, m]\) were significant \((\beta = 1.553, SE = 0.173, z = 8.952, p < 0.001)\).
3 Discussion

3.1 The image of softness

The results of the current experiment showed that the names that contain bilabial consonants were judged to be cuter names (kawaii namae in Japanese) than those that do not contain any bilabial consonants. This suggests that all the bilabial consonants used in Japanese ([p, b, m, φ, w]) were sound-symbolically associated with the image of softness. There are two possible explanations for these results. The first one is the physiological idea that softness results from the texture of the lips (Kumagai et al. to appear). The second explanation involves acoustic motivation. Ohala (1984, 1994) proposed the Frequency Code Hypothesis in which “low-frequency consonants are associated with low-frequency sounds, large size, softness, and heavy, slow movements” (Hinton Nichols and Ohala 1994/2006: 10). If bilabial consonants are assumed to have low-frequency energy (Berlin 2006), it is possible that Japanese speakers associate them with the image of softness.

However, the situation is different in English. The results showed that the consonants judged to be soft were [p, b, m, w] but not [f]. The labiodental [f] does not evoke the image of softness, as it involves only the lower lip and not both lips, which may not satisfy the condition that lips are softer than other articulatory organs in the oral cavity (Kumagai et al. to appear). Moreover, since both [pʰ] and [p] exhibited significant results for the image of softness, aspiration was not significant in the relationship between bilabial consonants and the image of softness, at least, in English.

The current results pertaining to the labiodental [f] require further discussion. The binomial test yielded the unexpected result that the response to [s] in the control group was significant, which implies that the alveolar [s] was more likely to be judged to be soft than the labiodental [f]. However, only from these results, it may be difficult to conclude that the alveolar [s] can evoke the image of softness. The reason for these results requires analysis with new experiments in further research.

3.2 The image of cuteness

The current results indicated that, as was the case with the image of softness, all the bilabial consonants were associated with the image of cuteness in Japanese, which also replicated Kumagai’s (2019) results of the image of cuteness. As mentioned in Section 1.3, one possibility for the origin of the image of cuteness is that it resulted from a pouting gesture produced by articulating bilabial consonants. Another possibility is that a sound-symbolic association resulted from another in the minds of language speakers. Previous studies have already shown that a meaning or image is mutually linked to another (see also French 1977; Gallace et al. 2011; Kawahara and Kumagai to appear). Thus, it is possible that the image of cuteness may be mutually associated with that of softness in Japanese. The sound-symbolic association between bilabials and the image of softness could have produced the sound-symbolic association between bilabials and the image of cuteness. The opposite case is also possible: the sound-symbolic association of cuteness may have produced that of softness.

The results for the image of cuteness in English were more complex: bilabial [p, b, m] were judged to be cute, while [pʰ, f, w] were not. Unlike the case of Japanese, the account only with the pouting gesture is insufficient. Then are there any additional possible reasons for these results? First, the English word cute is not semantically equal to the Japanese word kawaii, which may be responsible for the less clear results in English. The English word cute means not only “attractive” but also “clever/cunning,” which may have been reflected in the current results (see also Section 3.3). For example, although voiced obstruents evoke a negative image due to their aerodynamic challenge (Uno et al. 2020), if low-frequency consonants [b, m], including voiced obstruents, can evoke the image of villainess, then it is possible that the names that contain the low-frequency consonants [b, m] will be more likely judged to be cuter names (including the
meaning of cleverness/cunningness). This may account for why [b, m] is higher than \( ph \) and \( sp \) [\( p^h, p \)] in terms of the rate presented in Figure 4.

Second, the names with [\( p^h \)] were not significantly judged to be cute, which was not consistent with the results of Kawahara et al. (2020a) that the \( p \)-initial name was appropriate for the (cute) fairy type of Pokémon characters. Where did the discrepancy come from? It may be difficult to draw a conclusion at the moment because the experimental designs of each study were not completely equal, but, as the \( p \)-initial name usually involves aspiration, it is necessary to conduct an experiment with auditory stimuli, as there is a possibility that aspiration (with strong expelled breath) can evoke the image of strength, which may not have matched with the image of cuteness.

Third, the results for \([w]\) are rather puzzling, because it belongs to low-frequency consonants (Berlin 2006), which would show the same tendency as other bilabial consonants [\( b, m \)]. Moreover, the results for \([w]\) in the image of cuteness are contradictory to those in the image of softness in English. Rather, it was unpredicted that the response to [\( j ] < y > \) was significant, which implies that the alveolar [\( j \)] could be associated with the image of cuteness. This needs to be examined with new experiments, since it has the aforementioned problem of the ambiguous meaning of cuteness.

Fourth, apart from phonetic features, the profile of Japanese and English speakers was somewhat different: all the Japanese participants (\( N = 82 \)) were female or aged under 30 years, while there were only 14 profiles that overlapped in terms of age and gender among the English speakers, which might be responsible for the different results in sound symbolic effects. The author assumed that such a difference did not affect participants’ responses, as differences in gender and age in sound symbolism are understudied (cf. Klink 2009; Kawahara et al. 2020b). However, a study in psychology showed that females were more sensitive to cute infants (Lobmaier et al. 2010); thus, there is no denying that gender differences influence responses to the notion of cuteness. Thus, how differences in gender and age arise in the judgement of softness and cuteness would be an interesting avenue for further research.

Finally, the response to the labiodental [\( f \)] was lower than a chance level. If the sound-symbolic association of cuteness resulted from the pouting gesture with the (upper and lower) lips in articulating bilabial consonants, it is convincing that the labiodental [\( f \)] does not evoke such an image, as it involves only the lower lip and not both lips. On the other hand, like the image of softness, the response to the alveolar [\( s \)] was significant for the image of cuteness. However, it is unclear whether the alveolar [\( s \)] itself evokes the images of softness and cuteness. Thus, it seems best to refrain from drawing a conclusion until future experiments are conducted. To sum up, it is safe to conclude that the consonants to express cuteness in English are unaspirated, low-frequency bilabials.

### 3.3 The place of articulation of the glide /w/

Based on the experimental results, this section discusses whether the glide /\( w \)/ is specified with the place-of-articulation feature [labial] in Japanese and English. For the results with Japanese speakers, it is not only the bilabial consonants [\( p, b, m, f \)] but also [\( w \)] that is more likely to evoke the images of softness and cuteness, which suggests that the glide /\( w \)/ shares the feature [labial] with the bilabial consonants. Moreover, precisely speaking, the feature [labial] is associated with such images in Japanese. This accords with the claim made by Kumagai and Kawahara (2017, 2020) that the glide /\( w \)/ belongs to the [labial] category.

For the results with English speakers, the glide [\( w \)] and the bilabial consonants were more likely to evoke the image of softness. This suggests that it is specified with the feature [labial] being associated with

---

16 Voiced plosives have low-frequency energy that is referred to as voice bar (Reetz & Jongman 2009) and a lower adjacent \( F_0 \) contour (e.g. Kingston & Diehl 1994), and labial nasals show a lower frequency range in the anti-formant than did the alveolar and velar nasals (Reetz & Jongman 2009).
the image of softness in English. Meanwhile, the results for the image of cuteness were less clear. A possible reason for these results is that the glide /w/ shows labiality in one case but does not in another case within a single language; in fact, as mentioned in Section 1.2, both types of evidence have been reported thus far on the non-labiality of the Japanese glide /w/. Considering this peculiar behaviour, it may be unsurprising that the current results for English showed both labiality and non-labiality; however, even if this holds true, how the place of the articulation of the glide /w/ is analysed in phonological theories would be a serious challenge. To summarize, based on the current results for the image of softness, it is possible that the glide /w/ is phonologically labial in Japanese and English.

4 Concluding remarks

The experiments in this study showed that Japanese speakers associate all the bilabial consonants [p, b, m, ɸ, w] with the images of softness and cuteness. In addition, the images of softness and cuteness of bilabial consonants could be mutually linked in the minds of Japanese speakers. In contrast, the results for English were less clear; English speakers associate the bilabial consonants [pʰ, p, b, m, w] with the image of softness, while they associated some of the bilabial consonants (i.e. unaspirated, low-frequency bilabials) [p, b, m] with the image of cuteness. Overall, the current study demonstrated the pluripotentiality of sound symbolism, meaning that a set of linguistic sounds conveys multiple meanings both within a single language and across languages (Winter et al. 2019; Kawahara and Kumagai to appear) and also suggested that the glide /w/ is phonologically labial in Japanese and English.

A number of issues are remaining, which could possibly be addressed in the future research. First, the notion of a word in a language is not always the same as that in another. In fact, it seems that it is difficult to find a word in foreign languages that corresponds with the Japanese word  kawaii  (Buckley 2016). In the case of the current study, the meaning of cuteness in English is different from the notion of  kawaii  in Japanese, which may have caused the difference in results. Thus, it is necessary to examine the image of cuteness/  kawaii  by providing a clear definition of a particular word, based on experimental reports in psychology, such as Nittono (2016). Second, the current study assumed that the image of softness resulted from the physiological reason that lips are softer than other articulation-related organs such as the alveolar ridge and the velum in the oral cavity (Kumagai et al. to appear) and that the image of cuteness resulted from the pouting gesture that appears in articulating bilabial consonants. In consideration of these aspects, it is expected that such images are likely found across languages. Other languages should be tested in the future. Third, the current study suggests the possibility that the image evoked by a particular set of sounds arises from another image by the same set of sounds. Thus, since there is a possibility that new semantic associations can be discovered, it is worth examining the pluripotentiality of bilabial consonants across natural languages.

Acknowledgements: I would like to thank Shigeto Kawahara, Kazuko Shinohara, Ryoko Uno, as well as two anonymous reviewers, for useful discussion and helpful comments. My appreciation also goes to Keiko Nakamura for her advice on the design of the experiment for English speakers. The current study is supported by a Grant-in-Aid for Early-Career Scientists from JSPS (#19K13164).

References


