Sound symbolism can count three segments (whereas phonology presumably cannot)

Abstract

Some researchers have recently argued that sound symbolic requirements can cause phonological alternations, and others have claimed that the relationships between sounds and meanings can be likened to input-output relationships in phonology, both suggesting that sound symbolic patterns and phonological patterns may be governed by similar mechanisms. Against this theoretical development, this paper further addresses the question regarding how similar phonological systems and sound symbolic systems are, by focusing on their counting capacity. It has been known that phonological systems can count only up two segments. To examine whether a similar sort of restriction holds in sound symbolic patterns, we experimentally addressed the question of whether three segments of the same sort can cause stronger sound symbolic images than two segments. The results of three experiments using Pokémon names demonstrate that three segments can indeed cause stronger sound symbolic meanings than two segments. The overall results suggest that there is a non-negligible difference between phonological patterns and sound symbolic patterns.

Keywords:
sound symbolism, counting, voiced obstruents, [p], Pokémon, Japanese

1 Introduction

1.1 The issue addressed

Sound symbolism refers to systematic connections between sounds and meanings (e.g. Dingemanse et al. 2015; Hinton et al. 2006; Sidhu & Pexman 2018). For example, in many languages, low vowels like /a/ tend to be associated with images larger than high vowels like /i/ (Newman 1933; Sapir 1929; Shinohara & Kawahara 2016). However, in modern linguistic theories, sound symbolic patterns had usually been considered to lie outside the realm of linguistic inquiry, perhaps due to the influence of the Saussurian theorem of arbitrariness that the connections between sounds and meanings in natural languages are in principle arbitrary (Saussure 1916) (see also Hockett 1959 for another influential paper on arbitrariness). However, the field has recently witnessed a rapidly increasing rise of interest on sound symbolic patterns and related phenomena (see in particular Nielsen & Dingemanse 2021), and some scholars explicitly now argue that sound symbolic patterns can—and should—be a part of phonological research (see Kawahara 2020a for a review).

Alderete & Kochetov (2017) for instance point out that expressive palatalization—e.g. patterns of palatalization observed in child-directed speech—is caused by a formal requirement to use particular types of sounds (e.g. palatal consonants and high front vowels) to express particular types of meanings, such as smallness. They propose a family of Optimality Theoretic constraints (Prince & Smolensky 1993/2004)—Express(X)—and argue that this family of constraints interacts with other phonological constraints within a single grammatical system. See also Akimbo (2021), Akimbo & Bulkaam (2024), Akita (2020), Kumagai (2019), Kumagai (to appear) and Jang (2021) for other possible cases in which sound symbolic requirements affect phonological patterns.

Approaching this issue from a slightly different perspective, Kawahara (2020b, 2021) compared particular quantitative signatures of patterns of sound symbolic judgments and those found in stochastic phonological patterns, and argued that there appears to exist an interesting parallel between the two patterns. More concretely, he argues that both sound symbolic patterns and stochastic phonological patterns exhibit what Hayes (2020, 2022) refers to as "wug-shaped curves," a quantitative signature that is predicted by Maximum Entropy Harmonic Grammar, a framework that is now widely deployed to model a wide range of phonological—and other linguistic—patterns (Goldwater & Johnson 2003; Hayes 2022; Hayes & Wilson 2008; Smolensky 1986; Zuraw & Hayes 2017).

Building on these recent proposals to consider sound symbolic patterns on a par with phonological patterns, the current experiments examine the similarity—or dissimilarity—between the two, by focusing on the counting capability (or lack thereof) of the two systems. The focus of the topic that we explore in depth in the current experiments is the classic observation that phonological systems may count up to two but no more (e.g. Goldsmith 1976; Hewitt & Prince 1989; Ito & Mester 2003; McCarthy & Prince 1986; Myers 1997; Prince & Smolensky 1993/2004 among many others). The following quote from McCarthy & Prince (1986: 1) succinctly summarizes this now-classic observation, which the current project built upon:

> Consider first the role of counting in grammar. How long may a count run? General considerations of locality, now the common currency in all areas of linguistic thought, suggest that the answer is probably 'up to two': a rule may fix on one specified element and examine a structurally adjacent element and no other.

For example, there are many languages that prohibit two occurrences of the same segments or features (i.e. dissimilation patterns: see Bennett 2015, Hansson 2001 and Suzuki 1998 for typological surveys), but no known languages prohibit three occurrences while allowing for two (Ito & Mester 2003: 265). For instance, the native phonology of Japanese prohibits morphemes with two voiced obstruents, but no known languages prohibit morphemes with three voiced obstruents. Further, an experimental investigation by Kawahara & Kumagai (2023a) using nonce words shows that Japanese speakers do not distinguish between forms with two voiced obstruents and those with three voiced obstruents—forms with three voiced obstruents were treated on a par with forms with two voiced obstruents.

Paster (2019) recently challenged the thesis that phonology can only count up to two, demonstrating that there are cases that apparently involve counting. However, Paster also points out that all those patterns that apparently count are limited to suprasegmental patterns, and none involves segmental patterns (see §3 of Paster 2019). Another challenge to the classic no-counting thesis recently came from Kim (2022), who argues that Japanese disprefers a configuration in which a voiced obstruent is followed by two nasal consonants—a restriction that apparently involves counting three segments (i.e. *[D…N…N]). However, a later examination demonstrates that evidence for this claim in the existing words is very weak at best; neither can this alleged restriction be identified in a nonce word experiment (Kawahara & Kumagai 2023b). Thus, to the best of our knowledge, it is still safe to assume that phonological patterns can count up to two segments, but not three or more.[1]

1.2 The background: Pokémonastics

In the experiments reported below, we examined whether the same restriction holds or not in sound symbolic patterns, by specifically testing whether three segments can invoke stronger sound symbolic images than two segments. We took advantage of the Pokémonastics research paradigm, which explores the nature of sound symbolism in the context of Pokémon names (Kawahara et al. 2018) (for a discussion of why it is useful to use specifically Pokémon names to explore sound symbolic patterns, see e.g. Kawahara & Breiss 2021 for a recent summary). In the Pokémon world, some characters, when they get stronger, can evolve into a different character, and in so doing their names change (e.g. /iwaaku/ → /hageneeɾu/ and /messon/ → /ʑimeɾeon/).

A quantitative study of the names of the existing Pokémon names (including those up to the 6th generation) reported by Kawahara et al. (2018) shows that the number of voiced obstruents contained in their names tend to increase as Pokémon characters evolve, a correlation which was later replicated with a larger set of data by Shih et al. (2019). A number of experimental studies that followed used nonce words and demonstrated that Japanese speakers judge nonce names

[1] One candidate for a constraint that appears to require counting three segments in its description is the one that is responsible for intervocalic lenition, which needs to prohibit a configuration in which the target consonant is flanked by two vowels (e.g. *[VTV]). However, see Katz (2021) for arguments that intervocalic lenition is a matter of phonetic implementation rather than being a phonological process.

with voiced obstruents to be more likely as those of post-evolution characters than nonce names without voiced obstruents (Kawahara 2020b, 2021; Kawahara & Kumagai 2019a, 2021). The first experiment reported below took advantage of this sound symbolic connection between voiced obstruents and Pokémon evolution status to address the question of whether three segments cause stronger sound symbolic images than two segments.

1.3 Some previous studies

Before moving on, we review some previous studies which addressed the question regarding the counting capability of sound symbolism. First, Thompson & Estes (2011) built upon the observations that some sounds are associated with images of largeness (e.g. Sapir 1929 *et seq.*). In one of their experiments, they presented native speakers of English with pictures of an imaginary creature (referred to as "greeble") in different sizes, and different nonce names containing different numbers of "large phonemes." Their results showed that the larger the size of the named objects, the more "large phonemes" were contained in their chosen names. Their result shows that the counting behavior goes well beyond two; e.g. the largest greebles were assigned names with about 4.5 "large phonemes" on average.

However, this analysis collapsed three different classes of sounds (i.e. back vowels, sonorants, and voiced stops) into one set of "large phonemes," and therefore it is impossible to tell whether it truly instantiates an unambiguous case of counting—the pattern was instead likely to have arisen from additive effects of different factors influencing the judgment patterns. The first two experiments reported below improve upon this aspect by using a class of sounds that is unambiguously a natural class, both from the phonetic and phonological perspectives. The third experiment used only one kind of segment to unambiguously exclude the possibility that the counting behavior arise from influences of different types of segments adding up.

Another candidate of counting in sound symbolism in the previous literature comes from the Pokémonastics experiments reported in Kawahara (2020b, 2021), in which he varied the numbers of moras from two to six. The results showed that each mora count increased the post-evolution responses. However, to the extent that a mora is a suprasegmental property—which seems to be a fair assumption to make—it is not clear whether these results truly instantiate a case of counting at the *segmental* level: recall that Paster (2019) identifies phonological systems may be able to count, but only at the suprasegmental level. Moreover, given the well-established status of bimoraic feet in Japanese phonology (Poser 1990) and the possibility of recursive prosodic phrasing (Ito & Mester 2012), the apparent counting behavior may be recast in terms of different foot and prosodic word structures.

In short, the current experiments attempted to address the counting capability of sound symbolism at the segmental level in the least unambiguous way possible. The first two experiments also had an advantage of being able to make the fairly direct comparison with a phonological pattern, against the recent result reported by Kawahara & Kumagai (2023a), who tested the counting behavior of voiced obstruents in Japanese phonology.

2 Experiment 1

In this experiment, the participants were given one nonce word per trial and were asked to judge whether that name is more suitable for a pre-evolution Pokémon character or a post-evolution Pokémon character. The aim was to explore whether the numbers of voiced obstruents contained in nonce names, ranging from zero to three, would impact the sound symbolic judgment of these names, and more importantly, *how*. A previous study has shown that nonce words containing one voiced obstruent is more likely to be judged as post-evolution names than those without a voiced obstruent (Kawahara 2020b), and other several studies have found that, in addition to that difference, those words containing two voiced obstruents are more likely to be judged as post-evolution names than those containing only one (Kawahara 2021; Kawahara & Kumagai 2019a, 2021).

The novel addition of the current experiment is therefore to have explored the difference between the two voiced obstruent condition and the three voiced obstruent condition. This addition is an important one, however, because it will address the question of how (dis-)similar sound symbolic patterns are with respect to phonological patterns, as discussed in §1.1. If sound symbolic patterns can count only up to two, just like phonological systems, we should not expect a difference between those words with two voiced obstruents and those with three voiced obstruents. On the other hand, if sound symbolic patterns simply count without a restriction, and then we should observe a difference between the two conditions.

2.1 Method

Following the spirit of the open science initiative in linguistics (e.g. Cho 2021, Garellek et al. 2020 and Winter 2019), the raw data, the R markdown file as well as the Bayesian posterior samples are available at the OSF repository. The link to this repository is provided at the end of the paper.

2.1.1 Stimuli

The experiment had four conditions, differing in the numbers of voiced obstruents that they contain (zero, one, two and three). Each condition had 10 items, and they were all nonce names in Japanese. They consisted of three light CV syllables. The position of voiced obstruents was controlled within each condition; e.g. in one voiced obstruent condition, they were all placed at the word-initial position. Because /p/ is known to evoke a sound symbolic effect of cuteness (Kumagai (2019); see also Experiment III), it was not used in the current stimulus set. The actual list of the stimuli is shown in Table 1.

Table 1: The list of stimuli used in the first two experiments.

VcdObs=0	VcdObs=1	VcdObs=2	VcdObs=3
[kuɕiju]	[bitaɾe]	[gebiki]	[dagigo]
[suɸuma]	[biɾejo]	[dedaɾa]	[bigade]
[neɸuɾi]	[ganija]	[zodotɕi]	[zabade]
[neɾiɾu]	[bejumi]	[zugawa]	[zegizo]
[ɕihone]	[bojatɕi]	[zudani]	[buʑido]
[kaɾutsu]	[bikohe]	[zoɕike]	[bogebi]
[jakama]	[baheho]	[zadoja]	[gegige]
[sawake]	[geseɕi]	[ʑiboɾu]	[baʑizu]
[ɾihojo]	[ʑihana]	[babohi]	[gubebi]
[sojuki]	[bijuɾi]	[gibuse]	[bibogo]

2.1.2 Procedure

The experiment was administered online using SurveyMonkey. The participants were first presented with the basic background about the Pokémon world, namely, that some Pokémon characters can evolve, and that when they evolve, they tend to get heavier, bigger and stronger. In the main session, within each trial, the participants were presented with one nonce name and were asked to judge whether each name is suitable for a pre-evolution character or a post-evolution character. The stimuli were presented in the *katakana* orthography, which is used for real Pokémon names in general. Although the stimuli were presented in written forms, the participants were asked to read and pronounce each stimulus before they register each response. The order of the stimuli was automatically randomized for each participant by SurveyMonkey.

2.1.3 Participants

We obtained data from 110 native speakers of Japanese using the Buy Response function of SurveyMonkey. The qualification requirements for participation were that (1) they had to be a native speaker of Japanese, (2) they had not previously participated in an experiment on Pokémon names and (3) they had not studied sound symbolism before. Additional data from 38 native speakers of Japanese were collected using a snowball sampling method on the first author's X (formerly Twitter).

2.1.4 Statistics

For statistical analyses, we made use of a Bayesian mixed effects logistic regression model, using the `brms` package (Bürkner 2017). We will not attempt to explicate the mechanics of Bayesian analyses in detail here, but instead refer the interested readers to accessible introductory articles, including Franke & Roettger (2019), Kruschke & Liddell (2018) and Vasishth et al. (2018). In a nutshell, Bayesian analyses combine prior information with the obtained experimental data and produce a range of possible values—which are referred to as posterior distributions—for each estimated parameter.

One advantage of Bayesian analyses is that we can interpret the posterior distributions as representing the likely values of the estimated parameters. One heuristic to interpret the results of Bayesian modeling is to examine the middle 95% of the posterior distribution, known as 95% Credible Interval (henceforth, abbreviated as "95% CrI"), of the coefficient we are interested in. If the 95% CrI of a parameter does not include 0, then that parameter can be considered to be meaningful. However, unlike in a frequentist analysis, we do not have to rely on a strict dichotomy (i.e. "significant" vs. "non-significant" or "credible/meaningful" vs. "not credible/meaningful"). We can instead examine how many samples in the posterior distribution are in the expected direction, i.e., the probability of a particular hypothesis being true.

Another advantage of Bayesian analysis is that we can also address the question regarding with how much confidence we can conclude a null effect (Gallistel 2009). This is an important feature for the case at hand, because if sound symbolism were to behave like phonological patterns, we would expect a null difference between the two voiced obstruent condition and the three voiced obstruent condition (cf. Kawahara & Kumagai 2023a). If it turned out to be that way, we wanted to explore how likely it is that there are truly no differences, which is impossible to test with a frequentist analysis.

Moving on the the specifics of the model specifications for the current experiment, the binary dependent variable was whether each item was judged as a post-evolution character name (=1) or not (=0). The fixed independent variable was the number of voiced obstruents contained in the stimuli. This factor was coded as a categorical factor, so that we could make the targeted comparison between the two voiced obstruent condition and three voiced obstruent condition, which was implemented using `hypothesis` function. The baseline of this factor was arbitrarily chosen as the condition with zero voiced obstruents. In addition to this fixed factor, a random intercept of items and participants as well as the random slopes of participants for the fixed factor were included in the model. For prior specifications, a Normal(0, 1) weakly informative prior for the intercept (Lemoine 2019) and a Cauchy prior with scale of 2.5 for the slope (Gelman et al. 2018) were used.

Four chains with 2,000 iterations were run, and the first 1,000 iterations from each chain were discarded as warmups. All the $\hat{R}$-values for the fixed effects were 1.00 and there were no divergent transitions. See the R markdown file available at the OSF repository for complete details.

2.2 Results

Figure 1 shows the distribution of post-evolution response ratios for each voiced obstruent condition in the form of violin plots, in which the widths represent normalized probability distributions. Transparent light-blue circles, jittered slightly to avoid overlap, represent average responses for each condition from each participant. Solid red circles are the grand averages in each condition, with their 95% confidence intervals calculated by `ggplot:` (Wickham 2016).

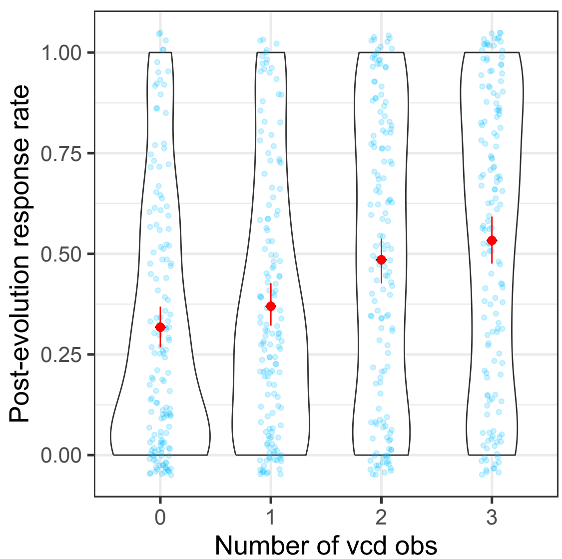

Figure 1: The results of Experiment 1, showing the distribution of post-evolution responses for each condition.

We observe a steady increase in the post-evolution responses, as the number of the voiced obstruents contained in the stimuli increase: the four conditions resulted in the following averages: 0.32 vs. 0.37 vs. 0.49 vs. 0.53.[2] The central coefficient estimate of the difference between the baseline (no voiced obstruents) and one voiced obstruent is 0.27, with its 95% CrI being [-0.16, 0.70]. Although this interval includes zero, the posterior distribution is skewed toward positive values, and 90% of the posterior samples were positive.

More importantly, a targeted comparison between the two voiced obstruent condition and three voiced obstruent condition shows that the central coefficient estimate for this difference is

[2]Even those words that contain three voiced obstruents were judged to be post-evolution names only slightly above 50%, which was a bit surprising. Some participants reported after the experiment that post-evolution names should be longer than three moras. See Kawahara et al. (2018) and Kawahara (2020b) for the effects of name length.

-0.39 with its 95% CrI being [-0.67, -0.11] and the posterior probability supporting this difference is 0.99. For the sake of completeness, the difference between one voiced obstruent condition and the two voiced obstruent condition was also calculated, which turned out to be robust, with its central coefficient and 95% CrI being -0.77 and [-1.1, -0.44], respectively. Its posterior probability was 1.00. In short, we observe that each difference between the four conditions was meaningful (although we can be only 90% confident about the difference between the first two conditions).

2.3 Discussion

The current experiment first of all replicated the findings of the previous studies that given nonce words, Japanese speakers do indeed generally associate voiced obstruents with post-evolution Pokémon names (Kawahara 2020b, 2021; Kawahara & Kumagai 2019a, 2021). It moreover found that names with three voiced obstruents were more likely to be associated with post-evolution characters than those with two voiced obstruents, suggesting that sound symbolic patterns can function in an additive fashion, and count at least up to three (cf. Thompson & Estes 2011). This case may be likened to cases of *counting cumulativity* in linguistic patterns (Breiss 2020; Hayes 2022; Jäger 2007; Jäger & Rosenbach 2006), in which multiple violations of the same phonological constraint function in an additive manner.[3]

The current result is particularly interesting in the light of the general question of how similar phonological patterns and sound symbolic patterns are, given the recent proposals that these two systems may have more in common than previously thought (e.g. Alderete & Kochetov 2017; Kawahara 2020a,b), as reviewed in §1.1. Assuming that it is indeed a true property of phonological systems that it can count only up to two segments (e.g. Ito & Mester 2003; McCarthy & Prince 1986; Prince & Smolensky 1993/2004), just as Japanese phonology counts only up to two voiced obstruents (Kawahara & Kumagai 2023a), the fact that sound symbolic patterns related to voiced obstruents can count up to three would instantiate a non-trivial difference between the two systems. At least within Japanese, the way its phonology handles voiced obstruents and the way voiced obstruents invoke their sound symbolic images differ from one another.

[3] The cases of counting cumulativity may raise the question of whether phonology may indeed be able to count. However, we are not aware of any convincing case of counting cumulativity that involves three loci, either in phonological alternations or in phonotactics. See Breiss (2020) for an informative review of the cases of counting cumulativity, as well as the other type of cumulativity (i.e. ganging-up cumulativity).

3 Experiment II

3.1 Preamble

To extend the scope of the findings from Experiment I, we tested another semantic dimension that can be symbolically signaled by voiced obstruents. In Japanese (and perhaps other languages), voiced obstruents are associated with general negative images (Hamano 1998), and in the context of Pokémon names, they are overrepresented in the names of villainous characters (Uno et al. 2020). More specifically, some Pokémon characters belong to particular "types", and it has been found that voiced obstruents are overrepresented in the names of the "dark type" characters (Uno et al. 2020). The productivity of this sound symbolic relationship has been confirmed by an experiment using nonce words (Kawahara & Kumagai 2019b). Experiment II made use of this previously identified sound symbolic relationship to further address the counting capability of sound symbolic patterns.

There are a few differences between Experiment I and Experiment II. In Experiment II, the participants were asked whether each name was suitable for a dark-type character or normal-type character. Before the main trials, they were told that all Pokémon characters belong to at least one type, with two examples; /çitokage/ 'Charmander (fire lizard)' belong to the "fire" type, and /goosu/ belong to both "ghost" type and "dark" type. The stimuli used in the experiment were identical to those used in Experiment I. The participants were university students from Meiji University. After excluding data from those who are not native speakers of Japanese and those who were familiar with research on sound symbolism, the data from 141 native speakers entered the statistical analysis. The details of the statistical modeling were identical to those of Experiment I.

3.2 Results

Figure 2 shows the results of Experiment II. As with Experiment I, we observe a steady increase in the dark-type response ratio, as the number of voiced obstruents contained in the stimuli increase. The grand averages for each conditions were 0.18 vs. 0.43 vs. 0.71 vs. 0.79.

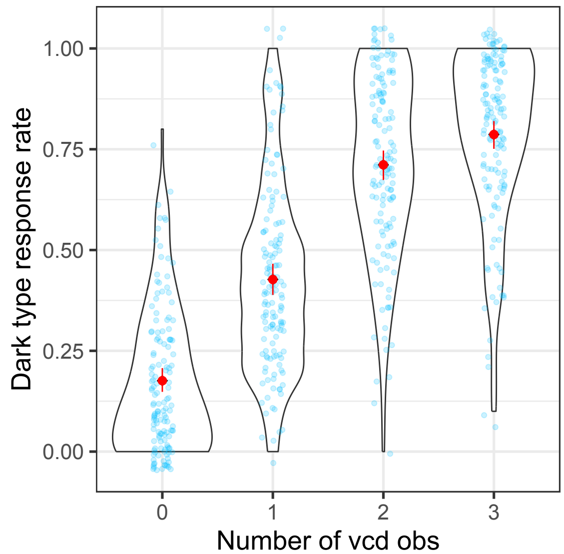

Figure 2: The results of Experiment II. The distribution of dark-type response ratios for each voiced obstruent condition.

This effect of voiced obstruents is a very robust one according to the Bayesian modeling. The difference between the no voiced obstruent condition (i.e. the baseline) and one voiced obstruent was very credible, with its central coefficient estimate being 1.48, with its 95% CrI being [0.78, 2.16]. All the posterior samples were positive. More importantly, the difference between the two voiced obstruent condition and the three voiced obstruent condition was also credible. The central coefficient estimate is -0.62 and its 95% CrI is [-1.15, -0.09]. The posterior probability of this targeted comparison is 0.97. The difference between the one voiced obstruent and two voiced obstruents was also robust (central estimate = -1.15, 95% CrI=[-2.09, -1.03], posterior probability supporting the difference = 1).

3.3 Discussion

The sound symbolic effects of voiced obstruent were clearer in Experiment II than in Experiment I—names with zero voiced obstruents were unlikely to be judged as dark-type characters, whereas names with three voiced obstruents were very likely to be judged as dark-type characters. And most importantly for the current purpose, we have found a solid distinction between the two voiced obstruent condition and the three voiced obstruent condition. The fact that this difference holds is unlike how voiced obstruents are treated by the Japanese phonological system (Kawahara & Kumagai 2023a), which is arguably a general property of phonological systems in natural languages (Ito & Mester 2003; McCarthy & Prince 1986; Prince & Smolensky 1993/2004).

4 Experiment III

4.1 Introduction

The previous two experiments have shown that a distinction between two segments and three segments matters when it comes to sound symbolic patterns—a distinction that phonological systems arguably do not make. However, in both experiments, the target sounds were voiced obstruents, so it seemed important to us to examine how generalizable this counting property is, i.e. whether this counting capability is observed for sound symbolic patterns that are caused by segments other than voiced obstruents.

Also, we felt it useful to address the possibility that the patterns we observed in the previous two experiments arose from different types of voiced obstruents—e.g. [z] and [d]—"ganging-up" rather than the patterns arising from pure counting (cf. Jäger & Rosenbach 2006; Jäger 2007). We reiterate that it is safe to say that a voiced obstruent is a coherent set of sounds both from the phonetic and phonological perspective in Japanese (Ito & Mester 2003). Ultimately, the distinction between counting cumulativity and ganging-up cumulativity depends on whether what adds up is multiple violations of the *same* constraint or those of *different* constraints, in the parlance of constraint-based theories such as Optimality Theory (Prince & Smolensky 1993/2004). Thus, the results of the first two experiments instantiate *counting* cumulativity to the extent that we can formulate Express(VcdObs) as a single constraint. Given that voiced obstruents behave as a natural class both in phonological and sound symbolic patterns in Japanese (Ito & Mester 2003; Hamano 1998), we believe that this is a fairly safe assumption to make.

Nevertheless, it is safer to be conservative and entertain the possibility that different effects of voiced obstruents are governed by different constraints, or to put it in a more theory-neutral term, different sound symbolic forces. To this end, we took advantage of the sound symbolic connection between [p] and "cuteness" (Kumagai 2019, 2022), which also manifests itself in the fact that labial sounds, including [p] are, overrepresented in the cute, fairy type Pokémon characters (Kawahara & Kumagai 2019b; Uno et al. 2020).

4.2 Method

Experiment III used the set of stimuli shown in Table 2. The experiment, like Experiments I and II, varied the number of [p]s that are contained in the stimuli. The position of [p] was controlled within each condition. Each condition had 10 items, all of which contain only light CV syllables. Since there could be a difference between sonorants and obstruents in terms of their impact on cuteness judgments (Kawahara et al. 2015; Shinohara & Kawahara 2013), the syllables not containing [p] all had a voiceless obstruent onset.

Table 2: The list of stimuli used in Experiment III.

[p]=0	[p]=1	[p]=2	[p]=3
[kuɕisu]	[pitahe]	[pepiki]	[papipe]
[sutsuka]	[piketo]	[papeka]	[pipape]
[kusuki]	[patɕiha]	[pepotɕi]	[popape]
[teɕiku]	[pekuɕi]	[pupata]	[pepipo]
[ɕihake]	[posatɕi]	[popaɕi]	[pupipo]
[kesutsu]	[pikohe]	[popike]	[popepi]
[tokaha]	[paheto]	[papoka]	[pepipe]
[sahake]	[peseki]	[popitsu]	[papupi]
[tɕihoto]	[pihaka]	[papoçi]	[pupepi]
[sokuki]	[pisutɕi]	[pipuse]	[pipope]

The responses were gathered using the Buy Response function of SurveyMonkey. Data from a total of 150 native speakers of Japanese were obtained. In this experiment, the participants were asked, for each name, whether the name is more suitable for a normal type character or a cute fairy type character. The details of the statistical analysis were identical to those of Experiments I and II.

4.3 Results

The results are presented in Figure 3, which shows the distribution of the fairy type character responses for each condition having different numbers of [p]. Similar to the two previous experiments, we observe a steady increase in the fairy response, as the number of [p]s contained in the names increases. The grand averages were: zero [p] = 0.21; one [p] = 0.39; two [p]s = 0.47; three [p]s = 0.57.

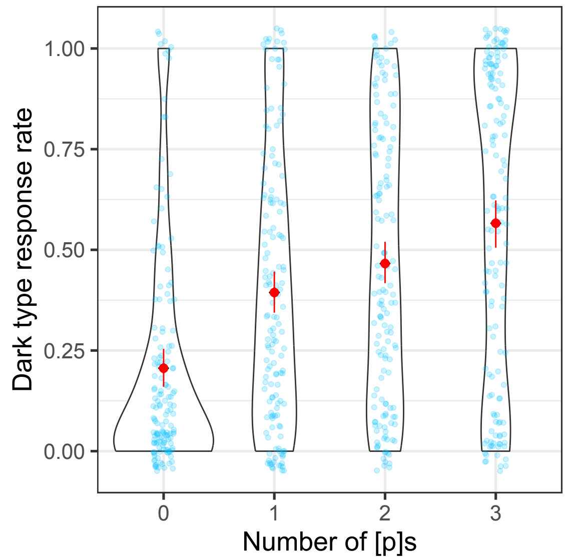

Figure 3: The results of Experiment III. The distribution of fairy-type response ratios for each condition, which contained different numbers of [p]s.

The results of the Bayesian logistic regression show that there is a clear difference between the baseline ([p]=0) and the condition which contains one [p] ($\beta = 1.50$, its 95% CrI = [0.97, 2.03], with all their posterior samples supporting the difference). The difference between the two [p] condition and the three [p] condition, which is most important for the purpose of the current study, was also very robust (β = -0.8 with its 95% CrI [-1.24, -0.38], and all the posterior samples support this difference). To be complete, the difference between the one [p] condition and the two [p] condition was also a reliable one (β = -0.45, its 95% CrI [-0.82, -0.08] and 97% of the posterior samples supporting the difference). In short, every addition of [p] in the names reliably increased the fairy responses.

4.4 Discussion

This experiment again shows that sound symbolism can count up to three. In order words, the counting capability is not a specific property of voiced obstruents, possibly different kinds of voiced obstruents "ganging-up" (Jäger & Rosenbach 2006; Jäger 2007), but it holds with one kind of segment—[p]—invoking the image of cuteness.

5 General discussion

We started with a general question—how (dis-)similar sound symbolic patterns are with respect to phonological patterns. To address this question, we focused on one property of phonological systems which seems to hold robustly across languages; at least when it comes to the patterns of segmental phonology, it can count only up to two segments, but no more. No known languages have been identified to prohibit three occurrences of the same segment/feature, whereas there are a plethora of examples in which two occurrences of the same segment are banned. Japanese precisely instantiates a case of this kind in which two voiced obstruents within morphemes are prohibited (Ito & Mester 2003), and experiment-wise too, Japanese speakers treat forms with three voiced obstruents on a par with forms with two voiced obstruents (Kawahara & Kumagai 2023a).

To the extent that sound symbolic patterns and phonological patterns are governed by the same system (see Alderete & Kochetov 2017 and Kawahara 2020b, in particular), we would have expected that a similar restriction would hold—that Japanese speakers would treat forms with three voiced obstruents just like forms with two voiced obstruents, when they make sound symbolic judgements. However, the results of two experiments show that this expectation did not hold up, when Japanese speakers make sound symbolic judgments of forms with different numbers of voiced obstruents. These results were further corroborated by an additional experiment which shows that three [p]s can evoke stronger sound symbolic images than two [p]s. It thus seems safe to conclude, given these results, that there is a non-negligible difference between the phonological patterns and sound symbolic patterns, at least in terms of their counting capabilities.

As Alderete & Kochetov (2017) and others have shown (Akimbo 2021; Akimbo & Bulkaam 2024; Akita 2020; Kumagai 2019, to appear; Jang 2021), sound symbolic requirements can affect phonological patterns. To the extent that our conclusion is on the right track, then, when such sound symbolic effects are incorporated into a phonological grammar, there should be some kind of filter that "strips off" the counting capability of sound symbolic mechanisms. Otherwise, we would expect there to be a constraint like Express(ThreeVcdObs), which requires that there be at least three voiced obstruents to express a particular semantic notion. While it remains to be seen that such patterns are indeed impossible in human languages, at this point we find it very unlikely. And if such filtering mechanism is to be required, it may be something that is akin to an abstraction mechanism that is at work when phonetic effects are grammaticalized into a phonological system (Gordon 2002; Hayes 1999; Smith 2002), which reflects a general observation that even when phonetic factors appear to drive phonological generalizations, some details are abstracted away from in the phonology system.

Conflicts of interest

We declare no conflicts of interest.

Availability of data and material

The data are available at `https://osf.io/zhnda/?viewonly=de5ffbd83dc24a1eb6db3b11af08c550`

Code availability (software application or custom code)

The code is also available at `https://osf.io/zhnda/?viewonly=de5ffbd83dc24a1eb6db3b11af08c550`

References

Akimbo, Samuel. 2021. Featural affixation and sound symbolism in Fungwa. *Phonology* 38(4). 537–569.

Akimbo, Samuel & Michael Bulkaam. 2024. Iconicity as the motivation for the signification and locality of deictic grammatical tones in Tal. *Glossa* 24(9).

Akita, Kimi. 2020. A typology of depiction marking: The prosody of japanese ideophones and beyond. *Studies in Language* .

Alderete, John & Alexei Kochetov. 2017. Integrating sound symbolism with core grammar: The case of expressive palatalization. *Language* 93. 731–766.

Bennett, Wm. G. 2015. *The phonology of consonants*. Cambrige: Cambridge University Press.

Breiss, Canaan. 2020. Constraint cumulativity in phonotactics: Evidence from artificial grammar learning studies. *Phonology* 37(4). 551–576.

Bürkner, Paul-Christian. 2017. brms: An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software* 80(1). 1–28.

Cho, Taehong. 2021. Where we are at: Impact, special collections, open science and registered report at the *journal of phonetics*. *Journal of Phonetics* 89.

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

Franke, Michael & Timo B. Roettger. 2019. Bayesian regression modeling (for factorial designs): A tutorial. Ms. `https://doi.org/10.31234/osf.io/cdxv3`.

Gallistel, Randy C. 2009. The importance of proving the null. *Psychological Review* 116(2). 439–453.

Garellek, Marc, Matthew Gordon, James Kirby, Wai-Sum Lee, Alexis Michaud, Christine Mooshammer, Oliver Niebuhr, Daniel Recasens, Timo B. Roettger, Adrian Simpson & Kristine Yu. 2020. Toward open data policies in phonetics: What we can gain and how we can avoid pitfalls. *Journal of Speech Sciences* 9(1).

Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau & Yu-Sung Su. 2018. A weakly informative default prior distribution for logistic and other regression models. *Annual Applied Statistics* 2(4). 1360–1383.

Goldsmith, John. 1976. *Autosegmental phonology*: MIT Doctoral dissertation.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. *Proceedings of the Workshop on Variation within Optimality Theory* 111–120.

Gordon, Matthew. 2002. A phonetically driven account of syllable weight. *Language* 78(1). 51–80.

Hamano, Shoko. 1998. *The sound-symbolic system of Japanese*. Stanford: CSLI Publications.

Hansson, Gunnar Olafur. 2001. *Theoretical and typological issues in consonant harmony*: University of California, Berkeley Doctoral dissertation.

Hayes, Bruce. 1999. Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In Michael Darnell, Edith Moravscik, Michael Noonan, Frederick Newmeyer & Kathleen Wheatly (eds.), *Functionalism and formalism in linguistics, vol. 1: General papers*, 243–285. Amsterdam: John Benjamins.

Hayes, Bruce. 2020. Assessing grammatical architectures through their quantitative signatures. Talk presenetd at BLS.

Hayes, Bruce. 2022. Deriving the wug-shaped curve: A criterion for assessing formal theories of linguistic variation. *Annual Review of Linguistics* 8. 473–494.

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379–440.

Hewitt, Mark & Alan Prince. 1989. OCP, locality and linking: The N. Karanga verb. In E. J. Fee & K. Hunt (eds.), *Proceedings of West Coast Conference on Formal Linguistics 8*, 176–191. Stanford: Stanford Linguistic Association.

Hinton, Leane, Johanna Nichols & John Ohala. 2006. *Sound symbolism, 2nd edition*. Cambridge: Cambridge University Press.

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

Ito, Junko & Armin Mester. 2003. *Japanese morphophonemics*. Cambridge: MIT Press.

Ito, Junko & Armin Mester. 2012. Recursive prosodic phrasing in Japanese. In Toni Borowsky, Shigeto Kawahara, Takahito Shinya & Mariko Sugahara (eds.), *Prosody matters*, 280–303. London: Equinox Publishing.

Jäger, Gerhard. 2007. Maximum Entropy Models and Stochastic Optimality Theory. In Joan W. Bresnan (ed.), *Architectures, rules, and preferences: Variations on themes*, 467–479. CSLI.

Jäger, Gerhard & Anette Rosenbach. 2006. The winner takes it all—almost: Cumulativity in grammatical variation. *Linguistics* 44(5). 937–971.

Jang, Hayeun. 2021. How cute do I sound to you?: Gender and age effects in the use and evaluation of Korean baby-talk register, Aegyo. *Language Sciences* 83. 101289.

Katz, Jonah. 2021. Intervocalic lenition is not phonological: Evidence from Campidanese Sardinian. *Phonology* 38(4). 651–692.

Kawahara, Shigeto. 2020a. Sound symbolism and theoretical phonology. *Language and Linguistic Compass* 14(8). e12372.

Kawahara, Shigeto. 2020b. A wug-shaped curve in sound symbolism: The case of Japanese Pokémon names. *Phonology* 37(3). 383–418.

Kawahara, Shigeto. 2021. Testing maxent with sound symbolism: A stripy wug-shaped curve in japanese pokémon names. *Language (Research Report)* .

Kawahara, Shigeto & Canaan Breiss. 2021. Exploring the nature of cumulativity in sound symbolism: Experimental studies of Pokémonastics with English speakers. *Laboratory Phonology* 12(1). 3, `https://doi.org/10.5334/labphon.280`.

Kawahara, Shigeto & Gakuji Kumagai. 2019a. Expressing evolution in Pokémon names: Experimental explorations. *Journal of Japanese Linguistics* 35(1). 3–38.

Kawahara, Shigeto & Gakuji Kumagai. 2019b. Inferring Pokémon types using sound symbolism: The effects of voicing and labiality. *Journal of the Phonetic Society of Japan* 23(2). 111–116.

Kawahara, Shigeto & Gakuji Kumagai. 2021. What voiced obstruents symbolically represent in Japanese: Evidence from the Pokémon universe. *Journal of Japanese Linguistics* 37(1). 3–24.

Kawahara, Shigeto & Gakuji Kumagai. 2023a. Lyman's law counts only up to two. *Laboratory Phonology* .

Kawahara, Shigeto & Gakuji Kumagai. 2023b. Rendaku is not blocked by two nasal consonants: A reply to Kim (2022). *Glossa* .

Kawahara, Shigeto, Atsushi Noto & Gakuji Kumagai. 2018. Sound symbolic patterns in Pokémon names. *Phonetica* 75(3). 219–244.

Kawahara, Shigeto, Kazuko Shinohara & Joseph Grady. 2015. Iconic inferences about personality: From sounds and shapes. In Masako Hiraga, William Herlofsky, Kazuko Shinohara & Kimi Akita (eds.), *Iconicity: East meets west*, 57–69. Amsterdam: John Benjamins.

Kim, Seoyoung. 2022. A maxent learner for super-additive counting cumulativity. *Glossa* 7(1).

Kruschke, John K. & Torrin M. Liddell. 2018. The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychological Bulletin and Review* 25. 178–206.

Kumagai, Gakuji. 2019. A sound-symbolic alternation to express cuteness and the orthographic Lyman's Law in Japanese. *Journal of Japanese Linguistics* 35(1). 39–74.

Kumagai, Gakuji. 2022. What's in a Japanese *kawaii* 'cute' name? A linguistic perspective. *Frontiers in Psychology* https://doi.org/10.3389/fpsyg.2022.104.

Kumagai, Gakuji. to appear. EXPRESS[P] in expressive phonology: Analysis of a nicknaming pattern using 'princess' in Japanese. *Phonology* .

Lemoine, N.P. 2019. Moving beyond noninformative priors: Why and how to choose weakly informative priors in bayesian analyses. *Oikos* 128. 912–928.

McCarthy, John J. & Alan Prince. 1986. Prosodic morphology. Ms., University of Massachusetts and Rutgers University.

Myers, Scott. 1997. OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15(4). 847–892.

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

Nielsen, Alan K. S. & Mark Dingemanse. 2021. Iconicity in word learning and beyond: A critical review. *Language and Speech* 64(1). 52–72.

Paster, Mary. 2019. Phonology counts. *Radical* 1. 1–61.

Poser, William. 1990. Evidence for foot structure in Japanese. *Language* 66. 78–105.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar.* Malden and Oxford: Blackwell.

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, Hayeun Jang, Jessica Johnson, Darya Kavitskaya, Shigeto Kawahara, Miran Oh, Rebecca L Starr & Alan Yu. 2019. Crosslinguistic and language-specific sound symbolism: Pokémonastics. Ms. University of Southern California, University of California, Merced, University of California, Berkeley, Keio University, National University of Singapore and University of Chicago.

Shinohara, Kazuko & Shigeto Kawahara. 2013. The sound symbolic nature of Japanese maid names. *Proceedings of the 13th Annual Meeting of the Japanese Cognitive Linguistics Association*

13. 183–193.
Shinohara, Kazuko & Shigeto Kawahara. 2016. A cross-linguistic study of sound symbolism: The images of size. In *Proceedings of the Thirty Sixth Annual Meeting of the Berkeley Linguistics Society.*, 396–410. Berkeley: Berkeley Linguistics Society.
Sidhu, David & Penny M. Pexman. 2018. Five mechanisms of sound symbolic association. *Psychonomic Bulletin & Review* 25(5). 1619–1643.
Smith, Jennifer. 2002. *Phonological augmentation in prominent positions*: University of Massachusetts, Amherst Doctoral dissertation.
Smolensky, Paul. 1986. Information processing in dynamical systems: Foundations of harmony theory. In D. Rumelhart, J. McClelland & PDPR Group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 1: Foundations, 194–281. Cambridge, MA: Bradford Books/MIT Press.
Suzuki, Keiichiro. 1998. *A typological investigation of dissimilation*: University of Arizona Doctoral dissertation.
Thompson, Patrick D. & Zachary Estes. 2011. Sound symbolic naming of novel objects is a graded function. *Quarterly Journal of Experimental Psychology* 64(12). 2392–2404.
Uno, Ryoko, Kazuko Shinohara, Yuta Hosokawa, Naho Ataumi, Gakuji Kumagai & Shigeto Kawahara. 2020. What's in a villain's name? Sound symbolic values of voiced obstruents and bilabial consonants. *Review of Cognitive Linguistics* 18(2). 428–457.
Vasishth, Shravan, Bruno Nicenboim, Mary Beckman, Fangfang Li & Eun Jong Kong. 2018. Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics* 71. 147–161.
Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis.* New York: Springer-Verlag.
Winter, Bodo. 2019. *Statistics for linguists.* New York: Taylor & Francis Ltd.
Zuraw, Kie & Bruce Hayes. 2017. Intersecting constraint families: An argument for Harmonic Grammar. *Language* 93. 497–548.