To appear in Phonology

Sound symbolism can count three segments (whereas phonological constraints presumably cannot)*

Shigeto Kawahara and Gakuji Kumagai
Keio University and Kansai University

Abstract

Some researchers have recently argued that sound symbolic requirements can cause phonological alternations, suggesting that sound symbolic patterns and phonological patterns may be governed by similar—or perhaps the same—mechanisms. Against this theoretical development, this paper further addresses the question of how similar phonological systems and sound symbolic systems are, by focusing on their counting capability. It has been known that phonological constraints can count only up to 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 do indeed cause stronger sound symbolic meanings than two segments. The overall results suggest that phonological systems and sound symbolic systems have a distinct characteristic, in that only the latter systems have a certain type of counting capability.

Keywords: sound symbolism, counting, voiced obstruents, [p], Pokémon, Japanese
Approximate word count: 6,000

*We are grateful to four anonymous reviewers and the associate editor of the journal for their very constructive suggestions. The research reported in this paper is supported by JSPS grants (#22K00559, #25K04035 and #19K13164). All remaining errors are ours.

1 Introduction

1.1 The relationship between phonology and sound symbolism

Sound symbolism refers to systematic connections between sounds and meanings (e.g. Akita 2015; Dingemanse et al. 2015; Hinton et al. 2006; Perniss et al. 2010; 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; Thompson & Estes 2011). 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 for some quantitative evidence). Some scholars now explicitly argue that exploration of sound symbolic patterns can—and should—be a part of phonological research (see Kawahara 2020a for a review of the arguments for this view).

For instance, Alderete & Kochetov (2017) 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 Akinbo (2021), Akinbo & Bulkaam (2024), Akita (2020), Klamer (2002), Dingemanse & Thompson (2020), Kumagai (2019, 2023) and Jang (2021) for other possible cases in which sound symbolic requirements affect—or at least, interact with—phonological patterns; see also Mithun (1982) and Monaghan & Roberts (2021) for possible influences of sound symbolic effects on diachronic changes, where expressive vocabularies resisted diachronic sound changes that applied to other regular, non-iconic vocabulary items.

Approaching this issue from a slightly different perspective, Kawahara (2020b) 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 (MaxEnt HG), 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; McPherson

& Hayes 2016; Shih 2017; Smolensky 1986; Zuraw & Hayes 2017).

In short, an increasing number of studies have recently argued that sound symbolic patterns and phonological patterns are governed by similar—or perhaps, the same—mechanisms.

1.2 Counting capability of phonology or lack thereof

Building on these recent proposals which treat 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. To preview the conclusions that follow from the current experimentation, we will show in this paper that phonological systems and sound symbolic systems have a clearly distinct characteristic, in that only the sound symbolic systems have a certain type of counting capability.

In order to address the (dis)similarity between the phonological systems and sound symbolic systems, the current experiments make use of the classic observation that phonological systems may count up to two but no more (e.g. Goldsmith 1976; Hayes 1995; Hewitt & Prince 1989; Ito & Mester 2003; McCarthy & Prince 1986; Myers 1997; Nelson & Toivonen 2000; Prince & Smolensky 1993/2004; Walker 2001 among many others).[1] While some apparent cases of counting have recently been pointed out in the literature, the following generalizations still hold robustly across known languages:

(1) No counting in phonology

a. No phonological constraints require the presence of three segments/features.

b. No phonological constraints prohibit three occurrences of the same feature/segment.

Let us now review the critical observations made in the literature on this topic in further detail. This now-classic thesis of "no-counting" in phonology was tacitly assumed in many phonological analyses, but was clearly expressed by McCarthy & Prince (1986: 1), who stated:

> 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.

[1] The same thesis is likely to hold in syntax (Chomsky 1965; Haspelmath 2014). In *Aspects of the Theory of Syntax*, Chomsky (1965) lists a number of syntactic operations that would be possible if syntax had the capability to count, which seem nevertheless be impossible in natural languages. To quote, "reflection of an arbitrary string (that is, replacement of any string $a_1....a_n$, where each a_i is a single symbol, by $a_n...a_1$), or interchange of the $(2n-1)^{th}$ word with the $2n^{th}$ word throughout a string of arbitrary length, or insertion of a symbol in the middle of a string of even length (pp. 55-56).

To be more concrete, McCarthy & Prince (1986) for instance argue that there exist no reduplicative patterns which copy exactly three segments from the base. Schematically, such a reduplicative pattern would look like [bad-badupi], [bia-biadupi], [adu-adupi] and [bla-bladupi], with the reduplicant's shape varing from CVC, CVV, VCV to CCV. To the best of our knowledge, no such reduplicative patterns have been found even after 1986.

Also, 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 extensive typological surveys), but no known languages prohibit three occurrences while allowing for two (Ito & Mester 2003: 265). A well-known example comes from the native phonology of Japanese, which prohibits morphemes with two voiced obstruents; on the other hand, no known languages prohibit morphemes with three voiced obstruents, while allowing for two. 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.

Prince & Smolensky (1993/2004), as they proposed Optimality Theory (OT), spend some good portions of their book discussing why their proposed system does not involve counting; for example, they state that a comparison between two candidates based on the numbers of violations of a particular constraint "is not numerical counting, but simply comparisons of *more* and *less*" (p. 83) (see also their §10.1.1). McCarthy (2003) also argues that OT constraints should not count or assess "degrees of violations", stating that "no language requires the presence of at least three round vowels to initiate rounding harmony, nor do we ever find that complementisers may be doubly but not trebly filled" (p. 80).

However, some possible exceptions to the non-counting thesis have been pointed out in some recent work, although as we will see, the generalizations in (1) still seem to hold. First, Paster (2019) challenged the thesis that phonology can only count up to two, demonstrating that there are cases that apparently involve counting. She, for example, proposes a tonal association rule for Kuria, by which the H-tone is associated with the *fourth* mora from the left edge of a stem. 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, implying the presence of a constraint 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 the productivity of this alleged restriction be identified in a nonce word experiment (Kawahara & Kumagai 2023b).

Finally, some studies have demonstrated that multiple reduplications can induce more intensified meanings, for instance in Fungwa (Akinbo 2023). These patterns may mean that morphological operations (i.e. reduplication) can apply multiple times, and that each operation has a semantic impact. However, these patterns do not necessarily imply that a single phonological constraint has a capability to count beyond two segments.

To summarize, to the best of our knowledge, it is still safe to assume that the general "no-counting" principles, or at least those specific implementations stated in (1), hold as a property of the phonological systems at the segmental level in natural languages. Put from a slightly different perspective, phonological constraints—as we formulate them in OT analyses—related to segmental phonology can count up to two segments, but not three or more in their structural description (McCarthy 2003).[2]

1.3 The background about the current experiments: Pokémonastics

In the experiments reported below, we examined whether the non-counting nature observed in phonological systems would hold 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 in general, see e.g. Kawahara & Breiss 2021 for a 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 with voiced obstruents to be more likely as those of post-evolution characters than nonce names without voiced obstruents (Kawahara 2020b; Kawahara & Kumagai 2019a). 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.

[2] One candidate for a constraint that appears to require counting three segments in its structural 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.

1.4 Previous observations about sound symbolisms

Before moving on, we review some previous studies which addressed 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": Gauthier & Tarr 1997) 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, reproduced below as Figure 1, 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.

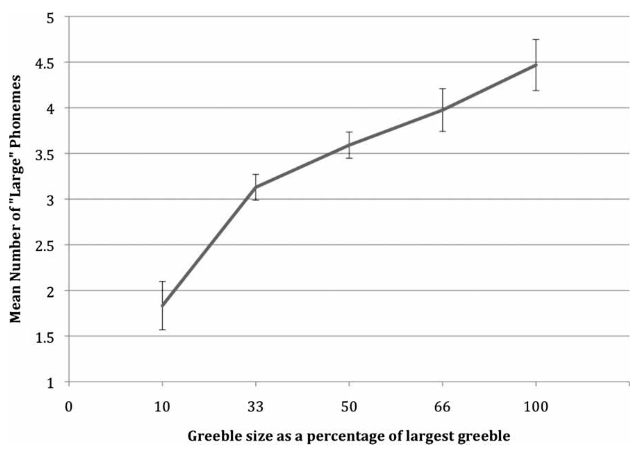

Figure 1: Results of Thompson and Estes 2011 (their Figure 3), in which the larger the named objects, the more "large phonemes" their names contained.

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 three different factors influencing the judgment patterns.[3] Similarly, there exist several other studies which showed cumulative effects of sound symbolism (Cuskley 2013; D'Onofrio 2014; Dingemanse & Thompson 2020; Priestly 1994), but their results are likely to have arisen from additive effects of different factors, just like the results of Thompson & Estes (2011) in Figure 1.

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

[3] In the parlance of recent linguistic theorization, this would be comparable to a case of ganging-up cumulativity (Jäger 2007; Jäger & Rosenbach 2006).

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.[4]

Another candidate of counting in sound symbolism in the previous literature comes from the Pokémonastics experiments reported in Kawahara (2020b), 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 (McCarthy & Prince 1986)—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 (Ito 1990; Mester 1990; Poser 1990) and the possibility of recursive prosodic phrasing (Ito & Mester 2012, 2013), 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 a fairly direct within-language 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 I

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

[4]One way to understand the current framing of the question from the perspective of modern phonological theorization may be that our experiments reported below address the question of whether sound symbolic patterns would show a pattern of counting cumulativity. We hasten to reiterate here, however, that while phonological patterns may show evidence for counting cumulativity (Breiss 2020; Hayes 2022), it still holds true that a single constraint cannot count three segments in their structural description, as discussed in detail in §1.2.

The cases of counting cumulativity in phonology may raise the question of whether phonology may indeed be able to count. However, there are no convincing case of counting cumulativity that involves three loci, either in phonological alternations or in phonotactics (McCarthy 2003—see Breiss 2020 for an informative review of the cases of cumulativity), and therefore, in this sense, the counting capability of phonological systems has to be limited.

If we are to deploy a theoretical mechanism like MaxEnt HG which allows for counting cumulativity, then we would have to make sure that constraints do not assign a violation mark based on a structural description that involves more than two segments. In other words, the grammar may be able to count the number of violations so that it can multiply them by the constraint weights, but the constraints themselves cannot count the number of segments to calculate their violations. See Kawahara & Kumagai (2023a) for further discussion on this point.

obstruent (Kawahara 2020b), and other 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 (e.g. Kawahara & Kumagai 2019a).

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 the nature of segmental, phonological constraints, as discussed in §1.1 and §1.2.

If sound symbolic patterns can count only up to two, just like phonological constraints, we should not expect a difference between those words with two voiced obstruents and those with three voiced obstruents—recall that in terms of phonological Lyman's Law, three voiced obstruents are no different from two 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

The raw data, the R markdown file as well as the Bayesian posterior samples are available at the OSF repository (for the importance of the open science policy in linguistic studies, see e.g. Cho 2021, Garellek et al. 2020 and Winter 2019). 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 consisted of 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 (see Adelman et al. 2018 for the importance of word-initial position in sound symbolism). Because [p] is known to have a sound symbolic effect associated with cutenes (Kumagai 2019, 2022, 2023; 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 account (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 (if any) 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 directly 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 credible/meaningful. However, unlike in a frequentist analysis, we do not have to rely on a strict—but yet arguably arbitrary—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, which reflect 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), which is impossible in frequentist analyses. This feature of Bayesian analysis is particularly important 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 regression analysis.

Moving on to 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 contrast-coded using the backward-reference coding method, in which a particular level is compared against the prior adjacent level, i.e. 3 is compared against 2; 2 is compared against 1; 1 is compared against 0. 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 further details.

2.2 Results

Figure 2 shows the distribution of the proportion of the post-evolution responses 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 the average response 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).

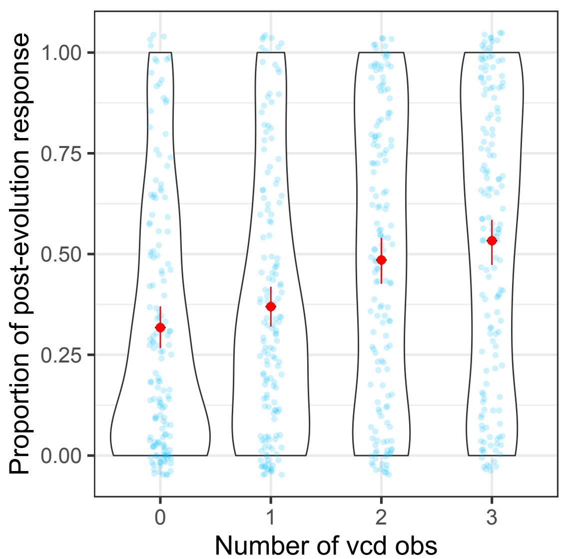

Figure 2: The results of Experiment 1, showing the distribution of the proportion of the post-evolution responses for each number of voiced obstruents contained in the stimuli.

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.[5] The central coefficient estimate of the difference between the zero voiced obstruent condition and the one voiced obstruent condition is 0.35, with its 95% CrI being [-0.09, 0.78]. Although this 95% CrI interval includes zero, the posterior distribution is heavily skewed toward positive values, and about 94% of the posterior samples were positive.

More importantly, the comparison between the two voiced obstruent condition and three voiced obstruent condition shows that the central coefficient estimate for this difference is 0.39 with its 95% CrI being [0.08, 0.72] and the posterior probability supporting this difference is 0.99.

[5]Even those nonce 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.

Finally, the difference between one voiced obstruent condition and the two voiced obstruent condition was also robust, with its central coefficient and 95% CrI being 0.78 and [0.40, 1.17], respectively. Its posterior probability being positive was 1.00.

In short, we observe that each difference between the four conditions was meaningful (although we can be only 94% 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; Kawahara & Kumagai 2019a). 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).

The current result is particularly interesting in the light of the general question regarding 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 constraints that it can count only up to two segments (e.g. Ito & Mester 2003; McCarthy 2003; McCarthy & Prince 1986; Prince & Smolensky 1993/2004), just as Japanese phonology counts only up to two voiced obstruents (Ito & Mester 2003; 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.

An anonymous reviewer has asked if the current results—especially the most crucial difference between the two voiced obstruent condition and the three voiced obstruent condition—could have arisen from the knowledge that the participants had about the existing Pokémon names. This interpretation is unlikely, because there are only 12 existing Pokémon characters whose name contains three voiced obstruents (e.g. [diguda]), and 6 of them are post-evolution characters (the ratio is 0.5, with its binomial 95% confidence interval being [0.25–0.75]).[6] On the other hand, there are 121 characters whose names contain two voiced obstruents, and 81 of them are post-evolution characters (the ratio is 0.67, with its binomial 95% confidence interval being [0.58–0.75]).

[6]This analysis is based on the data gathered by Kawahara et al. (2018), which includes more than 700 characters. Pokémon characters can actually evolve twice in the actual Pokémon world, but we collapsed this distinction between "evolved once" and "evolved twice", because in the experiment, we asked the participants to make a binary "pre-evolution" vs. "post-evolution"' judgment. The confidence interval was calculated using the `binom.confint` function of the `binom` package (Dorai-Raj 2022), whose syntax is available at the osf repository mentioned above.

Thus, there are not many examples from the existing names that support the association between "three voiced obstruents" and "post-evolution" in the first place—the confidence interval for this estimate ([0.25–0.75]) is very large, suggesting that the pattern found in the existing names is not very informative about this association. And if anything, the evidence from the existing names goes in the opposite way from the experimental result: those with two voiced obstruents are more likely to be post-evolution characters than those with three voiced obstruents, although we note that the latter confidence interval is properly contained in the former confidence interval ([0.58–0.75] vs. [0.25–0.75]).

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; Kubozono 1999; Suzuki 1962), and in the context of Pokémon names, they are overrepresented in the names of villainous characters (Hosokawa et al. 2018; 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. 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.[7] After excluding data from those who were not native speakers of Japanese and those who were familiar with research on sound symbolism, the data from 141 native speakers entered into the subsequent statistical analysis. The details of the statistical modeling were identical to those of Experiment I.

[7]We would like to thank Tomoko Monou for her assistance with the participant recruitment for this experiment.

3.2 Results

Figure 3 shows the results of Experiment II. As with Experiment I, we observe a steady increase in the dark-type responses, 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.

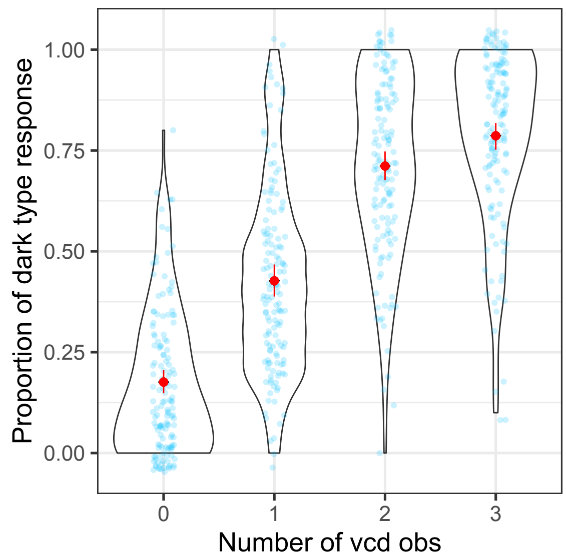

Figure 3: The results of Experiment II. The proportion of the dark-type responses for each voiced obstruent condition.

This effect of voiced obstruents between each level is very robust according to the Bayesian modeling. The difference between the no voiced obstruent condition and one voiced obstruent was very credible, with its central coefficient estimate and its 95% CrI being 1.61 and [0.95, 2.27], respectively. All the posterior samples were positive.

More importantly, the difference between the two voiced obstruent condition and the three voiced obstruent condition was also fairly credible. The central coefficient estimate is 0.59 and its 95% CrI is [-0.03, 1.22]. The posterior probability of this crucial comparison being positive is 0.97. The difference between the one voiced obstruent and two voiced obstruents was also robust (the central coefficient estimate = 1.54, its 95% CrI=[0.89, 2.19], the posterior probability being positive = 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 (Ito & Mester 2003; Kawahara & Kumagai 2023a), which is arguably a general property of phonological constraints at the segmental level in natural languages (McCarthy 2003; McCarthy & Prince 1986; Prince & Smolensky 1993/2004).

The observed difference between the two voiced obstruent condition and the three voiced obstruent condition in this experiment could not have arisen from an analogical inference from existing Pokémon names, because there were no dark type Pokémon characters whose name contains three voiced obstruents.

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 constraints 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. [b] 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 1986, 2003; Hamano 1998; Kubozono 1999; Suzuki 1962).

Nevertheless, it is safer to be conservative and entertain the possibility that effects of different voiced obstruents are governed by different sound symbolic forces. To this end, we took advantage of the sound symbolic connection between [p] and “cuteness” (Kumagai 2019, 2022, 2023), which also manifests itself in the fact that labial sounds, including [p] are, overrepresented in the cute, fairy type Pokémon characters (Hosokawa et al. 2018; 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 consisted of 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 (Perfors 2004; 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, except that in this analysis, we ran, for each chain, 5000 iterations with 4000 warm-ups in order to avoid inappropriate ESS (effective sample size) values and divergent transitions.

4.3 Results

The results are presented in Figure 4, which shows the distribution of the proportions 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: the zero-[p] condition = 0.21; the one-[p] condition = 0.39; the two-[p] condition = 0.47; the three-[p] condition = 0.57.

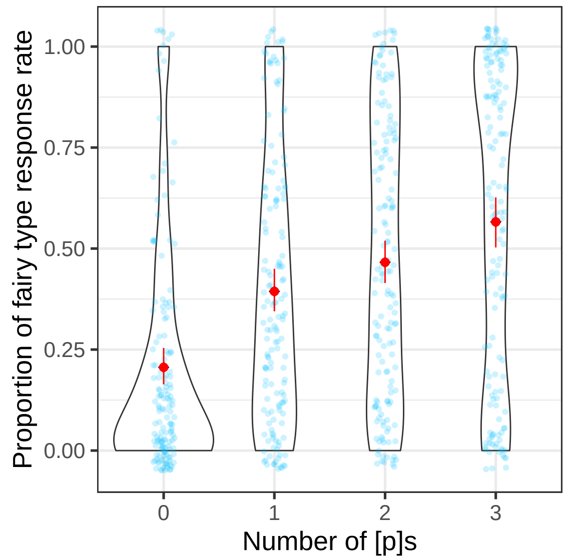

Figure 4: The results of Experiment III. The distribution of the proportion of the fairy type responses for each condition, which contained different numbers of [p]s.

The results of the Bayesian logistic regression show that there is a clear difference between the zero-[p] condition and the one-[p] condition (the central coefficient estimate =1.60, its 95% CrI = [1.06, 2.17]), 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 (the central coefficient estimate = 0.80 with its 95% CrI being [0.30, 1.29], and 99.9% of 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 (central coefficient estimate = 0.47, its 95% CrI [0.03, 1.29] and 98% of the posterior samples support this difference). In short, every addition of [p] in the names reliably increased the fairy-type 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. The difference between the two-[p] condition and the three-[p] condition could not have arisen from the analogical extension from existing names, because there were no fairy characters whose names contain three [p]s.

5 General discussion

5.1 Summary of the results

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 constraints which seems to hold robustly across languages; at least when it comes to the constraints related to 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 segmental, phonological constraints and sound symbolic patterns, at least in terms of their counting capabilities.

5.2 Some alternative interpretations

An anonymous reviewer pointed out an interesting alternative interpretation of the current results, regarding the counting capability of sound symbolism. More specifically, the difference between “2” and “3” that we identified in the three experiments above may instead be the differences between “2” and “all”, given that our “3” condition had three target segments in trisyllabic words (i.e. $[\text{D...D...D}]_{Wd}$, where “D” represents a voiced obstruent). We admit that this is a valid interpretation, and if this was the case, it is comparable to a property that phonological systems routinely exhibit; e.g. a vowel harmony pattern that targets all the vowels within a domain.

A follow-up experiment is necessary to address this alternative interpretation, by using four-syllable words which contained two target sounds and those which contain three target sounds;

schematically, $[\text{D...D...X...X}]_{Wd}$ vs. $[\text{D...D...D...X}]_{Wd}$, where “D” represents a voiced obstruent, and “X” represents a segment other than a voiced obstruent. Then the latter condition would be “3” but not “all”.

Another question that was raised was as follows: in this paper, we made a within-language comparison between the behavior of Lyman’s Law and the sound symbolic effects of voiced obstruents, and showed that only the latter can count up to three. However, while Lyman’s Law is a negative restriction on the presence of multiple voiced obstruents, the current experiment is about how the presence of particular segments positively impact sound symbolic judgments. Thus, the comparison between Kawahara & Kumagai’s (2023a) results and the current experiments may have to do with a difference about a negative restriction vs. a positive influence.

While this interpretation is not impossible, and more studies are warranted to fully address it, we find this explanation not very likely, given that for example, no languages seem to require that reduplicative patterns copy three segments; neither do we find phonological patterns which require three tokens of the same feature/segment. In other words, the “non-counting” thesis is not just about negative restrictions but also holds true about positive presence of particular structures (McCarthy 2003; McCarthy & Prince 1986). Therefore, it is not clear if we can explain the current findings vis-a-vis Kawahara & Kumagai’s (2023a) based on the positive vs. negative nature of the patterns at issue.

5.3 Phonology and sound symbolism again

To the extent that the current experiments have identified a non-trivial difference between phonological systems and sound symbolic systems, should we conclude that they are completely separate systems? We feel that this conclusion may be going too far as well. Recall that as Alderete & Kochetov (2017) and others have argued (Akinbo 2021; Akinbo & Bulkaam 2024; Akita 2020; Klamer 2002; Dingemanse & Thompson 2020; Kumagai 2019, 2023; Jang 2021; Mithun 1982; Monaghan & Roberts 2021), sound symbolic requirements may be able to affect—or at least interact with—phonological patterns.

To the extent that our conclusion is on the right track, then, when 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) (cf. Alderete & Kochetov 2017), 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 highly 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.

An alternative way of reconciling the current results with the view that phonology and sound symbolism interact in non-negligible ways, as suggested by an anonymous reviewer, may be to posit that phonology actually has an iconic component and a non-iconic component, cf. the "co-phonology" approach which posits several phonological sub-systems within a single language (Inkelas et al. 1996; Inkelas & Zoll 2007; Orgun 1996; Sande 2020). Once we accept this assumption, we can further posit that only the former has a counting capability.

Japanese sound symbolic words (i.e. mimetics) are characterized by a set of phonological characteristics that distinguish them from non-iconic words, such as the presence of singleton [p]s and active use of reduplication based on bimoraic feet (Ito & Mester 1995), which is compatible with the idea that phonology can consist of an iconic component and a non-iconic component. This idea that only an iconic component of phonology—to the extent that such a component exists—can count appears compatible with the view advanced by Akinbo (2023), for example, who points out that the number of reduplications correlates with the strengths of their expressive power (see also Kumagai 2023). Thus, this general idea appears to be worth extensive exploration in future studies.

However, one potential concern of this hypothesis is that reduplicative patterns, which can be iconic, as is the case with Japanese mimetics, are predicted to be able to count, but this prediction is incompatible with the general no-counting thesis discussed throughout the present paper. Even if a certain reduplication pattern is expressive, the phonological system does not allow that reduplication pattern to copy three segments (McCarthy & Prince 1986). There also remains a deeper question regarding why only an iconic component has the privilege to count.

All in all, reconciling the increasing number of proposals regarding the similarity between phonological systems and sound symbolic systems on the one hand, and the current finding that these two nevertheless show a distinct characteristic in terms of counting capability on the other, will continue to present an interesting challenge for phonological theorization.

Conflicts of interest

We declare no conflicts of interest.

Availability of data and code

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

References

Adelman, James S., Zachary Estes & Martina Coss. 2018. Emotional sound symbolism: Languages rapidly signal valence via phonemes. *Cognition* 175. 122–130.

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

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

Akinbo, Samuel Kayode. 2023. Reduplication, repetition and sound symbolism in Fungwa. *Phonological Data and Analysis* 5(2). 1–24.

Akita, Kimi. 2015. Sound symbolism. In Jan-Ola Östman & Jef Verschueren (eds.), *Handbook of pragmatics, installment 2015*, Amsterdam and Philadelphia: John Benjamins.

Akita, Kimi. 2020. A typology of depiction marking: The prosody of Japanese ideophones and beyond. *Studies in Language* 45(4). 865–886.

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.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.

Cuskley, Christine. 2013. Mappings between linguistic sound and motion. *The Public Journal of Semiotics* 5(1). 39–62.

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.

Dingemanse, Mark & B. Thompson. 2020. Playful iconicity: structural markedness underlies the relation between funniness and iconicity. *Language and Cognition* 12(1). 203–224.

D'Onofrio, Annette. 2014. Phonetic detail and dimensionality in sound-shape correspondences: Refining the *bouba-kiki* paradigm. *Language and Speech* 57(3). 367–393.

Dorai-Raj, Sundar. 2022. binom: Binomial confidence intervals for several parameterizations. R package.

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).

Gauthier, I & M. J. Tarr. 1997. Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision research* 38(12). 1673–1682.

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.

Haspelmath, Martin. 2014. Comparative syntax. In *The Routledge handbook of syntax*, 490–508. Routledge.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies.* Chicago: The University of Chicago Press.

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.

Hosokawa, Yuta, Naho Atsumi, Ryoko Uno & Kazuko Shinohara. 2018. Evil or not? Sound symbolism in Pokémon and Disney character names. Talk presented at the 1st international conference on Pokémonastics.

Inkelas, Sharon, Orhan Orgun & Cheryl Zoll. 1996. Subregularities as cogrammars: the theoretical status of nonproductive patterns and exceptions in grammars. In *Proceedings of the nels 26*, Ms., University of California, Berkeley.

Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45. 133–172.

Ito, Junko. 1990. Prosodic minimality in Japanese. In Michael Ziolkowski, Manual Noske & Karen

Deaton (eds.), *Proceedings of Chicago Linguistic Society: Parasession on the syllable in phonetics and phonology*, 213–239. Chicago: Chicago Linguistic Society.

Ito, Junko & Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17. 49–73.

Ito, Junko & Armin Mester. 1995. Japanese phonology. In John Goldsmith (ed.), *The handbook of phonological theory*, 817–838. Oxford: Blackwell.

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.

Ito, Junko & Armin Mester. 2013. Prosodic subcategories in Japanese. *Lingua* 124. 20–40.

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 & 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. 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.

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

Klamer, Marian. 2002. Semantically motivated lexical patterns: A study of Dutch and Kambera expressives. *Language* 78(2). 258–286.

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.

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

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. 2023. EXPRESS[P] in expressive phonology: Analysis of a nicknaming pattern using 'princess' in Japanese. *Phonology* 40. 267–290.

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

McCarthy, John J. 2003. OT constraints are categorical. *Phonology* 20(1). 75–138.

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

McPherson, Laura & Bruce Hayes. 2016. Relating application frequency to morphological structure: The case of Tommo So vowel harmony. *Phonology* 33. 125–167.

Mester, Armin. 1990. Patterns of truncation. *Linguistic Inquiry* 21. 475–485.

Mithun, Marianne. 1982. The synchronic and diachronic behavior of plops, squeaks, croaks, sighs, and moans. *International Journal of American Linguistics* 48(1). 49–58.

Monaghan, Padraic & Seán G. Roberts. 2021. Iconicity and diachronic language change. *Cognitive Science* 45(4). e12968.

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

Nelson, Diane & Ida Toivonen. 2000. Counting and the grammar: Case and numerals in Inari Sami. In Diane Nelson & Paul Foulkes (eds.), *Leeds working papers in linguistics 8*, 179–182.

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.

Orgun, C. Orhan. 1996. *Sign-based morphology and phonology: With special attention to Optimality Theory*: University of California, Berkeley Doctoral dissertation.

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

Perfors, Amy. 2004. What's in a name?: The effect of sound symbolism on perception of facial attractiveness. In K. Forbus, D. Gentner & T. Regier (eds.), *26th annual conference of the cognitive science society*, Chicago, IL: Psychology Press.

Perniss, Pamela, Robin L. Thompson & Gabriella Vigiliocco. 2010. Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology* doi:10.3389/fpsyg.2010.00227.

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

Priestly, Tom. 1994. On levels of analysis of sound symbolism in poetry, with an application to Russian poetry. In Leane Hinton, Johanna Nichols & John Ohala (eds.), *Sound symbolism*, 237–238. Cambridge: Cambridge University Press.

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

Sande, Hannah. 2020. Cophonologies by Ph(r)ase. *Natural Language and Linguistic Theory* 38. 1211–1261.

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. 2017. Constraint conjunction in weighted probabilistic grammar. *Phonology* 34(2). 243–268.

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. Cross-linguistic 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.

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.

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.

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.

Walker, Rachel. 2001. Round licensing, harmony, and bisyllabic triggers in Altaic. *Natural Language and Linguistic Theory* 19. 827–878.

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.