

Exploring the nature of cumulativity through sound symbolism: Experimental studies of Pokémonastics with English speakers

Shigeto Kawahara & Canaan Breiss
Keio University & UCLA

Abstract

Recent formal phonological analyses as well as a body of laboratory phonology experiments have shown that phonological knowledge is inherently cumulative. These results have opened up a set of new questions regarding *how* cumulativity operates in linguistic systems. We addressed these questions by bringing in a new type of evidence: data from sound symbolism. Two experiments on the sound symbolism in Pokémon names with native speakers of English show that (1) more than two factors (e.g. vowel quality, consonant voicing and name length) can cumulatively interact, (2) when two different factors are present (ganging cumulativity), their interaction is linearly cumulative, (3) when multiple instances of the same factor are present (counting cumulativity), their effects are sub-linearly cumulative, and (4) counting cumulativity and ganging-up cumulativity can coexist in a single system. While the current experiments may not provide a final answer to the questions on how cumulativity functions in linguistic patterns in general, our experiments can be understood as cases studies which shed light on how various linguistic factors interact when speakers make a linguistically informed judgment on sound symbolic patterns. The current case studies also suggest that it may be informative to study sound symbolism to explore the nature of linguistic patterns.

Keywords: cumulativity, sound symbolism, Pokémon names, voicing, the iconicity of quantity

1 Introduction

1.1 Lexicographic optimization vs. numerical optimization

When speakers make a decision about language (what to say, how to say it, etc.), they usually have available various pieces of evidence that bear upon that decision. More often than not, pieces of evidence differ in their cogency, and can conflict with each other. How to optimally resolve this sort of situation, implemented in terms of constraint conflict and interaction, is at the heart of constraint-based theories of grammar. In such cases, there are two major strategies to resolve the conflict: optimization based on numeric ordering vs. optimization based on lexicographic ordering (Gigerenzer and Gaissmaier 2011). The first decision strategy is familiar from standard statistical regression models, which take into account all available evidence. Different pieces of evidence have different weights, and the ultimate decision is based on the weighted sum of each constraint’s importance. The second decision strategy is a “fast-and-frugal” decision model, in which a decision is made based solely on the most important piece of evidence that bears upon that choice. For a non-linguistic example of this type of decision-making strategy, Wübben and Wagenenheimer (2008), cited by Gigerenzer and Gaissmaier (2011), show that in order to predict whether or not a certain customer will return again, experienced managers use a very simple fast-and-frugal decision rule of this type. Specifically, “[i]f a customer has not purchased within a certain number of months (the hiatus), the customer is classified as inactive; otherwise, the customer is classified as active” (Gigerenzer and Gaissmaier 2011: p.455). This simple decision-making strategy turns out to be at least as good as a more “sophisticated” Poisson regression model. An unintuitive but important feature of this fast-and-frugal approach is that it entirely disregards at least half (often more) of available evidence (Gigerenzer and Gaissmaier 2011), the exact opposite of numerical optimization approaches.

This distinction between numeric ordering vs. lexicographic ordering has a striking corollary in linguistic theory, albeit in a slightly different disguise.¹ The debate at issue is the comparison between Optimality Theory (OT) (Prince and Smolensky 1993/2004) and other related theories which use numerically-weighted constraints, such as Harmonic Grammar (HG) (Legendre, Miyata, and Smolensky 1990a; Legendre, Miyata, and Smolensky 1990b; Pater 2009; Pater 2016). These two theories share the same general optimization architecture—both of them can be conceived of as devices that can map input representations to output representations. Common to both is the tripartite structure of GEN, CON and EVAL: GEN

¹Few linguistic studies frame this debate in terms of the perspective of general decision making processes, the latter of which is arguably domain-general (though see Prince and Smolensky 1993/2004; Prince and Smolensky 1997; Tesar 2007), but there seems to be a clear parallel between the debate in linguistics and the one in the decision-making literature.

generates a set of candidates that are under consideration, while CON is a set of (possibly) universal constraints. The primary difference between OT and HG lies in how EVAL operates. In OT, the constraints are *strictly ranked*, yielding outcomes which embody the fast-and-frugal approach. The decision between two candidates is solely made by the highest ranked constraint which distinguishes them, and the violation profiles of lower ranked constraints are ignored. On the other hand, in HG, constraints bear numerical weights, and the competition between two candidates is resolved by taking into consideration all the relevant constraints and their weights as sources of evidence. The candidate with the lowest weighted sum of constraint violations is chosen as optimal.

In recent years, there has been an active debate regarding which model—lexicographic OT vs. numeric HG—offers a better device to model speakers’ linguistic knowledge (Anttila and Magri 2018; Anttila, Borgeson, and Magri 2019; Boersma and Hayes 2001; Boersma and Pater 2016; Breiss 2019; Breiss and Albright 2020; Hayes 2017; Hayes 2020; Jäger and Rosenbach 2006; Jäger 2007; O’Hara 2017; Pater 2009; Pizzo 2015; Potts et al. 2010; Prince and Smolensky 1993/2004; Prince and Smolensky 1997; Smith and Pater 2017; Tesar 2007; Zuraw and Hayes 2017 among many others). The dominant framework in phonological theories since the early ’90s was OT, but recent studies, many inspired by Pater (2009), have shown that at least some linguistic patterns are better modelled with HG because they demonstrate cases of cumulative constraint interactions, which is predicted only under the latter theory.² This body of studies has opened up a new set of questions—to the extent that linguistic patterns show cumulativity, how precisely does this cumulativity manifest itself in linguistic patterns? The current experiments aim to contribute to this debate by bringing in a new type of evidence: data from sound symbolism. Before laying out the specific purposes of the experiments, however, we first expand on the difference between the two competing grammatical frameworks to explicate the questions that we address experimentally.

1.2 Cumulativity

The key difference between OT and HG is their predictions regarding *cumulativity*.³ In the strict-ranking OT, constraint effects are predicted not to be cumulative because candidate competition is decided solely by the highest relevant constraint. On the other hand, in HG, no constraints are ever ignored outright, although some constraints may receive very low weights so that their effects may be essentially invisible, dwarfed by the weights of other more important constraints.

The question regarding whether the effects of constraints are cumulative or not has been

²Historically speaking, the Maximum Entropy formulation of HG predates OT (e.g. Smolensky 1986).

³Discussion in this section heavily draws upon Breiss (2019) and Kawahara (2020a).

actively discussed since the inception of OT (Prince and Smolensky 1993/2004). One tenet of OT is *strict domination*, which states that constraint ranking is strict, predicting non-cumulative effect. Many subsequent studies in OT followed this spirit in assuming that constraints are ranked with strict domination instead of being numerically weighted.

However even at during the early days of research in OT, it was pointed out that there are apparent cases in which two lower-constraints can “gang-up” to defeat a higher-ranked constraint (e.g. Alderete 1997, Guy and Boberg 1994 and Kirchner 1996). Such cases have generally been handled by positing a constraint that is violated if and only if these *two* lower-ranked constraints are violated. This conjoined constraint is created by the mechanism known as “local conjunction” (Crowhurst 2011; Smolensky 1995; Smolensky 1997).⁴

In the LabPhon research tradition, partly as a reaction to OT, it has been pointed out that phonological patterns can be deeply cumulative. The cumulative nature of linguistic patterns seems particularly clear in wellformedness judgments of surface phonotactic patterns. For instance, Coleman and Pierrehumbert (1997) show that a word bearing a combination of several “minor” phonotactic violations (e.g. *spleitisak*) is judged to be worse than a word with one “major” phonotactic violation (e.g. *mrupation*), an observation recently reiterated by Pierrehumbert (2020). In the conclusion section, they state:

According to Optimality Theory, “impossible words” are those in which a constraint is so strong that a null parse is preferred to a parse in which the constraint is violated. This means that impossible words are those which are egregious according to a single constraint.

However, the probability of the worst part is not the best score of acceptability: the log probability of the whole word is a better measure, a result at odds with standard generative phonology and OT alike (p. 55).

Hay, Pierrehumbert, and Beckman (2003) further generalized this finding and concluded that “well-formedness reflects the cumulative effect of the likelihood of the subparts (p.13).” Bailey and Hahn (2001) show that both phonotactic probabilities and neighborhood density cumulatively affect word-likeness judgment patterns by English speakers. Pizzo (2015) found that phonotactic violations in onsets and those in codas cumulatively add up (e.g. *tlavb* is judged to be worse than *tlag* and *plavb*). Generally speaking, wellformedness judgment patterns of surface phonotactics are stochastic and cumulative, and we know of no

⁴We should bear in mind, however, that Smolensky did not originally propose local conjunction as a mechanism to generate new constraints based on the already existing constraints. He instead proposed local conjunction as a way to understand the internal structure of constraints that should be admitted in CON. In practice, however, subsequent researchers have used local conjunction as a mechanism to generate new constraints, usually to handle these types of apparently-cumulative cases.

controlled studies which showed that phonotactic judgment patterns are non-cumulative, solely determined by the most important constraint, as predicted by OT.⁵

Cumulative phonotactic patterns have been found in other laboratory phonology experiments as well. Rose and King (2007) demonstrate that speakers of Amharic and Chaha show speech errors more frequently when a target word violates two phonotactic restrictions than when there is only one violation. Kawahara (2011) examined whether phonological devoicing in Japanese is judged to be natural when it is caused by a restriction against a voiced geminate (*DD) and a restriction against two voiced obstruents (a.k.a. Lyman’s Law). The results of this phonological judgment experiment suggest that the two phonological constraints—*DD and Lyman’s Law—cumulatively make devoicing more natural. Breiss (2019) and Breiss and Albright (2020) report a series of artificial language learning experiments to show that cumulativity is the default strategy in phonotactic learning for English speakers.

There is an accumulating body of evidence from phonological analyses that some phonological alternations work in a similar, cumulative fashion as well. For example, the probabilities of *t/d*-deletion in English are cumulatively affected by different types of featural OCP constraints (Guy and Boberg 1994; Guy and Boberg 1997). In Japanese, singleton [p] and voiced geminates are both banned from native phonology but they are tolerated in loanwords, as long as they appear independently (Ito and Mester 1995; Ito and Mester 1999). However, devoicing of geminates occurs when voiced geminates co-occur with [p] within the same word, instantiating cumulative interaction of the two constraints—*[p] and *DD—both of which are independently motivated in the native phonology (Fukazawa et al. 2015; Kawahara and Sano 2016). Kim (2019) shows that Rendaku voicing in Japanese compound formation (Vance 2015) is not blocked by a single instance of a nasal segment, but it is blocked by two nasal segments, arguing that Rendaku can create two voiced consonants in a morpheme but not three. Blust (2012) reports cases from Austronesian and Australian languages in which two instances of marked segments—including geminates and prenasalized segments—are avoided via dissimilation. Smith and Pater (2017) show that the distribution of schwas in French is affected by both the number of surrounding consonants and its prosodic position within a word, and that these effects interact cumulatively. Green and Davis (2014) found

⁵One could take the position that phonological grammar is not responsible for modelling phonotactic judgment patterns (see de Lacy 2009 for a related discussion), potentially relegating these experimental results as a matter of “performance.” However, we know of no good theory of phonological performance that can account for such gradient and cumulative patterns. One general problem of this position is that a categorical constraint that can trigger a phonological alternation (e.g. OCP constraints: Yip 1988) in one language can gradiently affect surface phonotactic patterns in other languages (e.g. Coetzee and Pater 2008; Frisch, Pierrehumbert, and Broe 2004; Kawahara and Sano 2014). Another problem is that gradient judgment patterns are demonstrably best modelled using distinctive features (Albright 2009), a toolkit that is typically used to model competence.

various cases of cumulative interactions among restrictions on complex syllable structures in Colloquial Bambara. Zuraw and Hayes (2017) offer corpus-based analyses of Tagalog, French and Hungarian, showing cumulative action of different types of alternation triggers in all three languages.

This increasing body of work appears to have established that cumulativity is a general property of natural languages. At the same time, it has opened up several new questions regarding how cumulativity works in linguistic systems, including:

- (1) Can more than two factors interact cumulatively?
- (2) When two different factors are present (ganging cumulativity: Jäger and Rosenbach 2006), are the results linearly cumulative or sub/super-linear?
- (3) When multiple instances of the same factor are present (counting cumulativity: Jäger and Rosenbach 2006), are the effects linearly cumulative or sub/super-linear?
- (4) Can counting cumulativity and ganging-up cumulativity coexist in a single system?

Regarding the first question, many if not all previous studies have explored the interaction of two phonological factors, and it is only poorly understood whether, for example, three factors can interact cumulatively (though cf. Kim 2019). In the context of OT, we may understand this question as whether or not we should allow local conjunction of more than two constraints, or relatedly, whether local conjunction can be recursive (Ito and Mester 2003; Smolensky and Legendre 2006). Ito and Mester (2003) carefully entertain the possibility of introducing “recursive local conjunction,” which would essentially make it possible for any number of factors to cumulatively add up. They are cautious about the consequence of positing such a mechanism, stating that:

With local conjunction as a recursive operation, ternary (and higher) conjunction...[is] formally derivable... No convincing evidence has been found so far that [a ternary conjoined constraint] is ever linguistically operative separate from [a binary conjoined constraint], which tends to support the old idea in generative linguistics (cf. syntactic movement theory) that the genuine contrast in grammars is not “1 vs. 2 vs. 3 vs. 4 vs...”, but “1 vs. greater than 1.” (p. 265)

The second question can be understood as follows: can the probability of a candidate that violates two constraints be predicted from the probabilities of the two candidates that violate each of these two constraints alone (Breiss and Albright 2020)? This question is related to the issue of whether we need local conjunction in addition to frameworks such as HG which predict cumulative phonological effects by default (Shih 2017). As implied above, both local conjunction and numeric optimization have been proposed as means to account

for cumulative patterns, and therefore a natural expectation is that one of these suffices; however, when the cumulative effects are substantially super-linear, that may require local conjunction of constraints even in a framework with weighted constraints. Shih (2017) argues that this is actually the case in a parasitic tone harmony pattern in Dioula d’Odienné; thus, this question bears examining in the experiments carried out here.

The third question, related to the second, concerns the cases in which there are multiple instances of a particular structure—here, we call that structure [A]. If the effects of [A] (on phonotactic acceptability, the probability of undergoing a phonological alternation, etc.) increase as the function of the number of [A], that would instantiate a case of counting cumulativity (Jäger and Rosenbach 2006). When counting cumulativity holds, are the outcomes linear, sub-linear or super-linear (Breiss and Albright 2020)? This question is related to the general issue of what the linking function is between the number of constraint violations (or harmony scores) and the probabilities of the candidate that violates that constraint (Hayes 2020; McPherson and Hayes 2016).

Finally, it remains underexplored whether counting cumulativity and ganging-up cumulativity can co-exist in a single system. As far as we know, there are only a few studies which directly addressed this question (McPherson and Hayes 2016; Zuraw and Hayes 2017; Zuraw and Hayes 2017).

To summarize, the studies reviewed in this section have shown that linguistic patterns, at least in some contexts, show cumulative effects; these studies opened up several questions regarding *how* cumulativity works in naturally languages. In this theoretical context, we shed light on four specific questions concerning the nature of cumulativity from a hitherto underexplored perspective; namely, sound symbolism, systematic associations between sounds and meanings (e.g. Hinton, Nichols, and Ohala 1994). Concretely, we report two Pokémonastics experiments, which use Pokémon names to study sound symbolic patterns in natural languages (Shih et al. 2019). An assumption that underlies our studies is that there are non-trivial parallels between sound symbolic patterns and phonological patterns, and hence studying one of these domains can shed light on the nature of the other (Kawahara, Katsuda, and Kumagai 2019; Kawahara 2020a; Shih 2020).

1.3 Pokémonastics

The main purpose of the current paper is to address the nature of cumulativity in natural languages; however, the current studies can also be understood as case studies of Pokémonastics, a research paradigm in which researchers explore the nature of sound symbolism in human languages (Kawahara, Noto, and Kumagai 2018; Shih et al. 2019), where sound symbolism is a cover term that refers to systematic connections between sounds and meanings (Hinton,

Nichols, and Ohala 1994).

Shih et al. (2019) summarize various research advantages of using Pokémon names to study sound symbolism. One clear advantage is the fact that Pokémon characters (of which there are more than 800 characters now) are specified for various attributes, such as weight, height, strengths, and evolution levels. This aspect of the Pokémon universe makes it possible to conduct a quantitative study of sound symbolic patterns in an ecologically realistic setting.

Another distinct advantage is the fact that the set of denotations that are assigned a name is fixed across all languages in the Pokémon universe. This is not quite the case in natural languages, because languages differ in terms of which real world object to name, and this is the case even at the level of basic vocabulary. For instance, Japanese needs to distinguish “older brother” and “younger brother” and does not have a word that corresponds to the English word “brother.” Neither does it have a gender neutral term “siblings.” This sort of cross-linguistic difference makes cross-linguistic studies of sound symbolism using existing words difficult (although it is not impossible: e.g. Wichmann, Holman, and Brown 2010). The Pokémon universe therefore has a potential to provide a well-controlled test ground for cross-linguistic comparisons of sound symbolism. In this spirit, Shih et al. (2019) report a cross-linguistic study of Pokémon names in Cantonese, English, Japanese, Korean, Mandarin and Russian. Even if Pokémon names are not available in a particular language, one can run an elicitation study to explore how Pokémon creatures would be named in that language. Godoy et al. (2019) report a study of this sort targeting speakers of Brazilian Portuguese.

In this general research paradigm, it is also possible to conduct experiments to explore which Pokémon properties are symbolically expressed how in what languages. For example, how evolution levels are symbolically expressed have been explored in Japanese (Kawahara and Kumagai 2019), Brazilian Portuguese (Godoy et al. 2019) as well as in English (Kawahara and Moore to appear). Evolved Pokémon characters are generally larger, heavier, and stronger. These studies have revealed interesting cross-linguistic commonalities as well as differences. For instance, in all these three languages, nonce names with [a] tend to be judged more suitable for post-evolution Pokémon characters than nonce names with [i]. The effects of voiced obstruents are tangible in all these three languages—larger, post-evolution characters tend to be associated with names with voiced obstruents—but the effect size is evidently larger for Japanese than for English and Brazilian Portuguese.

The majority of the experimental studies, however, is still limited to those targeting Japanese speakers. In order for the Pokémonastic paradigm to provide a useful resource for cross-linguistic studies of sound symbolism in general, more studies targeting languages other than Japanese are hoped for. This is one gap that the current experiments are intended to address.

2 Experiment 1

2.1 Introduction

In order to address the theoretical and empirical issues outlined above, the experiment manipulated three factors: (i) vowel quality ([a] vs. [i]), (ii) voicing of consonants (voiced vs. voiceless), and (iii) length (short vs. long). The main purpose of the experiment was to examine whether these three factors interact cumulatively or not. This design also allows us to address another question regarding the nature of cumulativeness—whether the effects are linearly cumulative or sub/super-linear.

In addition to addressing the nature of cumulativeness in sound symbolism, each of the sound symbolic associations that is tested in the experiment has a plausible phonetic basis. The experiment thus serves an additional purpose of testing the robustness of sound symbolic effects that are grounded in our speech behavior, even if the sound symbolic effects at issue are not evidently observable in the Pokémon lexicon.

The first factor of vowel quality ([a] vs. [i]) instantiates a well-known sound symbolic effect, in which the vowel [a] is associated with large images, whereas the vowel [i] is associated with small images (Jespersen 1922; Sapir 1929). One plausible phonetic basis of this sound symbolic principle is the difference in oral aperture size: the mouth is much wider open for [a] than for [i], and this difference may be projected onto the different size ratings (Jespersen 1922; Sapir 1929). Another plausible explanation is their difference in F2: [i] has very high F2, whereas [a] has low F2. Given that formant frequency is inversely proportional to the size of a resonating chamber, sounds with high frequency energy are generally associated with small images (Ohala 1983b; Ohala 1994).

These sound symbolic associations ([a] = big, [i] = small) are known to hold for English speakers in some classic experiments on sound symbolism (Newman 1933; Sapir 1929). Within Pokémonistics, a previous experimental experiment has shown that English speakers indeed associate names with [a] with post-evolution names and names with [i] with pre-evolution names (Kawahara and Moore to appear), even though these sound symbolic associations do not seem to be inferrable from the existing English Pokémon lexicon (Shih et al. 2019). A similar sound symbolic effect is observed in other Pokémonistics experiments targeting Japanese speakers (Kumagai and Kawahara 2019) and Brazilian Portuguese speakers (Godoy et al. 2019).

The second factor that is manipulated in the experiment is the effects of consonant voicing. Newman (1933) found that English speakers tend to judge nonce words with voiced obstruents to be larger than those with voiceless obstruents. Articulatorily speaking, the production of voice obstruents requires expansion of supralaryngeal cavity (Ohala 1983a;

Proctor, Shadle, and Iskarous 2010; Westbury 1983)—this expansion occurs because it is necessarily to keep the intraoral airpressure sufficiently low with respect to the subglottal airpressure level in order to sustain vocal fold vibrations (Ohala 1983a). Acoustically speaking, voiced obstruents involve low frequency energy in three respects: (1) they are characterized by low f_0 as well as low F1 in surrounding vowels (Kingston and Diehl 1994), (2) burst energies are lower for voiced obstruents than for voiceless obstruents (Chodroff and Wilson 2014), and (3) at least word-internally, voiced obstruents are characterized by low frequency energy which reflect vocal fold vibration (Stevens and Blumstein 1981). These low frequency properties, which are demonstrably integrated into one perceptual property (Kingston and Diehl 1995; Kingston, Diehl, et al. 2008), can be mapped onto large images, because of the general inverse relationship between the size of a resonator and its frequency (Ohala 1983b; Ohala 1994).

Shih et al. (2019) did not identify a correlation between evolution levels and the number of voiced obstruents contained in their names in the set of existing English Pokémon names. The first Pokémonastics experiment by Kawahara and Kumagai (2019) showed that English speakers tend to associate nonce names with voiced obstruents with post-evolution characters, whereas they tend to associate nonce names with voiceless obstruents with pre-evolution characters, although the difference between groups in that experiment was very small. The primary target of this experiment, moreover, was Japanese speakers, and hence the stimuli were Japanese-sounding words, and could have sounded unnatural to English speakers. A follow-up experiment by Kawahara and Moore (to appear) identified a similar trend for English speakers to associate voiced obstruents with larger post-evolution characters, although the effect of voicing was not statistically significant in that study. The current experiment therefore addresses, with a larger number of test items and participants, whether we can identify the effects of consonant voicing on the judgment of evolution in Pokémon names.

The third factor, phonological length, was first identified as an active sound symbolic principle in the existing set of Japanese Pokémon names (Kawahara, Noto, and Kumagai 2018). They found that those Pokémon characters with longer names are generally larger, heavier, stronger and more evolved. They attribute this observation to a previously known sound symbolic principle, “the iconicity of quantity,” in which larger quantity is expressed via phonological length (Haiman 1980; Haiman 1985). A follow-up cross-linguistic study of existing Pokémon names by Shih et al. (2019) targeting Cantonese, English, Japanese, Korean, Mandarin and Russian shows that the iconicity of quantity is the sound symbolic principle that most robustly holds across these languages, including English. Two experimental studies confirmed the productivity of this principle using nonce names with English speakers (Kawahara and Kumagai 2019; Kawahara and Moore to appear).

To recap, the current experiment manipulated three phonological dimensions (vowel, consonant voicing, and length) to examine whether each of this factor impacts the judgment of evolvedness in Pokémon names. Perhaps more importantly, to the extent that these factors impact the judgment of evolvedness, an ensuing question is whether they do so cumulatively, and if so, how.

2.2 Methods

2.2.1 Stimuli

Experiment 1 had three factors which were fully crossed; six items were included in each cell. The full list of the stimuli is given in Table 1. The stimuli either had two voiceless stops or two voiced stops in onset; two items had labial stops, two items had coronal stops, and two items had dorsal stops. The short forms were CVC.CVC, where coda consonants were sonorants. Long forms are of the form CrVC.CIVC; the first complex onset was created using an additional [r], because this is the only consonant that can form a complex cluster with any preceding stop in English. The second complex onset was created using [l] in order to avoid two occurrences of [r] within the same word, which can be perceptually confusing (Hall 2012). The vowels were either [i] or [a].

Table 1: The list of stimuli for Experiment 1

Voiceless			
[i]-[i] short	[i]-[i] long	[a]-[a] short	[a]-[a] long
pinkin	prinklin	pankan	pranklan
pintil	prinslim	pantal	pranslam
tinsin	trinslin	tansan	translan
timpim	trimplim	tampam	tramplam
kimpin	krimplin	kampan	kramplan
kintil	krinslin	kantal	kranslan
Voiced			
[i]-[i] short	[i]-[i] long	[a]-[a] short	[a]-[a] long
bingin	bringlin	bangan	branglan
bindil	brinzlim	bandal	branslam
dinzin	drinzlin	danzan	dranzlan
dimbim	drimblim	dambam	dramblam
gimbin	grimblin	gamban	gramblan
gindil	grinzlin	gandal	granzlan

2.2.2 Participants

The experiment was distributed online via SurveyMonkey. The responses were collected using the “buy response” function of SurveyMonkey. A total of 150 participants, who passed all the inclusion criteria (see below), completed the experiment.

2.2.3 Procedure

The first page of the experiment was a consent form, which was approved by the first author’s institute. The second page presented the qualification questions, and only those who fulfilled all four of the following conditions were allowed to proceed: (1) they were a native speaker of English, (2) they were familiar with Pokémon, (3) they were not already familiar with sound symbolism, (4) and they had not participated in a Poémonastics experiment before.

In the instructions, the participants were told that the experiment is about how they would name new Pokémon characters. They were also told that there are two aspects of Pokémon that are crucial: (1) Pokémon characters undergo evolution, and when they do, they are called by a different name, (2) when Pokémon characters undergo evolution, they generally become larger, heavier, and stronger. The participants were provided with an example pair that illustrates the difference between pre-evolution character and post-evolution character using a pair of non-existing Pokémon characters, shown in Figure 1.



Figure 1: Pictures used to illustrate pre-evolution vs. post-evolution Pokémon characters drawn by a digital artist *toto-mame*. They are generally judged to be authentic by Pokémon practitioners, and were used in the experiment with the permission from the artist.

Within each trial, the participants were given one nonce name and asked to judge whether that name is better for a pre-evolution character or a post-evolution character, i.e., the task was to make a binary decision. The stimuli were presented in the English orthography,

although they are asked to read each stimulus in their head before making their responses.⁶ They were asked to base their decision on their intuition, without thinking too much about “right” or “wrong” answers. The order of the stimuli was randomized per participants.

2.2.4 Statistical analyses

Results for this experiment, as well as Experiment 2 below, were analyzed using Bayesian hierarchical logistic regression using the `brms` R package (Bürkner 2017), with evolvedness (0 = pre-evolution, 1 = post-evolution) as the dependent variable, and binary fixed effects of length, vowel, and voicing, with all two- and three-way interactions, plus random intercepts for subject and item, with random slopes of all fixed effects and interactions by subjects.⁷ Bayesian models yield a distribution of possible values for each parameter of interest, which are interpreted by examining the middle 95% of these values, called the 95% Credible Interval (abbreviated 95% CI, followed by bracketed upper and lower bounds). We can interpret these values directly as our degree of belief in the estimate of the role of the factor in explaining the data (Vasishth et al. 2018), with positive coefficients (β ’s) indicating that the factor was associated with post-evolution Pokémon. In addition to allowing us to estimate the evidence in favor of a specific effect—rather than simply rejecting a point null hypothesis—taking a Bayesian approach to analysis allowed us to fit the complex model structure with multiple interaction terms justified by the experimental design without convergence issues.

2.3 Results

The results are graphically represented in Figure 2. We observe that each phonological factor affected the judgment of evolvedness in the expected direction. Long names were more likely to be associated with post-evolution characters than short names (left panels vs. right panels). Names with [a] are more likely to be associated with post-evolution characters than names with [i] (top panels vs. bottom panels). The names with voiced obstruents were more likely to be associated with post-evolution characters than names with voiceless consonants (comparisons within each panel).

The results of the hierarchical Bayesian mixed effects model show that length (long vs. short, $\beta = 0.90$, 95% CI [0.53, 1.26]), vowel ([a] vs. [i], $\beta = 1.85$, 95% CI [1.41, 2.29]),

⁶Since Pokémon names are often communicated in written forms, and since the previous Pokémonastics experiments used orthographic stimuli, the current experiment followed that methodology (Kawahara and Kumagai 2019; Kawahara and Moore to appear). It is possible that orthography has some effects on sound symbolic patterns (Cuskley, Simner, and Kirby 2017), but it has also been shown that sound symbolism holds beyond the influences of orthography (Sidorov, Pexman, and Saint-Aubin 2016).

⁷We ran four chains of 2,000 iterations each, retaining for analysis samples from the second 1,000 from each chain. All \hat{R} values were between 1 and 1.01, indicating that the chains mixed successfully; weakly-informative, “default” priors were used.

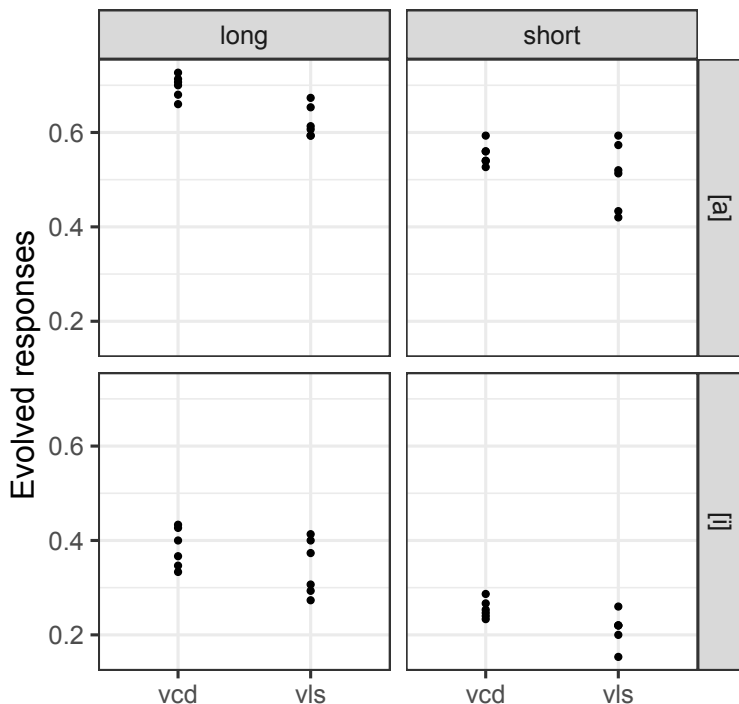


Figure 2: The results of Experiment 1. Each dot represents average “evolved response” for one item.

and voicing (voiced vs. voiceless, $\beta = 0.60$, 95% CI [0.28, 0.92]) all meaningfully predicted participants’ responses. This result indicates that each linguistic factor cumulatively affected the judgment of evolution status for English speakers. The credible intervals for all of the interaction terms included 0, indicating that at least for the case at hand, the cumulativeness was linear in nature.

2.4 Discussion

The current results first of all show that three phonological factors can independently impact the judgment of evolvedness in naming new Pokémon characters. Further, the fact that each factor exerts its own effect regardless of the presence of other factors is evidence that cumulativeness of three factors is possible in judgment concerning sound symbolism. We submit that this is a novel result.

The current results show that none of the interactions were substantial for the case at hand. When two or three factors are relevant, the probabilities of the outcomes can be predicted solely based on the joint probabilities of each factor at play (cf. Shih 2017). For the case at hand, then, the cumulative effects were linear. We reiterate that with the

Bayesian analysis, we can interpret the results as suggesting that the interaction terms are indeed not meaningful, instead of simply failing to reject a null hypothesis.

Finally, as discussed in section 2.1, the sound symbolic effects of vowels and voiced obstruents on evolution levels are not observed in the existing English Pokémon lexicon (Shih et al. 2019), while these sound symbolic effects are observed in other experimental settings (Kawahara and Moore to appear; Kumagai and Kawahara 2019). The current results lend further support to the thesis that English speakers are able to apply these sound symbolic associations to new Pokémon names, even when they are not evidently apparent in the existing Pokémon patterns.

3 Experiment 2

In order to further our understanding of the nature of cumulativity in sound symbolic patterns, Experiment 2 tested counting cumulativity by varying name lengths in three degrees (short vs. medium vs. long). To test whether counting cumulativity and ganging-up cumulativity can co-exist, this factor was crossed with the binary vowel quality measure from Experiment 1.

3.1 Methods

The experimental procedures were almost identical to those of Experiment 1, so we only highlight the important differences.

3.1.1 Stimuli

The list of the stimuli in Experiment 2 is shown in Table 2. Short forms are disyllabic CV.CV words, and the vowels are either [i] or [a]; the first consonants were voiceless stops (three items for [p], [t], [k] each), and the second consonants were sonorants. Medium forms were of the form CVC.CVC, in which onset consonants were voiceless stops and coda consonants were sonorants, which guaranteed sonority fall across the syllable boundary (Vennemann 1988). Long forms were of the form CCVC.CCVC. Each consonant cluster in onset is an attested sequence in the English phonotactics, and there was a sonority fall across the syllable boundaries.

3.1.2 Participants

A total of 147 native speakers of English participated in this experiment. They were recruited online from two universities in the United States, as well as from the “Psychological research

Table 2: The list of stimuli for Experiment 2

[i] short	[i] medium	[i] long
pini	pinkin	prinklin
pimi	pimpil	primplim
pili	piltim	prilslim
tini	tinsin	trinkrin
timi	timpim	trimplim
tili	tilpil	trilspil
kimi	kimpin	krimplin
kini	kintil	krinslin
kili	kiltim	kriltrim
[a] short	[a] medium	[a] long
pana	pankan	pranklan
pama	pampal	pramplam
pala	paltam	pralslam
tana	tansan	trankran
tama	tampam	tramplam
tala	talpal	tralspal
kama	kampan	kramplan
kana	kantal	kranslan
kala	kaltam	kraltram

on the net” website.⁸ Nine participants were excluded from the subsequent analysis because they did not satisfy all of the inclusion criteria: (1) they are a native speaker of English, (2) they are familiar with Pokémon, (3) they are not familiar with sound symbolism, (4) and they have not participated in a Poémonastics experiment before.

3.1.3 Statistical analyses

Taking the theoretical quantity of length as a continuous scale, we coded the length factor numerically as 1, 2, and 3. Other aspects of the analysis were identical to those of Experiment 1, although we report an additional analysis to examine the question of whether the counting cumulativity pattern is linear or sub/super-linear.

⁸<https://psych.hanover.edu/research/exponnet.html>

3.2 Results

3.2.1 General results

Figure 3 illustrates the general results by presenting average “evolved response” for each item. We observe that as the names get longer, they were more likely to be judged as names for post-evolution characters. Within each panel, we observe that names with [a] were judged to be more likely for post-evolution characters than those names with [i].

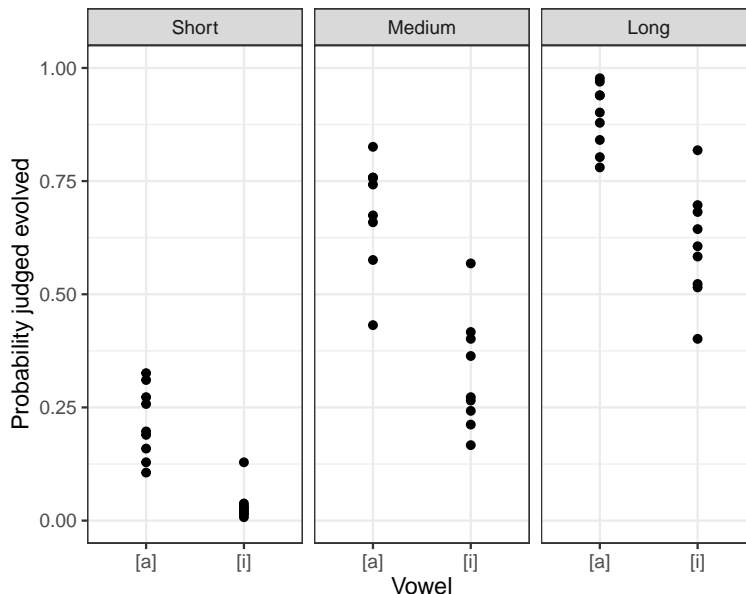


Figure 3: The results of Experiment 2, showing the cumulative effect of length modulated by vowel. Each dot represents average “evolved response” for one item.

The statistical analysis shows that the binary variable of vowel ([a] vs. [i], $\beta = 2.04$, 95% CI [0.68, 3.36]) as well as the continuous variable of length (1-unit increase $\beta = 2.67$, 95% CI [2.19, 3.17]) both meaningfully predicted participants’ judgements of evolvedness, while the interaction of the two did not ($\beta = 0.15$, 95% CI [-0.46, 0.78]). We conclude that both counting (length 1 vs. 2 vs. 3) and ganging (vowel + length) cumulativity obtained, and that since the interaction did not meaningfully differ from zero, the ganging cumulativity between vowel quality and name length was linear in nature.

3.2.2 Probing the linearity of counting cumulativity

To assess whether the counting cumulativity was linear, we re-fit the model above with length as a three-level unordered factor. We then used the samples returned by the Bayesian model

to calculate the distributions of probable values of the log-odds of being judged evolved for each combination of length and vowel quality. These are plotted in Figure 4.

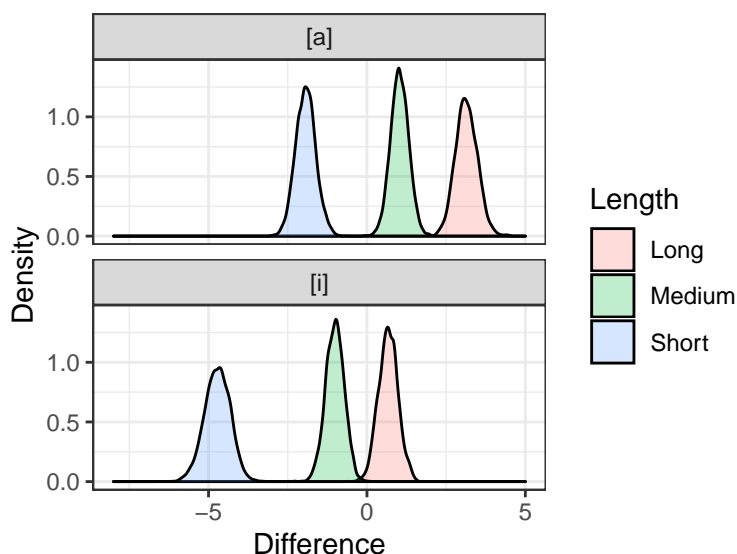


Figure 4: Estimates of the log-odds of being judged evolved for each level of length, divided by vowel quality.

Since we are interested in whether the change in log-odds when moving from short to medium is different than that of moving from medium to long, we subtracted the adjacent categories from each other, yielding distributions over differences seen in Figure 5.

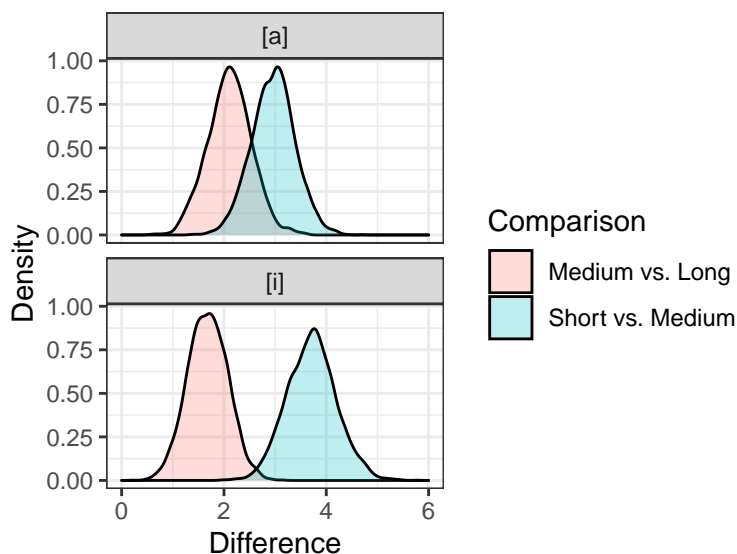


Figure 5: Differences in effect of length by category, divided by vowel quality.

Finally, we can use these distributions to answer the question of whether counting cumulativity is linear, sub-linear, or super-linear. As noted above, linear cumulativity means that the log-odds of being judged evolved increases by the same amount for each adjacent pair of categories; if this were the case in Experiment 2, we expect the pink and blue distributions to be entirely overlapping; to the degree that they are not, the cumulativity is sub-linear (pink below blue) or super-linear (blue below pink).

To more directly visualize the linearity of counting cumulativity, we can simply examine whether the *difference* between these two distributions is positive, negative, or centered on zero. Further, we can average across the two vowel qualities, since our hypothesis in question 3 above does not concern whether the linearity of counting cumulativity itself differs by what it is ganged with; such a question is interesting, but beyond the scope of conclusions that can be reasonably drawn using this experiment. This yields difference in differences in log-odds of being judged evolved between the two levels of counting cumulativity, plotted in Figure 6. We find that the vast majority of credible values for this difference are above zero: $\beta = 1.44$, 95% CI [0.45, 2.45]. We therefore conclude that Experiment 2’s counting cumulativity is sub-linear: the increase in likelihood of perceived evolvedness going from medium to long length is less than that associated with going from short to medium.

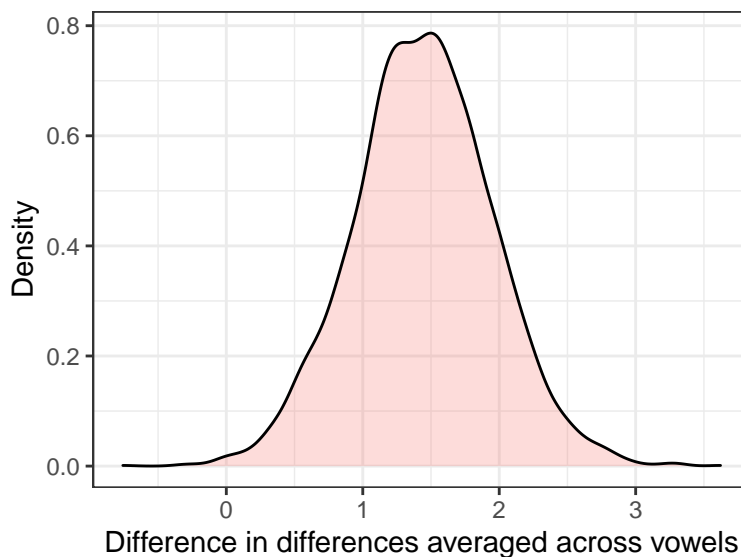


Figure 6: Differences in effect of length by category, divided by vowel quality.

3.3 Discussion

Experiment 2 demonstrated that counting and ganging cumulativity both obtain in the domain of sound symbolism, paralleling recent findings from other areas of phonological

competence (Hayes 2020; McPherson and Hayes 2016; Zuraw and Hayes 2017). This result supports further the idea that phonology operates with numeric optimization, since the two types of cumulativity interact as predicted by this theory—the multiple levels of length cleanly intersect with vowel without complication. We can also examine the nature of the numerical optimization at play: we find that ganging cumulativity (the overlap of length and vowel) is linearly cumulative, while counting cumulativity (the gradual increase along the continuous *length* dimension) is sub-linear. Although it is tempting to attribute this difference to the type of cumulative effect (that is, it is sub-linear *because* it is counting cumulativity), the current data and the state of the phonological literature do not allow us to draw this conclusion; such a query is left for future research.

4 General discussion

The two experiments have shown that sound symbolic effects operate cumulatively when English speakers are provided with new names for Pokémon characters and are asked to judge their evolution status. One may ask if the observed cumulative patterns are surprising at all; i.e. they could have been otherwise. Our answer is positive: going back to the general issue of how speakers make linguistic decisions (§1.1), the participants could have resorted to a fast and frugal decision strategy (Gigerenzer and Gaissmaier 2011); for example, they could have assigned all names with [i] to pre-evolution characters and those with [a] to post-evolution characters, especially given that these sound symbolic effects are so robust (Jespersen 1922; Sapir 1929 *et seq*). Or, given that the iconicity of quantity (Haiman 1980) is such a robust principle in the Pokémon universe (Shih et al. 2019), they could have assigned all long names to post-evolution characters, and could have made their decision solely based on that criterion. However, the current results show that English speakers did not resort to such fast-and-frugal decision strategies: they instead probabilistically took all factors (vowel quality, consonant voicing, and different degrees of length) into consideration when they decided whether each name belonged to a pre-evolution character or a post-evolution character. The current results thus suggest that they took an optimization approach based on numeric ordering. As such, we do not take the current results to be trivial—they provide further evidence that speakers deploy numeric optimization instead of lexicographic optimization in linguistic judgment patterns, *a la* the increasing body of evidence reviewed in §1.2.

Previous studies have focused on the question of *whether* cumulativity holds of linguistic patterns (§1.2), but less attention has been paid to questions of *how* it does so. We demonstrate that (1) three factors can interact cumulatively, (2) when they do, their ganging cumulativity is linear, (3) but at least in the case of length, we found their counting

cumulativity is sub-linear, and (4) counting cumulativity and ganging-up cumulativity can co-exist within the same system. We do not pretend that the current experiments offer a final answer to these questions, but we nevertheless submit that only through case studies of this kind will we understand how speakers take different sorts of information into account when they make a linguistic decision. Once we know that linguistic effects are generally cumulative, the next research questions should be *how* cumulative they are—our experiments can be understood as offering a first set of answers to these questions.

The current experiments have also contributed toward expanding on the Pokémonastics database, a resource that can be used for cross-linguistic comparisons of sound symbolic patterns (Shih et al. 2019). As reviewed in §1.3, the effects of vowel quality were known to affect the judgment of evolution status for Japanese speakers (Kumagai and Kawahara 2019) and Brazilian Portuguese speakers (Godoy et al. 2019). While this effect was also shown to be productive among English speakers (Kawahara and Moore to appear), the current replication of the effects further lends support for the robustness of this sound symbolic patterns across languages.

The fact that we found an effect of voiced obstruents in Experiment 1 is also encouraging, as in one of the previous studies, the effect was not significant (Kawahara and Moore to appear). The current experiment shows that with a sufficient number of items and speakers, we can, with a reasonably amount of confidence, identify a sound symbolic effect of voiced obstruents even among English speakers. This sound symbolic effect too was previously identified as active among Japanese speakers (Kumagai and Kawahara 2019) and Brazilian Portuguese speakers (Godoy et al. 2019), a cross-linguistic parallel that we find intriguing and important. It may be the case that sound symbolic values of voiced obstruents are grounded in the articulatory/acoustic properties of these sounds, and hence may be available to speakers of different languages (Shinohara and Kawahara 2016). Again, we do not pretend as if studying three languages suffices, but it points to a hypothesis that phonetically-motivated sound symbolic patterns are universally available to speakers of different languages (Imai and Kita 2014).

One assumption that we made throughout this paper is that sound symbolic judgments are governed by—or at least related to—phonological knowledge, or put differently, we can use data from sound symbolic patterns to explore our linguistic knowledge (Kawahara, Katsuda, and Kumagai 2019; Kawahara 2020a; Shih 2020). This assumption was not widely shared in the mainstream generative linguistics at least until recently, in which sound symbolism was considered as residing outside the realm of linguistic knowledge proper. However, sound symbolism and phonological patterns share several non-trivial properties, as recently argued by Kawahara (2020b) (see also Alderete and Kochetov 2017).

To be more specific, phonological patterns and sound symbolic patterns seem to share two fundamental properties in common; namely, stochasticity (Dingemanse 2018) and cumulativity, the latter of which was the focus of this paper. As reviewed in section 1.2, cumulativity seems to be the norm for phonological patterns; the current results, which show that cumulativity is also the norm in sound symbolic patterns, thus strengthen the conclusion that there exist non-trivial parallels between phonological patterns and sound symbolic patterns. The parallel between phonological input-output mapping and sound symbolic sound-meaning mapping seems to require a unified explanation, suggesting the possibility that they are governed by a similar mechanism (Kawahara, Katsuda, and Kumagai 2019; Kawahara 2020a; Shih 2020).

In turn, we hope to have demonstrated by way of detailed case studies that we can study sound symbolic patterns to explore the nature of linguistic patterns. In particular, we were able to shed light on several questions regarding how speakers take into account different sorts of information when they make a linguistic judgment. In this sense, sound symbolism provides a new test ground for phonological hypotheses. On the other hand, cumulativity is one aspect that has been underexplored in the studies of sound symbolism (Kawahara 2020a; Kawahara 2020b).⁹ Since cumulativity is one of the central issues in contemporary phonological studies, it allowed us to examine this underexplored aspect of sound symbolism, which in and of itself has descriptive values. All in all, we conclude that formal phonology and studies on sound symbolism can and should mutually inform one another.

Data Accessibility Statement

The experimental data from the current experiments as well as the R markdown files are available as supplementary materials.

Ethics and consent

Experiments 1 and 2 were conducted under the ethical approval granted by the first author’s institution. A subset of the participants for Experiment 2 was recruited from the UCLA experiment participant pool, which was approved by the second author’s institution. A consent form was provided to the participants before the experiments.

⁹A few exceptions which directly explored this aspect of sound symbolism include Ahlner and Zlatev (2010), Kawahara and Kumagai (to appear) and Thompson and Estes (2011).

Funding

This project is supported by the JSPS grant #17K13448 (PI: Shigeto Kawahara) as well as the research money granted to the Keio Institute of Cultural and Linguistic Studies. This work is also supported in part by NSF Graduate Research Fellowship DGE-1650604 to Canaan Breiss.

Competing interests

We declare no competing interests.

References

- Ahlner, Felix and Jordan Zlatev (2010). “Cross-modal iconicity: A cognitive semiotic approach to sound symbolism”. In: *Sign Systems Studies* 38.1/4, pp. 298–348.
- Albright, Adam (2009). “Feature-based generalisation as a source of gradient acceptability”. In: *Phonology* 26.1, pp. 9–41.
- Alderete, John (1997). “Dissimilation as local conjunction”. In: *Proceedings of the North East Linguistics Society* 27. Ed. by Kiyomi Kusumoto. Amherst: GLSA, pp. 17–31.
- Alderete, John and Alexei Kochetov (2017). “Integrating sound symbolism with core grammar: The case of expressive palatalization”. In: *Language* 93, pp. 731–766.
- Anttila, Arto, Scott Borgeson, and Giorgio Magri (2019). “Equiprobable mappings in weighted constraint grammars”. In: *SIGMORPHON 2019*.
- Anttila, Arto and Giorgio Magri (2018). “Does MaxEnt overgenerate? Implicational universals in Maximum Entropy Grammar”. In: *Proceedings of the Annual Meeting of Phonology 2017*.
- Bailey, Todd and Ulrike Hahn (2001). “Determinants of wordlikeness: Phonotactics or lexical neighborhoods?” In: *Journal of Memory and Language* 44, pp. 568–591.
- Blust, Robert (2012). “One mark per word? Some patterns of dissimilation in Austronesian and Australian languages”. In: *Phonology* 29.3, pp. 355–381.
- Boersma, Paul and Bruce Hayes (2001). “Empirical tests of the Gradual Learning Algorithm”. In: *Linguistic Inquiry* 32, pp. 45–86.
- Boersma, Paul and Joe Pater (2016). “Convergence properties of a Gradual Learning Algorithm for Harmonic Grammar”. In: *Harmonic Grammar and Harmonic Serialism*. Ed. by John J. McCarthy and Joe Pater. London: Equinox, pp. 389–434.

- Breiss, Canaan (2019). *Cumulativity by default in phonotactic learning: Experimental and computational evidence*. Ms. University of California, Los Angeles.
- Breiss, Canaan and Adam Albright (2020). *Cumulative markedness effects and (non-)linearity in phonotactics*. Ms. UCLA and MIT.
- Bürkner, Paul-Christian (2017). *brms: An R Package for Bayesian Multilevel Models using Stan*. R package.
- Chodroff, Eleanor and Colin Wilson (2014). “Burst spectrum as a cue for the stop voicing contrast in American English”. In: *Journal of the Acoustical Society of America* 136.5, pp. 2762–2772.
- Coetzee, Andries W. and Joe Pater (2008). “Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic”. In: *Natural Language and Linguistic Theory* 26, pp. 289–337.
- Coleman, John and Janet Pierrehumbert (1997). “Stochastic phonological grammars and acceptability”. In: *Computational Phonology: Third Meeting of the ACL Special Interest Group in Computational Phonology*. Somerset: Association for Computational Linguistics, pp. 49–56.
- Crowhurst, Megan (2011). “Constraint conjunction”. In: *The Blackwell companion to phonology*. Ed. by Marc van Oostendorp et al. Oxford: Blackwell-Wiley, pp. 1461–1490.
- Cuskley, Christine, Julia Simner, and Simon Kirby (2017). “Phonological and orthographic influences in the bouba-kiki effect”. In: *Psychological Research* 81.1, pp. 119–130.
- de Lacy, Paul (2009). “Phonological evidence”. In: *Phonological Argumentation: Essays on Evidence and Motivation*. Ed. by Steve Parker. London: Equinox, pp. 43–77.
- Dingemanse, Mark (2018). “Redrawing the margins of language: Lessons from research on ideophones”. In: *Glossa* 3.1, 4, doi:org/10.5334/gjgl.444.
- Frisch, Stephan, Janet Pierrehumbert, and Michael Broe (2004). “Similarity avoidance and the OCP”. In: *Natural Language and Linguistic Theory* 22, pp. 179–228.
- Fukazawa, Haruka et al. (2015). “Two is too much: Geminate devoicing in Japanese”. In: *Phonological Studies* 18, pp. 3–10.
- Gigerenzer, Gerd and Wolfgang Gaissmaier (2011). “Heuristic decision making”. In: *Annual Review of Psychology* 62, pp. 451–482.
- Godoy, Mahayana C. et al. (2019). “Gotta name’em all: An experimental study on the sound symbolism of Pokémon names in Brazilian Portuguese”. In: *Journal of Psycholinguistic Research*.
- Green, Christopher R. and Stuart Davis (2014). “Superadditivity and limitations on syllable complexity in Bambara words”. In: *Perspectives on Phonology, Theory, and Acquisition*:

- Papers in honor of Daniel A. Dinnsen*. Ed. by Ashely W. Farris-Trimble and Jessica Barlow. Amsterdam: John Benjamins, pp. 223–247.
- Guy, Gregory and C. Boberg (1994). “The Obligatory Contour Principle and sociolinguistic variation”. In: *Toronto Working Papers in Linguistics*.
- (1997). “Inherent variability and the obligatory contour principle”. In: *Language Variation and Change* 9, pp. 149–164.
- Haiman, John (1980). “The iconicity of grammar: Isomorphism and motivation”. In: *Language* 56.3, pp. 515–540.
- ed. (1985). *Iconicity in Syntax*. Amsterdam: John Benjamins.
- Hall, Nancy (2012). “Perceptual errors or deliberate avoidance? Types of English /r/-dissimilation.” In: *Proceedings of the Thirty-Fourth Annual Meeting of the Berkeley Linguistics Society*, pp. 133–144.
- Hay, Jennifer, Janet Pierrehumbert, and Mary Beckman (2003). “Speech perception, well-formedness, and the statistics of the lexicon”. In: *Papers in Laboratory Phonology VI: Phonetic interpretation*. Ed. by John Local, Richard Ogden, and Rosalind Temple. Cambridge: Cambridge University Press, pp. 58–74.
- Hayes, Bruce (2017). “Varieties of Noisy Harmonic Grammar”. In: *Proceedings of AMP*.
- (2020). *Assessing grammatical architectures through their quantitative signatures*. Talk presented at BLS.
- Hinton, Leane, Johanna Nichols, and John Ohala (1994). *Sound Symbolism*. Cambridge: Cambridge University Press.
- Imai, Mutsumi and Sotaro Kita (2014). “The sound symbolism bootstrapping hypothesis for language acquisition and language evolution”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 369.1651.
- Ito, Junko and Armin Mester (1995). “Japanese phonology”. In: *The Handbook of Phonological Theory*. Ed. by John Goldsmith. Oxford: Blackwell, pp. 817–838.
- (1999). “The phonological lexicon”. In: *The Handbook of Japanese Linguistics*. Ed. by Natsuko Tsujimura. Oxford: Blackwell, pp. 62–100.
- (2003). *Japanese Morphophonemics*. Cambridge: MIT Press.
- Jäger, Gerhard (2007). “Maximum Entropy Models and Stochastic Optimality Theory”. In: *Architectures, rules, and preferences: Variations on themes*. Ed. by Joan W. Bresnan. CSLI, pp. 467–479.
- Jäger, Gerhard and Anette Rosenbach (2006). “The winner takes it all—almost: Cumulativity in grammatical variation”. In: *Linguistics* 44.5, pp. 937–971.
- Jespersen, Otto (1922). “Symbolic value of the vowel *i*”. In: *Linguistica: Selected Papers in English, French and German*. Vol. 1. Copenhagen: Levin and Munksgaard, pp. 283–303.

- Kawahara, Shigeto (2011). “Japanese loanword devoicing revisited: A rating study”. In: *Natural Language and Linguistic Theory* 29.3, pp. 705–723.
- (2020a). *Cumulative effects in sound symbolism*. Ms. Keio University.
- (2020b). “Sound symbolism and theoretical phonology”. In: *Language and Linguistic Compass*.
- Kawahara, Shigeto, Hironori Katsuda, and Gakuji Kumagai (2019). “Accounting for the stochastic nature of sound symbolism using Maximum Entropy model”. In: *Open Linguistics* 5, pp. 109–120.
- Kawahara, Shigeto and Gakuji Kumagai (to appear). “What voiced obstruents symbolically represent in Japanese: Evidence from the Pokémon universe”. In: *Journal of Japanese Linguistics*.
- (2019). “Expressing evolution in Pokémon names: Experimental explorations”. In: *Journal of Japanese Linguistics* 35.1, pp. 3–38.
- Kawahara, Shigeto and Jeff Moore (to appear). “How to express evolution in English Pokémon names”. In: *Linguistics*.
- Kawahara, Shigeto, Atsushi Noto, and Gakuji Kumagai (2018). “Sound symbolic patterns in Pokémon names”. In: *Phonetica* 75.3, pp. 219–244.
- Kawahara, Shigeto and Shin-ichiro Sano (2014). “Identity Avoidance and Lyman’s Law”. In: *Lingua* 150, pp. 71–77.
- (2016). “/p/-driven geminate devoicing in Japanese: Corpus and experimental evidence”. In: *Journal of Japanese Linguistics* 32, pp. 57–77.
- Kim, Seoyoung (2019). *Modeling super-gang effects in MaxEnt: Nasal in Rendaku*. Talk presented at NELS 50.
- Kingston, John and Randy Diehl (1994). “Phonetic knowledge”. In: *Language* 70, pp. 419–454.
- (1995). “Intermediate properties in the perception of distinctive feature values”. In: *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. Ed. by Bruce Connell and Amalia Arvaniti. Cambridge: Cambridge University Press, pp. 7–27.
- Kingston, John, Randy Diehl, et al. (2008). “On the internal perceptual structure of distinctive features: The [voice] contrast”. In: *Journal of Phonetics* 36, pp. 28–54.
- Kirchner, Robert (1996). “Synchronic chain shifts in Optimality Theory”. In: *Linguistic Inquiry* 27, pp. 341–350.
- Kumagai, Gakuji and Shigeto Kawahara (2019). “Effects of vowels and voiced obstruents on Pokémon names: Experimental and theoretical approaches [in Japanese]”. In: *Journal of the Linguistic Society of Japan* 155, pp. 65–99.

- Legendre, G  raldine, Yoshiro Miyata, and Paul Smolensky (1990a). “Harmonic Grammar – A formal multi-level connectionist theory of linguistic well-formedness: An application”. In: *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 884–891.
- (1990b). “Harmonic Grammar – A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations”. In: *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 388–395.
- McPherson, Laura and Bruce Hayes (2016). “Relating application frequency to morphological structure: The case of Tommo So vowel harmony”. In: *Phonology* 33, pp. 125–167.
- Newman, Stanley (1933). “Further experiments on phonetic symbolism”. In: *American Journal of Psychology* 45, pp. 53–75.
- O’Hara, Charlie (2017). “How abstract is more abstract? Learning abstract underlying representations”. In: *Phonology* 34.2, pp. 324–345.
- Ohala, John (1983a). “The origin of sound patterns in vocal tract constraints”. In: *The Production of Speech*. Ed. by Peter MacNeilage. New York: Springer-Verlag, pp. 189–216.
- (1983b). “The phonological end justifies any means”. In: *Proceedings of the 13th International Congress of Linguists*. Ed. by S. Hattori and K. Inoue. Tokyo: Sanseido, pp. 232–243.
- (1994). “The frequency code underlies the sound symbolic use of voice pitch”. In: *Sound Symbolism*. Ed. by Leane Hinton, Johanna Nichols, and John Ohala. Cambridge: Cambridge University Press, pp. 325–347.
- Pater, Joe (2009). “Weighted constraints in generative linguistics”. In: *Cognitive Science* 33, pp. 999–1035.
- (2016). “Universal grammar with weighted constraints”. In: *Harmonic Grammar and Harmonic Serialism*. Ed. by John J. McCarthy and Joe Pater. London: Equinox, pp. 1–46.
- Pierrehumbert, Janet (2020). “70+ years of probabilistic phonology”. In: *Oxford Handbook on the History of Phonology*. Ed. by B. Elan Dresher and Harry van der Hulst. Oxford University Press.
- Pizzo, Presley (2015). “Investigating properties of phonotactic knowledge through web-based experimentation”. Doctoral dissertation. University of Massachusetts, Amherst.
- Potts, Christopher et al. (2010). “Harmonic grammar with linear programming: From linear systems to linguistic typology”. In: *Phonology* 27.1, pp. 1–41.

- Prince, Alan and Paul Smolensky (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden and Oxford: Blackwell [the original manuscript circulated in 1993].
- (1997). “Optimality: From neural networks to universal grammar”. In: *Science* 275, pp. 1604–1610.
- Proctor, Michael I., Christine H. Shadle, and Khalil Iskarous (2010). “Pharyngeal articulation differences in voiced and voiceless fricatives”. In: *Journal of the Acoustical Society of America* 127.3, pp. 1507–1518.
- Rose, Sharon and Lisa King (2007). “Speech error elicitation and co-occurrence restrictions in two Ethiopian Semitic languages”. In: *Language and Speech* 50.4, pp. 451–504.
- Sapir, Edward (1929). “A study in phonetic symbolism”. In: *Journal of Experimental Psychology* 12, pp. 225–239.
- Shih, Stephanie S (2017). “Constraint conjunction in weighted probabilistic grammar”. In: *Phonology*, pp. 243–268.
- (2020). “Gradient categories in lexically-conditioned phonology: An example from sound symbolism”. In: *Proceedings of the 2019 Annual Meeting on Phonology*.
- Shih, Stephanie S et al. (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 and 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*. Berkeley: Berkeley Linguistics Society, pp. 396–410.
- Sidorov, V. N., Penny M. Pexman, and Jean Saint-Aubin (2016). “From the Bob-Kirk effect to the Benoit-Éric effect: Testing the mechanism of name sound symbolism in two languages”. In: *Acta Psychologica* 169, pp. 88–99.
- Smith, Brian W. and Joe Pater (2017). *French schwa and gradient cumulativity*. Ms. University of California, Berkeley and University of Massachusetts, Amherst (ROA-1702).
- Smolensky, Paul (1986). “Information processing in dynamical systems: Foundations of harmony theory”. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Ed. by D. Rumelhart, J. McClelland, and PDPR Group. Vol. 1: Foundations. Cambridge, MA: Bradford Books/MIT Press, pp. 194–281.
- (1995). *On the internal structure of the constraint component CON of UG*. Talk presented at the University of California, Los Angeles (ROA-86).

- Smolensky, Paul (1997). *Constraint interaction in generative grammar II: Local conjunction, or random rules in Universal Grammar*. Handout of talk given at Hopkins Optimality Theory Workshop/Maryland Mayfest, Baltimore.
- Smolensky, Paul and Géraldine Legendre (2006). *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar Volume I: Cognitive Architecture Volume II: Linguistic and Philosophical Implications*. Cambridge: MIT Press.
- Stevens, Kenneth and Sheila Blumstein (1981). “The search for invariant acoustic correlates of phonetic features”. In: *Perspectives on the Study of Speech*. Ed. by Peter Eimas and Joanne D. Miller. New Jersey: Earlbaum, pp. 1–38.
- Tesar, Bruce (2007). *A comparison of lexicographic and linear numeric optimization using violation difference ratios*. Ms. Rutgers University.
- Thompson, Patrick D. and Zachary Estes (2011). “Sound symbolic naming of novel objects is a graded function”. In: *Quarterly Journal of Experimental Psychology* 64.12, pp. 2392–2404.
- Vance, Timothy (2015). “Rendaku”. In: *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*. Ed. by Haruo Kubozono. Berlin: Mouton de Gruyter, pp. 397–441.
- Vasishth, Shravan et al. (2018). “Bayesian data analysis in the phonetic sciences: A tutorial introduction”. In: *Journal of Phonetics* 71, pp. 147–161.
- Vennemann, Theo (1988). *Preference Laws for Syllable Structure and the Explanation of Sound Change: With Special Reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Gruyter.
- Westbury, John R. (1983). “Enlargement of the supraglottal cavity and its relation to stop consonant voicing”. In: *Journal of the Acoustical Society of America* 73, pp. 1322–1336.
- Wichmann, Søren, Eric W. Holman, and Cecil H. Brown (2010). “Sound symbolism in basic vocabulary”. In: *Entropy* 12.4, pp. 844–858.
- Wübben, M. and F. von Wagenenheim (2008). “Instant customer base analysis: Managerial heuristics often “get it right””. In: *Journal of Marketing* 72, pp. 82–93.
- Yip, Moira (1988). “The Obligatory Contour Principle and phonological rules: A loss of identity”. In: *Linguistic Inquiry* 19, pp. 65–100.
- Zuraw, Kie and Bruce Hayes (2017). “Intersecting constraint families: An argument for Harmonic Grammar”. In: *Language* 93, pp. 497–548.