Cumulative effects in sound symbolism

Abstract

Sound symbolism, systematic associations between sounds and meanings, has not generally received much serious attention from theoretical phonologists. On the other hand, there is a dramatically growing interest in sound symbolism by psychologists and cognitive scientists. Against this background, overarching goals of this paper are (i) to show that sound symbolic associations and phonological mappings share a non-trivial property (i.e. cumulativity) and (ii) to demonstrate that the same analytical device—Maximum Entropy Harmonic Grammar—can straightforwardly handle this property in these two apparently disparate domains. By pointing out a non-trivial parallel between sound symbolic connections and phonological mappings, I hope to show theoretical phonologists that studying sound symbolism can be interesting and informative. I also hope to show those researchers who study sound symbolism that an analytical device that theoretical linguists employ is useful in that it allows us to model an important aspect of sound symbolism. My ultimate goal is thus to enhance the communication between theoretical phonologists and researchers who work on sound symbolism who are not yet interested in theoretical phonology. The paper also has a descriptive value in that it summarizes various cases of cumulative effects in sound symbolic patterns from a variety of languages.

Keywords: sound symbolism, theoretical phonology, cumulativity, MaxEnt, Optimality Theory

1 Introduction

1.1 Synopsis

Sound symbolism refers to systematic associations between sounds and meanings (Hinton et al. 2006). Perhaps the most famous case of sound symbolism is what is now known as the *bouba-kiki* effect, in which segments like [b] and [u] tend to be associated with round figures, whereas segments like [k] and [i] tend to be associated with angular shapes (Ramachandran & Hubbard 2001), as in Figure 1.

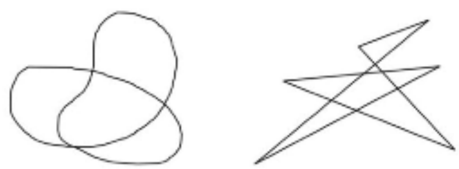

Figure 1: An illustration of the *bouba-kiki* effect. Given the two nonce words *bouba* and *kiki*, the round figure on the left tends to be named *bouba*, while the angular figure on the right tends to be named *kiki*. This effect is also known as the *takete-maluma* effect (Köhler 1947), in which obstruents, especially voiceless stops, are associated with angular shapes, whereas sonorants are associated with round shapes.

Another well-known case of sound symbolism is the observation that [i] is often associated with images of smallness (Jespersen 1922; Sapir 1929), and that this vowel is very often used to express diminutive meanings across languages (Blasi et al. 2016; Ultan 1978). Since sound symbolic patterns flout the widely-held view that the relationships between sounds and meanings are arbitrary in human languages (Hockett 1959; Saussure 1916), sound symbolism has not generally received serious attention from theoretical phonologists. On the other hand, there is a surprisingly growing interest in sound symbolism in other disciplines, which is now very actively studied by anthologists, phoneticians, psychologists, cognitive scientists, and even marketing researchers (Dingemanse et al. 2015, Kawahara 2019, Lockwood & Dingemanse 2015, Perniss et al. 2010, Nuckolls 1999 and Sidhu & Pexman 2018 offer recent reviews). Against this background, overarching goals of the present paper are to show that (i) sound symbolic connections and phonological mappings share a non-trivial property (i.e. cumulativity) and (ii) to demonstrate that the same analytical device—Maximum Entropy Harmonic Grammar (henceforth MaxEnt)—can naturally model this property in these two apparently disparate domains. My hope in writing this paper is therefore to show the theoretical phonologists that studying sound symbolism can be interesting and informative, by pointing out a non-trivial parallel between sound symbolic connections and phonological mappings (see also Shih 2019 who shares the same spirit). I also hope to show those researchers who study sound symbolism that a formal analytical device that theoretical linguists deploy is useful in that it allows us to model an important aspect of sound symbolism. The paper in addition has a descriptive value in that it summarizes various cases of cumulative effects in sound symbolic patterns from a variety of languages, including Brazilian Portuguese, English, Korean, Japanese and Swedish. The issue of whether sound symbolic effects are cumulative or not is an understudied area of research, which has been directly studied only by a few studies (Ahlner & Zlatev 2010; Thompson & Estes 2011). This is the descriptive gap that the current paper intends to address.

1.2 Sound symbolism and theoretical phonology

Let me elaborate on a few specific points that were summarized in the preceding subsection. Modern linguistic theories have assumed that the relationships between sounds and meanings in human languages are essentially arbitrary, since the influential work by Hockett (1959) and Saussure (1916). This assumption to the best of my knowledge has not been seriously challenged in modern linguistic theories until the present, including most phonological theories. However, I should note that there are studies which have analyzed sound alternation patterns that are demonstrably caused by sound symbolic principles. One recent prominent example is Alderete & Kochetov (2017), who propose that sound symbolism should be integrated with core phonological grammar, arguing that some patterns of palatalization found in baby-talk registers across different languages—"expressive palatalization"—are driven by sound symbolic considerations, rather than by phonological or phonetic considerations. What they analyze, however, is still sound alternation patterns (i.e. palatalization), and not sound-meaning connections *per se*. As Dingemanse (2018) succinctly summarizes (sect. 5.2), there have been other analyses of (morpho)phonological alternations caused by sound symbolic principles by theoretical phonologists, especially in the context of analyzing phonological properties of ideophones. McCarthy (1983), for example, analyzes feature-size morphemes in ideophones in various languages, but, crucially, he sets aside the sound-meaning correspondences observed in these constructions. To quote:

> Many of the examples that I discuss here may be subsumed under the general designation of sound symbolism because they make phonetic distinctions that stand in an essentially iconic relationship with their meaning. My concern is entirely with the formal properties of these systems – what sorts of segments they affect, what phonetic properties they exploit, and how they might apply throughout a word. Thus, I have nothing to say about the issue of iconic versus symbolic meaning nor have I attempted to review here the extensive literature on this topic (p. 136).

Similarly, Mester & Ito (1989) offer a famous analysis of distribution of palatalized segments in

Japanese mimetic forms, in which palatalization carries distinctive sound symbolism meanings, such as "childishness" and "uncontrolledness" (Hamano 1996). However, they set aside the analysis of these sound symbolic meanings, stating that:

> We are interested here in the intricate and, at first glance, puzzling surface distribution of the palatal prosody, and not in its somewhat elusive semantic-pragmatic contribution to the meaning of the base (p. 269).

In no way am I implying anything negative by quoting these remarks, but instead, I believe that these quotes represent the general view that formal phonologists share: they may be interested in morphophonological properties of sound alternations triggered by sound symbolic considerations, but not in sound-meaning relationships *per se*.[1] This situation is what I would like to challenge in the present paper.

Before I delve into my main argument, I admit that different phonologists are willing to accept different types of evidence to construct phonological theories. de Lacy (2009), for example, admits neither phonotactic patterns nor loanword adaptation patterns as domains of phonological inquiry. Some researchers are willing to use patterns of verbal art, like rhyming and text-setting, as evidence for phonological theories, while others are less willing to do so.[2] Nevertheless, I would like to attempt to show in this paper that phonologists may find interesting parallels between sound symbolic correspondences and phonological mappings. At a more general level, phonologists and researchers on sound symbolism have common interests and address similar issues, and they can thus potentially inform one another (Kawahara 2019). To provide one example, it has long been noted in the studies of sound symbolism that some sound-meaning connections have clear bases in the articulatory and/or acoustic properties of the sounds at issue. For instance, [b] and [u] cause images of roundness across different languages (see Figure 1), and it is natural to conjecture that

[1] A clear exception is Jakobson (1978), who discusses sound symbolic values of distinctive features: "[o]wing to neuropsychological laws of synaesthesia, phonetic oppositions can themselves evoke relations with musical, chromatic, olfactory, tactile, etc. sensations. For example, the opposition between acute and grave phonemes has the capacity to suggest an image of bright and dark, of pointed and rounded, of thin and thick, of light and heavy, etc. This 'sound symbolism'...[the] inner value of the distinctive features, although latent, is brought to life as soon as it finds a correspondence in the meaning of a given word and in our emotional or aesthetic attitude towards this word and even more towards pairs of words with two opposite meanings (pp. 113)." See also Jakobson (1971) who offers cross-linguistic sound symbolic analyses of kinship terms, which make crucial use of distinctive features.

[2] I cannot think of any paper other than de Lacy (2009) who explicitly denies the use of such evidence for phonological argumentation. However, I have encountered such negative remarks against the use of "external evidence" (Bagemihl 1995; Churma 1979) in personal conversations as well as in anonymous reviews. On this note, Gouskova (2013) states "[p]honology is changing rapidly. Whereas in the past, we used the same methodologies and largely agreed on the goals of the field, the nature of evidence, and the assumptions about representations, no such agreement exists today—as the field grows, so does diversity of opinion. Some phonologists collect the evidence for their theories using introspection, fieldwork, and descriptive grammars, while others trust only quantitatively robust experimental or corpus data. Some test phonological theories computationally, aiming to replicate human behavior or sound patterns in an explicit model, whereas others prefer to compare theories on conceptual grounds (p. 173)."

lip gestures of these sounds are iconically mapped onto their meaning of roundedness (D'Onofrio 2014). This observation may remind theoretical phonologists of a classic observation that at least some sound alternations are driven by articulatory and/or acoustic considerations (e.g. Hayes et al. 2004), setting aside the issue of whether these connections hold synchronically or diachronically. In other words, both sound-meaning connections and phonological patterns may be grounded in the phonetic properties of sounds under question. Another possible parallel, which this paper explores in further depth, is *cumulativity*.

1.3 The starting point

At first glance, phonological mappings and sound-meaning connections may appear to involve very different mechanisms—after all, sound symbolism has long been considered as residing outside the realm of theoretical phonology. However, Kawahara et al. (2019) point out that these mechanisms may not be as different as they first appear to be. The crucial starting point is to recognize that generative phonology has always been a function that maps one representation (e.g. underlying representation) to another representation (e.g. surface representation). Then, there is nothing that prevents us from using the same mechanism to model mapping between sound and meaning.[3]

One reason that traditional phonological devices had not been used for analyses of sound symbolism may be the fact that sound symbolic patterns are almost always stochastic—the relationships between sounds and meanings always manifest themselves as stochastic tendencies rather than deterministic connections, and traditional models of theoretical phonology were not designed to handle such stochastic generalizations. However, the situation has radically changed recently—there is a growing body of evidence demonstrating that phonological knowledge is stochastic (see e.g. Coetzee 2012; Coetzee & Pater 2011; Hayes & Londe 2006; Pierrehumbert 2001), and there are now various theoretical models of phonology which can account for such stochastic phonological knowledge, one of which is Maximum Entropy (MaxEnt) Harmonic Grammar (Goldwater & Johnson 2003). With this theoretical development as a background, Kawahara et al. (2019) used MaxEnt to model stochastic nature of sound symbolism.

Another potential reason why phonological theories were not applied to sound symbolic patterns may have been that in the SPE-style rules (Chomsky & Halle 1968) (A $\rightarrow$ B / C _ D), the

[3] I note in passing that there are recent models within generative phonology which no longer posit underlying representations. Hyman (2018) offers a summary of these models, as well as responses to these proposals, explaining why we want to keep underlying representations in phonological theorization. Whether we want to eliminate underlying representations from phonological theories or not, phonological theories have always been a mechanism that can accept an input (e.g. a nonce word or a novel root-suffix combination) and put out an output that is grammatically licit in that language. In other words, phonology has always been a grammatical device that produces a pronounceable output given a novel input. Likewise, as a large body of the literature on sound symbolism suggests, speakers can, given a nonce word, very often guess its meaning at a level that is higher than chance; i.e. they can, albeit stochastically, map novel sound sequences to meanings.

input-output mapping (A $\rightarrow$ B) and its environment (*CAD) are inseparably encapsulated in one format; on the other hand, sound symbolism is hardly sensitive to the surrounding phonological environments,[4] and what is crucial is the mapping from A to B. Constraint-based theories of phonology, of which MaxEnt is one example, have liberated the input-output mapping from the cause of its mapping (Prince & Smolensky 1993/2004), which makes the parallel between sound symbolism and phonological mappings clearer.

Building on Kawahara et al. (2019), the theoretical goal of this paper is to show that MaxEnt, combined with Optimality Theoretic constraints (Prince & Smolensky 1993/2004), can also model cumulative effects in sound symbolism, just as it can model cumulative effects in phonological patterns.

2 Cumulativity in phonology: A summary

To demonstrate that phonological patterns and sound symbolic patterns share an interesting property in that they both show cumulative effects, this section first summarizes evidence that phonological patterns are cumulative.[5] The question of whether phonological systems show cumulative effects or not has been one of the central issues in phonological theorization since the inception of Optimality Theory (OT: Prince & Smolensky 1993/2004). As discussed in further detail in section 4, in OT, when Constraint A dominates both Constraints B and C, the simultaneous violation of Constraints B and C is not considered to be worse than the violation of Constraint A. Neither are any numbers of violations of B or C worse than a single violation of Constraint A. These non-cumulative natures follow from *strict domination* of constraint rankings, one central tenet of OT. In principle, OT thus does not predict cumulative effects in phonological patterns, although in practice, since the early eras of Optimality Theoretic research, researchers have pointed out potential cases of cumulative effects in phonology, and local conjunction was often deployed to model such cumulative patterns (Smolensky 1995 *et seq*; see also Crowhurst 2011). In recent years, we witness revived interests in whether phonological patterns show cumulative effects or not in the context of comparing Optimality Theory with other related constraint-based frameworks that use weights rather than ranking, including Harmonic Grammar (Legendre et al. 1990; Pater 2009), Noisy Harmonic Grammar (Boersma 1998; Boersma & Pater 2016) and MaxEnt grammar (Goldwater & Johnson 2003; Zuraw & Hayes 2017). It is probably safe to say that examining the (non-)cumulative nature of phonological patterns is one of the most important issues in current phonological theorization. In what follows, I provide a brief review of evidence that has been put

[4]An exception may include cases of phonaesthemes, such as English *gl-* (Bergen 2004). In such cases, one can argue that [g] is related to the notion of light, but only before [l].

[5]This section is inspired by a summary presented in Breiss (2019). For cumulative effects in syntactic variations, see Featherston (2005), Jäger & Rosenbach (2006), and Kellar (2006).

forward to show that phonology does show cumulative patterns.

We start with a simple example in English phonotactics: English allows #Cl clusters (e.g. *clip*) as well as sC# clusters (e.g. *list*), but barely allows both within the same word, especially when the Cs are stops (e.g. **glVsp* or **plVst*: Albright 2012; Breiss & Albright 2020). Likewise, in a phonotactic judgment pattern, Coleman & Pierrehumbert (1997) show that the combination of several "minor" phonotactic violations (e.g. *spleitisak*) is judged to be worse than one "major" phonotactic violation (e.g. *mrupation*). Hay et al. (2003) further generalized this finding and argued that acceptability judgments of nonce words reflect the accumulation of the wellformedness—in their view, probabilistic likelihoods—of subparts. Pizzo (2015) in her large-scale acceptability judgment experiment found that phonotactic violations in onsets and those in codas show cumulative effects (e.g. *tlavb* is judged to be worse than *tlag* and *plavb*)—see also Bailey & Hahn (2001) for similar results. Beyond these examples from English phonotactics, Albright (2012) shows that in Lakota, for example, combinations of a fricative and a consonant cluster are heavily under-attested. In all of these patterns summarized so far, it is not the case that "the worst phonotactic violation" has a final say in determining the acceptability of the (nonce) words under question, as OT predicts.

Cumulative effects seem to be observed in the context of phonological alternations as well. For example, the probability of *t/d*-deletion in English can be best understood as cumulative effects of different types of featural OCP constraints (Guy & Boberg 1997) (see also Coetzee & Pater 2008 and Frisch et al. 2004 for cumulative and gradient effects in similarity avoidance patterns found in the lexicon of various languages). In Japanese loanwords, singleton [p] and voiced geminates are both tolerated, as long as they appear independently; however, devoicing of geminates occurs when they co-occur (Fukazawa et al. 2015; Kawahara & 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. Blust (2012) lists many cases from Austronesian and Australian languages in which two instances of marked segments—including geminates and prenasalized segments—are avoided by way of dissimilation. Smith & Pater (2017) show that the presence/absence of schwa 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 & Davis (2014) found various instances of cumulative interactions between restrictions on complex syllable structures in Colloquial Bambara. Zuraw & Hayes (2017) offer extensive corpus-based analyses of three languages (Tagalog, French and Hungarian), showing cumulative action of different types of constraints in all three languages. In addition, as mentioned above, in general, many patterns that have been analyzed with local conjunction in Optimality Theoretic research show cumulative natures.[6]

[6] I should note, however, that local conjunction has been deployed to model other phonological patterns, such as chain shift (Kirchner 1996), derived environment effects (Lubowicz 2002), and dominant-recessive harmony patterns (Bakovic 2000). Also, reanalyses of apparently cumulative patterns have been proposed without recourse to local

Evidence for cumulative nature of phonological knowledge has been put forth by some studies in laboratory phonology tradition as well. Rose & King (2007) show that in two Semitic languages (Amharic and Chaha), when a structure violates two phonotactic restrictions, that structure is more likely to induce speech errors than when there is only one violation. Kawahara (2011) explored whether phonological devoicing is judged to be natural when it is caused by a restriction against a voiced geminate and a restriction against two voiced obstruents (a.k.a. Lyman's Law)—the results suggest that both factors cumulatively make devoicing more natural for Japanese speakers. Breiss (2019) conducted a series of artificial language learning experiments to show that cumulativity is the default strategy in phonotactic learning for English speakers. Breiss & Albright (2020) report an additional experiment supporting the same conclusion. It thus seems safe to conclude based on this body of evidence that at least some aspect of phonological knowledge has a cumulative property.

To account for this general observation that phonological patterns are (at least partially) cumulative, one analytical framework that has been gaining popularity is MaxEnt Harmonic Grammar (Albright 2012; Breiss 2019; Breiss & Hayes to appear; Daland 2015; Goldwater & Johnson 2003; Hayes 2017; Hayes et al. 2012; Jurafsky & Martin 2019; Kim 2019; Pizzo 2015; Shih 2017; Smith & Pater 2017; Zuraw & Hayes 2017; Wilson 2006, 2014 among many others). This model has been shown to be successful in modeling various patterns of cumulative effects in phonology. In addition, it has an attractive mathematical property in that it has been shown to converge on best-fitting models (Della Pietra et al. 1997). For these reasons, this paper uses MaxEnt, together with Optimality Theoretic constraints, to analyze cumulative effects in sound symbolism. I hasten to add that other models, such as Noisy Harmonic Grammar, may just work as well for the current cases at hand (see e.g. Hayes 2017 and Zuraw & Hayes 2017 for comparisons between various stochastic models of phonology)[7]—my goal in this paper is to show that sound symbolism shows cumulative effects, and a model of theoretical phonology which has shown to be successful to model phonological cumulative effects can also account for the cumulative effects in sound symbolism.[8]

conjunction, often by splitting up the relevant constraints (e.g. Kawahara 2006; Padgett 2002).

[7]There are some studies that show (or briefly mention) that MaxEnt may fit the given data better than Noisy Harmonic Grammar (Breiss 2019; Breiss & Albright 2020; Breiss & Hayes to appear; Smith & Pater 2017).

[8]One caveat is in order. MaxEnt was not proposed as a model of theoretical phonology. Jaynes (1957) proposed this mechanism as a general way to make statistical inferences based on given knowledge that is limited. Smolensky (1986) proposed to utilize MaxEnt as a model of general cognition. MaxEnt is in fact equivalent to a general statistical device known as a log-linear model or multinomial logistic regression (Breiss & Hayes to appear; Jurafsky & Martin 2019). It is a framework that is widely used in natural language processing (Berger et al. 1996).

3 A brief review of MaxEnt grammar

This section briefly reviews how MaxEnt grammar works in the context of linguistic analyses. Readers who are familiar with this analytical framework can safely skip this section. For intuitive explications of each calculation step, see Breiss & Hayes (to appear) and/or Zuraw & Hayes (2017). MaxEnt grammar is similar to Optimality Theory (OT: Prince & Smolensky 1993/2004) in that a set of candidates is evaluated against a set of constraints. Unlike OT, however, constraints are *weighted* rather than *ranked*. Consider a toy example in (1). The set of candidates that are evaluated are listed in the leftmost column. The top row lists the set of constraints that are relevant, and each constraint is assigned a particular weight.[9] The tableau shows the violation profiles of each constraint—which candidate violates which constraints how many times.

(1) A toy example tableau of MaxEnt grammar

	Constraint A Weight = 3	Constraint B Weight = 2	Constraint C Weight 1	H-score	eHarmony	*Z*	P
Candidate 1	1			1*3=3	e^{-3}= 0.0498	0.0565	88
Candidate 2		2	1	2*2+1*1=5	e^{-5}= 0.0067	0.0565	12

Based on the constraint violation profiles, for each candidate x, its Harmony Score (H-Score(x)) is calculated using the formula in (1):

$$\text{H-score}(x) = \sum_{i}^{N} w_i C_i(x) \ \ (N \text{ is the number of the constraints}) \qquad (1)$$

where w_i is the weight of the i-th constraint, and $C_i(x)$ is the number of times candidate x violates the i-th constraint. For example, Candidate 2 in the tableau (1) violates Constraint B twice and Constraint C once; its H-Score is therefore $2 * 2 + 1 * 1 = 5$.

The H-Scores are negatively exponentiated (eHarmony, e^{-H} or $\frac{1}{e^H}$: Wilson 2014), which corresponds to the probability of each candidate. Intuitively, the more constraint violation a candidate incurs, the higher the H-Score, and hence the lower the eHarmony (e^{-H}) is. Therefore, more violations of constraints lead to lower probability of that candidate. The eHarmony values are relativized against the sum of the eHarmony values of all the candidates, which is sometimes referred to as Z:

$$Z = \sum_{j}^{M} (e^{-H})_j \ \ (M \text{ is the number of the candidates}) \qquad (2)$$

[9] Constraints are called “features” in the computational linguistics literature.

In the example in (1), Z is $0.0498 + 0.0067 = 0.0565$. The probability of each candidate $x_j, p(x_j)$, is $\frac{eHarmony(x_j)}{Z}$.

For an accessible introduction to how to find optimal weights in MaxEnt grammars given the observed dataset, see Hayes & Wilson (2008). To implement the analyses that follow, I used the MaxEnt Grammar tool (Hayes et al. 2009), software which calculates the best weights for each constraint given the observed frequencies of each candidate; the software also calculates the predicted probabilities of each candidate based on these weights.

4 Counting cumulativity and ganging-up cumulativity

Before discussing actual cases of cumulative effects in sound symbolism and how they may be modeled using MaxEnt, I introduce one distinction that has proven to be useful for theoretical phonologists (and theoretical linguists in general). Jäger & Rosenbach (2006) distinguishes two types of cumulativity—counting cumulativity and ganging-up cumulativity—which present a different type of challenge to OT (Prince & Smolensky 1993/2004). To illustrate these effects in the context of OT, starting with counting cumulativity, it instantiates a case in which more than one violation of a lower-ranked constraint takes precedence over a violation of a higher ranked constraint. Consider the illustrative tableaux in (2). The first comparison shows that Constraint A is ranked higher than Constraint B—this is why [X] is selected as the winner. However, given a candidate like [Z] which violates Constraint B twice, it can lose against the candidate [W]. This situation instantiates a case of counting cumulativity.

(2) An illustration of counting cumulativity

	Constraint A	Constraint B
→ [X]		*
[Y]	*	
[Z]		**
→ [W]	*	

In OT (Prince & Smolensky 1993/2004), this sort of situation is not predicted to arise because of strict domination of constraint rankings. In practice, however, we do seem to observe cases that are instantiated by the toy tableaux in (2). Such cases are often handled by positing a constraint that is violated if and only if a constraint is violated twice; i.e. OCP constraints (Leben 1973; McCarthy 1986 *et seq*), which prohibits the multiple occurrences of an identical segment/feature within a certain domain. This new constraint can be ranked above Constraint A to solve the ranking paradox. Analyses using local self-conjunction, in place of OCP constraints, have also been proposed (Alderete 1997; Blust 2012; Ito & Mester 2003).

In the context of sound symbolism, a pattern of counting cumulativity would hold if, when there is a sound S associated with meaning M, two or more instances of S evoke stronger images of M than a single instance of S. Sound symbolism, on the other hand, can be said to be non-cumulative, if one instance of the segment S suffices to signal meaning M, and the number of S does not matter.

The other kind of cumulativity, the ganging-up cumulativity, is illustrated by the toy tableaux in (3). Constraint A dominates both Constraint B and Constraint C, as [X] wins over [Y], and [Z] wins over [W]. However, given a candidate like [U] which violates both Constraint B and Constraint C, then this simultaneous violation can take precedence over Constraint A. In OT, these cases are often dealt with by positing a constraint that is violated if and only if both Constraint B and Constraint C are simultaneously violated within a certain domain, often in the form of constraint conjunction (Smolensky 1995 and subsequent works). As we will see below in detail, MaxEnt generally does away with the need for local conjunction because it can handle ganging-up cumulative constraint interactions (see in particular Zuraw & Hayes 2017; however, see also Shih 2017 who argues that there may be cases in which local conjunction is justified in MaxEnt grammar models).

(3) An illustration of ganging-up cumulativity

	Constraint A	Constraint B	Constraint C
→ [X]		*	
[Y]	*		
→ [Z]			*
[W]	*		
[U]		*	*
→[V]	*		

In the context of sound symbolism, a ganging-up cumulativity holds if segments S_1 and S_2 cause the same image M, and the image M is stronger when S_1 and S_2 co-occur than when S_1 and S_2 occur individually. On the other hand, sound symbolism is non-cumulative if one segment—either S_1 or S_2—determines that the word carries meaning M, and the presence of another segment does not affect the extent to which M is expressed.

5 Counting cumulativity in sound symbolism

This section summarizes reported cases of counting cumulativity in sound symbolism, and takes up two experimental studies which report quantitive data that can be analyzed using MaxEnt grammar. To recap, the crucial question is whether two or more instances of a segment, or a feature, can cause stronger sound symbolic images than one instance. There are some impressionistic reports that this can indeed be the case. First, Hamano (2013) reports that in the Tsugaru dialect of Japanese, there is an incremental increase in the strength of mimetic forms in proportion to the number of voiced obstruents contained in these expressions; e.g. [kata] < [gata] < [gada] and [kaki] < [gaki] < [gagi] (No glosses provided in the original source). Second, McCarthy (1983), citing Martin (1962), points out that tense consonants function as an intensifier in Korean ideophones, and we observe “greater intensity as the laryngeal features are expressed on both syllable-initial obstruents” (p. 144); e.g. [pancak] < [panc’ak], [p’ancak] < [p’anc’ak] ‘glittering.’

There is one study in psychology which directly addressed the question of cumulativity in sound symbolism; namely, Thompson & Estes (2011). They built upon the observations that some sounds are associated with images of largeness (e.g. Sapir 1929 *et seq.*). They presented to the participants—native speakers of British and American English—pictures of an imaginary creature (“greeble”) in different sizes, and different nonce names containing different numbers of “large phonemes.” They found that the larger the creature, the more likely it was for the judged nonce names to contain “large phonemes,” as shown in Figure 2 (reproduced from their Figure 3).

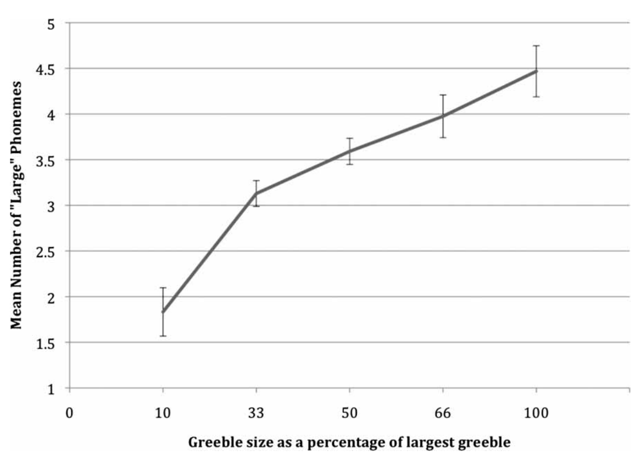

Figure 2: A cumulative sound symbolic effect found by Thompson and Estes 2011 (their Figure 3). The larger the size of the named objects, the more “large phonemes” were contained in their chosen names. This result is that of their Experiment 1, which used orthographic stimuli with native speakers of American English.

However, they collapsed three very different classes of sounds—back vowels, sonorants, and voiced stops—into one set of "large phonemes," and therefore it is impossible to tell whether this is a case of counting cumulativity or ganging-up cumulativity. Also, they counted the average numbers of large phonemes for each size condition (not vice versa), so it is hard to apply the sort of MaxEnt analyses presented below.

Another case study is reported by Kawahara et al. (2019), who analyzed the names of Takarazuka actress names. In Takarazuka, all actresses are biologically female, but some of them play a male role whereas other play a female role. Once they choose their Takarazuka gender, that gender is fixed throughout their career. Drawing on the previous observation that female names are more likely to contain sonorants than obstruents in Japanese (Shinohara & Kawahara 2013), Kawahara et al. (2019) show that the number of sonorants in the names positively correlates with the probability of those names being used for the female names (see Figure 3, reproduced from their Figure 1). While they collapsed a set of sounds into one class ("sonorant"), it is probably safe to consider this pattern to be a case of counting cumulativity, because we know from the previous literature that sonorants function as a natural class to represent femaleness in various languages (Perfors 2004; Shinohara & Kawahara 2013; Sidhu et al. 2019; Sullivan 2018; Wong & Kang 2019). Since Kawahara et al. (2019) have already developed a MaxEnt analysis, it will not be repeated here (although Kawahara et al. 2019 do not pay attention to the cumulative nature of this pattern). I thus only note here that this pattern seems to instantiate a good case of counting cumulativity in sound symbolism, and according to Kawahara et al. (2019), MaxEnt grammar can account for this pattern in a straightforward manner.

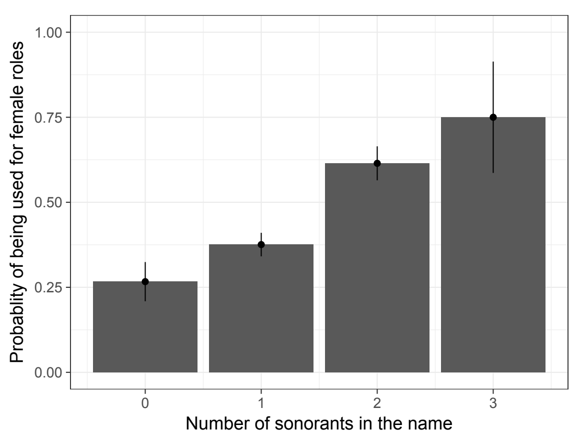

Figure 3: Cumulative sound symbolism found by Kawahara et al. (2019) (their Figure 1). The more sonorants are contained in the Takarazuka actress names, the more likely that the names are used for female roles.

A similar effect of a sonorant/obstruent distinction on name choices is experimentally examined by Kawahara (2012), targeting native speakers of English. This study tested how likely disyllabic nonce words containing different numbers of obstruents were judged to be male names, inspired by some of the references cited above. The stimulus conditions included OO, OS, and SS (where O stands for an obstruent, and S stands for a sonorant).[10] The stimuli were always disyllabic, and there were 10 vowel combinations, controlled across the different consonantal conditions. Two types of consonants were tested in each condition, resulting in 20 consonant-vowel combinations. The stimuli were auditory stimuli produced by native speakers of English. The tokens were resynthesized with a uniform falling pitch contour. The peak amplitude was adjusted to 0.7 Pa. For the judgment experiment, 25 native speakers of English listened to each stimulus in a randomized order, and judged whether it sounded like a male name or a female name. The results were that the male responses increased as the number of obstruents in nonce names increased: OO = 57.1%, OS = 48.2%, SS = 39.2%. Again, assuming that we can treat obstruents and sonorants as natural classes, it seems safe to conclude that this result instantiates a case of counting cumulativity: the higher the number of obstruents included in nonce words, the more likely they were judged to be male names.

Let us now build a MaxEnt analysis with OT-constraints to model this data, but before doing so, one caveat is in order. Since the structure of the constraints is rather simple (i.e. two constraints), I do not attempt to justify the inclusion of each constraint statistically, for example by way of log-likelihood tests (Breiss & Hayes to appear; Hayes et al. 2012; Shih 2017). To reiterate, my purpose is not to justify the existence of a particular constraint; it is instead to show that MaxEnt is a general, useful tool to model sound symbolic patterns, because it can capture a general nature of the observed data (i.e. cumulativity). My demonstration is therefore largely qualitative. As a quantitive measure of the fit between the observed and predicted measures, I report the Kullback-Leiber divergence (a.k.a. relative entropy) (Kullback & Leibler 1951), a measure of how one distribution (here predicted values) differs from another distribution (here the observed values).[11]

As stated above, the first step in developing a MaxEnt grammar analysis of sound symbolic patterns is to recognize that sound symbolic connections can be understood as mapping from sound (input) to meaning (output), just like phonology involves mapping from underlying representation to surface representation. To account for this sound symbolic pattern, I posit two constraints in (4). These constraints are similar to markedness constraints in Optimality Theory in that they only evaluate the wellformedness of output candidates (Prince & Smolensky 1993/2004).[12] The

[10]The experiment also included the SO condition, which is set aside here to simplify the discussion, because the response for this condition was very similar to that of the OS condition.

[11]Given two discrete probability distributions, this measure is calculated as $\sum_i P(i) log \frac{P(i)}{Q(i)}$, where P is the set of observed values and Q is the set of predicted values. Both the observed values and predicted values are rescaled so that they each sum to 1 and can be treated as treated as probability distributions. The closer the two distributions, the smaller this value.

[12]It is possible to posit constraints that militate against a particular mapping—not a surface structure—between

constraint formulation in this paper follows the format proposed by McCarthy (2003).

(4) Constraints for the analysis of the effects of obstruents/sonorants on gender

a. *OBSFEMALE: Assign a violation mark for each obstruent used in a female name.
b. *MALENAME: Assign a violation mark for each male name.

The first constraint reflects the preference for obstruents to be used in male names. The second constraint militates against male names in general. This second constraint is necessary so that male names receive some constraint violations. This constraint can be understood as belonging to the family of *STRUC constraint, a constraint banning a structure itself (Prince & Smolensky 1993/2004)—see especially Daland (2015) for the role of *STRUC constraint in MaxEnt analyses. The MaxEnt tableau appears in (5).

(5) A MaxEnt analysis of the effects of obstruents/sonorants on gender choices.

		w = 0.36	w = 0.44				
Input	Output	*OBS FEMALE	*MALE NAME	H-score	eHarmony	Predicted P	Observed P
SS	Female			0	1	60.8	60.8
	Male		1	0.44	0.64	39.2	39.2
OS	Female	1		0.36	0.70	51.8	51.9
	Male		1	0.44	0.64	48.2	48.1
OO	Female	2		0.72	0.49	65.7	42.9
	Male		1	0.44	0.64	34.3	57.1

It takes three types of inputs (OO, OS, and SS) and for each type of input, it calculates, based on the constraint weights and violation profiles, the predicted probabilities (percentages) of it being judged as female names and male names, which are shown in the "Predicted P" column.

The MaxEnt grammar tool (Hayes et al. 2009) found the optimum weights for the two constraints, given the observed percentages, which are shown at the top row in (5). We observe in the two rightmost columns in (5) that the percentages predicted by these weights match extremely well with the observed values. The Kullback-Leiber divergence is as small as 0.000001. This analysis shows that MaxEnt grammar can straightforwardly account for a counting cumulativity pattern in sound symbolism. What is particularly interesting about this analysis is the comparison between the Female candidates in the OS and OO conditions; since the Female candidate in the latter condi-

sounds and meanings, as in Kawahara et al. (2019).

tion violates *OBSFEMALE twice, it receives higher a H-score, and hence lower probability. This comparison illustrates how MaxEnt deals with counting cumulativity in sound symbolism. Viewed from the opposite angle, cumulativity is the default result in MaxEnt, and therefore it is suited as an analytical device of cumulative sound symbolic patterns.

One interesting aspect of MaxEnt grammar that this analysis reveals is that even if a candidate satisfies all the constraints (i.e. the SS names mapped onto female names), it is not predicted to get "all the share" i.e. 1.0 probability (Kawahara et al. 2019). This is because MaxEnt grammar calculates the probability distribution over all candidates that are considered; less than optimum candidates are assigned some non-zero probabilities, and hence even the perfect candidate is not assigned 1.0 probability. This nature of MaxEnt differs from OT: in OT, if there is a candidate that perfectly satisfies the whole constraint set, then it would harmonically bound the other candidates (Prince & Smolensky 1993/2004), and is predicted to always win. In Harmonic Grammar as well, if a candidate satisfies all the constraints, its H-Score is zero, and that candidate is deemed to win all the time. Even in its noisy version, with no constraint violations, no noise will be added so that its H-Score is predicted to be zero.

Another case of counting cumulativity in sound symbolism comes from an analysis of Pokémon names. Kawahara et al. (2018) found that in the existing set of Japanese Pokémon names, the evolved characters are more likely to contain voiced obstruents than non-evolved characters. Kawahara & Kumagai (2019) built on this observation and asked 58 native speakers of Japanese to judge whether particular nonce names are better suited for a pre-evolution or a post-evolution version of Pokémon. Their nonce word stimuli controlled for the number of voiced obstruents, varying from zero to two. The experiment used Japanese orthography to present the stimuli. Their results are reproduced in Figure 4, which instantiates a clear case of counting cumulativity: names with one voiced obstruent were more likely to be associated with post-evolution characters than names with no voiced obstruents (averages: 50.7 vs. 26.7); names with two voiced obstruents were judged to be more so than names with one voiced obstruent (averages: 65.7 vs. 50.7).

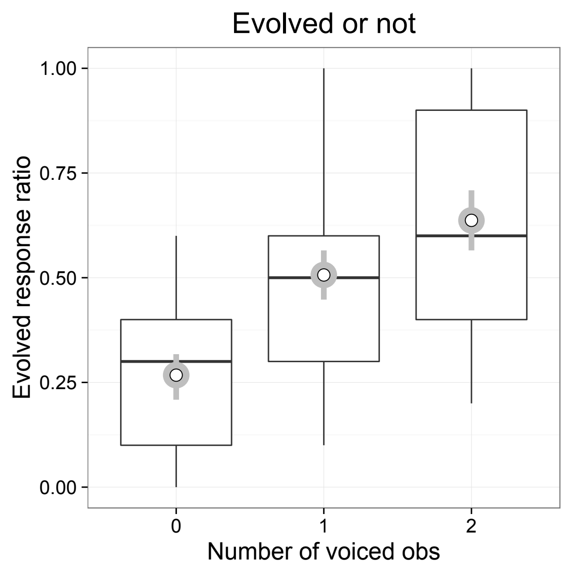

Figure 4: A case of counting cumulativity found by Kawahara & Kumagai (2019) (their Figure 4). The white circles represent the averages for each condition. The grey bars around the averages show 95% confidence intervals. The more voiced obstruents the nonce names contain, the more likely they were judged to be names of post-evolution Pokémon characters.

For this case at hand, I posit two constraints in (6).

(6) Constraints for the analysis of the Pokémon's evolution status

a. *VcdObsPreEvol: Assign a violation mark for each voiced obstruent used in a pre-evolution Pokémon character name.

b. *PostEvol: Assign a violation mark for each post-evolution Pokémon character name.

The first constraint reflects the tendency to use names with voiced obstruents for post-evolution characters.[13] The second constraint militates against evolved characters's names in general, which is again considered as an instance of *Struc constraint (Daland 2015).

The MaxEnt analysis tableau appears in (7). It takes three types of inputs (i.e. names with 0

[13]One might worry that this constraint refers to an abstract, complex and arguably Pokémon-specific notion like "evolution," to the extent that the set of constraint has to be universal, shared across all languages, as standardly assumed in Optimality Theory (Prince & Smolensky 1993/2004). This issue is related to the question of what notions can be symbolically represented in human language systems (Auracher 2017; Lupyan & Winter 2018; Westbury et al. 2018), and how universal they are. For the issue of universality and language specificity of sound symbolism, see for example a review by Imai & Kita (2014), and a recent paper by Bremner et al. (2013). For the current case, since post-evolution Pokémon characters are generally larger than pre-evolution Pokémons, the constraint can be formulated as referring to the notion of size rather than evolution itself. It is more likely that size is universally represented via sound symbolism (Shinohara & Kawahara 2016; Sidhu & Pexman 2018).

voiced obstruents, names with 1 voiced obstruent and names with 2 voiced obstruents), and for each type of input, it puts out the predicted probabilities (percentages) of how likely each name is assigned to a post-evolution category and a pre-evolution category.

(7) A MaxEnt analysis of Pokémon's evolution status (the counting cumulativity).

		w = 0.78	w = 0.91				
Input	Output	*VcdObsPreEvol	*PostEvol	H-score	eHarmony	Predicted P	Observed P
0 VcdObs	PostEvol		1	0.91	0.40	28.7	26.7
	PreEvol			0	1	71.3	73.3
1 VcdObs	PostEvol		1	0.91	0.40	46.8	50.7
	PreEvol	1		0.78	0.46	53.2	49.3
2 VcdObs	PostEvol		1	0.91	0.40	65.7	63.7
	PreEvol	2		1.56	0.21	34.3	36.3

The MaxEnt grammar tool found the optimum weights for the two constraints, given the observed percentages, and we observe in the two rightmost columns in (7) that the percentages predicted by these weights match very well with the observed values. The Kullback-Leiber divergence is 0.002. This analysis again demonstrates that MaxEnt grammar can straightforwardly account for counting cumulativity patterns. What is particularly relevant in the current analysis is how the MaxEnt grammar differentiates the two PreEvol candidates in the 1 VcdObs condition and the 2 VcdObs condition; since the candidate violates *VCDOBSPREEVOL twice in the second condition, it receives a higher H-score and is hence assigned lower probability. Counting cumulativity therefore naturally arises without further stipulations.

I also note that the effect of voiced obstruents is clearly sub-linear (averages: 26.7 vs. 50.7 vs. 63.7) in Figure 4: the slope between 0 and 1 is steeper than the one between 1 and 2. As MaxEnt involves a sigmoid function, it can account for this sub-linear pattern without stipulations.[14]

Before closing this section on counting cumulativity, a very similar sound symbolic effect of voiced obstruents on the evolution status in Pokémon names was identified by Godoy et al. (2019), who studied this issue with a free elicitation study and forced-choice judgment experiments targeting Brazilian Portuguese speakers. Their Experiment 3 had three conditions in which one member of a pair had no voiced obstruents, and the other member had either one, two, or three

[14]To be clear, MaxEnt can account for sub-linear, linear, and super-linear patterns, reflecting different portions of sigmoid curves (Breiss & Albright 2020). See Breiss & Albright (2020) and Kim (2019) for super-linear cumulative patterns in phonology.

voiced obstruents. They had 107 participants for this experiment, and the stimuli were presented in written Portuguese orthography. The rate in which the nonce names with voiced obstruents were associated with post-evolution characters increased as the number of voiced obstruents increased: 0 vs. 1: 55%; 0 vs. 2: 55.7%; 0 vs. 3: 63.3%. This pattern is very similar to the Japanese case analyzed in (7), so I will not repeat a MaxEnt analysis for this pattern. However, the results of this study show that the same counting cumulativity pattern holds for Brazilian Portuguese speakers.

6 Ganging-up cumulativity

We now turn to the cases of ganging-up cumulativity. One case study comes from another experiment on Pokémon names (Kumagai & Kawahara 2019).[15] The participants were 37 native speakers of Japanese. The experiment presented a pair of a pre-evolution character and a post-evolution character as well as a pair of nonce names written in Japanese orthography, and asked the participants to choose which name is better suited for which character; i.e. the experiment was presented in the 2 Alternative Forced Choice (2AFC) format. Based on two experiments, they found two generalizations: [a] is judged to be better for post-evolution character than [i], and voiced obstruents are judged to be better for post-evolution characters than voiceless obstruents. The results of their Experiment 2, which instantiate ganging-up cumulativity, are summarized in (8).

(8) Summary of the results of Kumagai & Kawahara (2019) (their Experiment 2). [p] stands for voiceless obstruents; [b] stands for voiced obstruents; [i] stands for two high vowels.

a. [pa] (as opposed to [pi]): 63% post-evolution.
b. [ba] (as opposed to [bi]): 58% post-evolution.
c. [pa] (as opposed to [bi]): 19% post-evolution.
d. [ba] (as opposed to [pi]): 90% post-evolution.

Conditions (a) and (b) show the effects of the vowel: [a] is better suited for post-evolution characters than a high vowel is. Comparing Conditions (a) and (d), we observe that voiced obstruents further increase the likelihood of the names being chosen as post-evolution characters. Accounting for the patterns in (8) requires only two constraints stated in (9).

(9) Constraints posited for the analysis of the Pokémon's evolution status

[15] Kumagai & Kawahara (2019) actually develop a MaxEnt analysis of their results using EXPRESS(X) constraints first proposed by Alderete & Kochetov (2017). Since their paper is written in Japanese, I believe that it is useful to present their reanalysis here. Although the analysis offered by Kumagai & Kawahara (2019) is similar to what is presented below, the way the constraints are formulated are different.

a. *[i]PostEvol: Assign a violation mark for each high vowel used in a post-evolution Pokémon name.

b. *VlsPostEvol: Assign a violation mark for each voiceless obstruent used in a post-evolution Pokémon name.

These constraints reflect the tendency to favor non-high vowels and voiced obstruents for post-evolution Pokémon characters. The MaxEnt analysis of this ganging-up cumulativity appears in (10).

(10) A MaxEnt analysis of Pokémon's evolution status (the ganging-up cumulativity).

			w = 0.41	w = 1.83				
Condition	Input	Output	*[i]PostEvol	*VlsPostEvol	H-score	eHarmony	Predicted P	Observed P
(a)	[pa]	PostEvol		1	1.83	0.16	60	63
	[pi]	PostEvol	1	1	2.24	0.11	40	37
(b)	[ba]	PostEvol			0	1	60	58
	[bi]	PostEvol	1		0.41	0.66	40	42
(c)	[pa]	PostEvol		1	1.83	0.16	19	19
	[bi]	PostEvol	1		0.41	0.66	81	81
(d)	[ba]	PostEvol			0	1	90	90
	[pi]	PostEvol	1	1	2.24	0.11	10	10

Like the analyses presented in section 5, the MaxEnt grammar tool found the optimum weights given the dataset provided in (8). The percentages predicted by these weights match very well with the observed percentages, demonstrating that MaxEnt grammar can account for a ganging-up cumulativity effect in sound symbolism. The Kullback-Leiber divergence is 0.007. Just like the two cases analyzed in section 5, the "perfect candidate" (the [ba]=PostEvol candidate in Condition (d)) is not assigned 1.0 probability. The cumulative nature of MaxEnt is most clearly observed in the behavior of the [pi]=PostEvol candidate in Condition (d)—this candidate violates both of the constraints, hence receives the highest H-score, and consequently, the lowest predicted probability.

The next example of ganging-up cumulativity comes from the experimental results reported by Ahlner & Zlatev (2010), which, as the authors admit, only partially support the cumulative nature of sound symbolism. Their empirical target is the *bouba-kiki* effect in which certain sounds are associated with angular objects, whereas other sounds are associated with round objects (see Figure 1: Ramachandran & Hubbard 2001). Their specific hypotheses were (i) [i] and voiceless stops are associated with angular objects, whereas (ii) [u] and sonorants are associated with round objects; like the experiment by Kumagai & Kawahara (2019) analyzed above, it was a 2 Alternative Forced Choice (2AFC) experiment. The participants were 20 native speakers of Swedish, and the stimuli were auditory stimuli. They presented different combinations of these sounds and asked whether each nonce name better matches with an angular shape or a round shape. Their results are summarized in (11).

(11) Summary of the results of Ahlner & Zlatev (2010).

a. [i]+sonorants = 90% angular vs. [u]+sonorants =10% angular.
b. [i]+voiceless stops = 80% angular vs. [i]+sonorants = 20% angular.
c. [u]+voiceless stops = 65% angular vs. [i]+sonorants = 35% angular.
d. [i]+voiceless stops = 90% angular vs. [u]+sonorants = 10% angular.

Condition (a) shows that [i] is better suited for angular objects than [u] is. Condition (b) shows that voiceless stops are better suited for angular objects than sonorants are. The comparison between Condition (a) ([u]+sonorants: 10% angular) and Condition (c) ([u]+voiceless stops = 65% angular) shows having sonorants lowers the angular responses; comparing Condition (b) (=[i]+voiceless stops: 80% angular) and Condition (c) ([u]+voiceless stops: 65% angular) shows that having back vowels lowers the angular responses. Ahlner & Zlatev (2010) thus conclude that "results showed that *both* vowels and consonants independently, and in combination, contribute to establishing the iconic ground in cross-modal iconicity (p. 329; emphasis in the original)." One complicating aspect of this data is its near-ceiling effect: comparing Conditions (a) and (d), we observe that the presence of [i] alone can make the nonce words angular-like, so much so that the effects of consonants are not visible.

Two constraints that are required to account for this pattern are listed in (12), and the MaxEnt tableaux are shown in (13).

(12) Constraints for the analysis for Ahlner & Zlatev's data

a. *BACKANGULAR: Assign a violation mark for each back vowel in a name for an angular shape.
b. *SONANGULAR: Assign a violation mark for each sonorant in a name for an angular shape.

(13) A MaxEnt analysis of Ahlner & Zlatev's data.

			w = 1.27	w = 1.58				
Condition	Input	Output	*BACKANGULAR	*SONANGULAR	H-score	eHarmony	Predicted P	Observed P
(a)	[i]+son	Angular		1	1.58	0.21	78.1	90
	[u]+son	Angular	1	1	2.85	0.058	21.9	10
(b)	[i]+vls	Angular			0	1	82.9	80
	[i]+son	Angular		1	1.58	0.21	17.1	20
(c)	[u]+vls	Angular	1		1.27	0.28	57.7	65
	[i]+son	Angular		1	1.58	0.21	42.3	35
(d)	[i]+vls	Angular			0	1	94.5	90
	[u]+son	Angular	1	1	2.85	0.06	5.5	10

As before, the analysis is generally successful, as observed in the two rightmost columns. Importantly, however, the MaxEnt analysis was not able to account for the near-ceiling effect: the angular shape in Condition (a) (=[i]+sonorants) receives lower predicted probability than the angular shape in Condition (d) (=[i]+voiceless stops). Consequently, the Kullback-Leiber divergence is larger than those in the previous analyses, although it is not very large (=0.02). This is inevitable because the former candidate's violation profile is a superset of that of the latter candidate. In other words, as long as the constraint *SONANGULAR is necessary, which indeed it is, the angular shape in Condition (a) receives lower predicted probability than the angular shape in Condition (d). This is not to say that MaxEnt cannot handle ceiling effects in general: Zuraw & Hayes (2017) demonstrate that it can—for the current case at hand, it is a near-ceiling effect, which is probably very difficult to handle. Ahlner & Zlatev (2010) themselves note (p. 330), however, that the number of the participants in their experiment was small, and that the true cumulative effect would have predicted that the angular response for the [i]+voiceless stops candidate in Condition (d) should be higher, as predicted by the current MaxEnt analysis.

The final case study of the ganging-up cumulativity comes from the study by D'Onofrio (2014). Her study targets the *bouba-kiki* effect, and she tested three phonological dimensions that may affect the judgement of shapes: vowel backness, consonant voicing, and three places of articulation. This case is particularly interesting, because all the factors are fully-crossed, thus instantiating what Zuraw & Hayes (2017) refer to as "intersecting constraint families," as schematically illustrated in Figure 5. Each dimension dictates only one phonological dimension: i.e. x-axis = consonant voicing, y-axis = consonant place of articulation, z-axis = vowel quality. Each cell occurs at the intersection of these three dimensions.

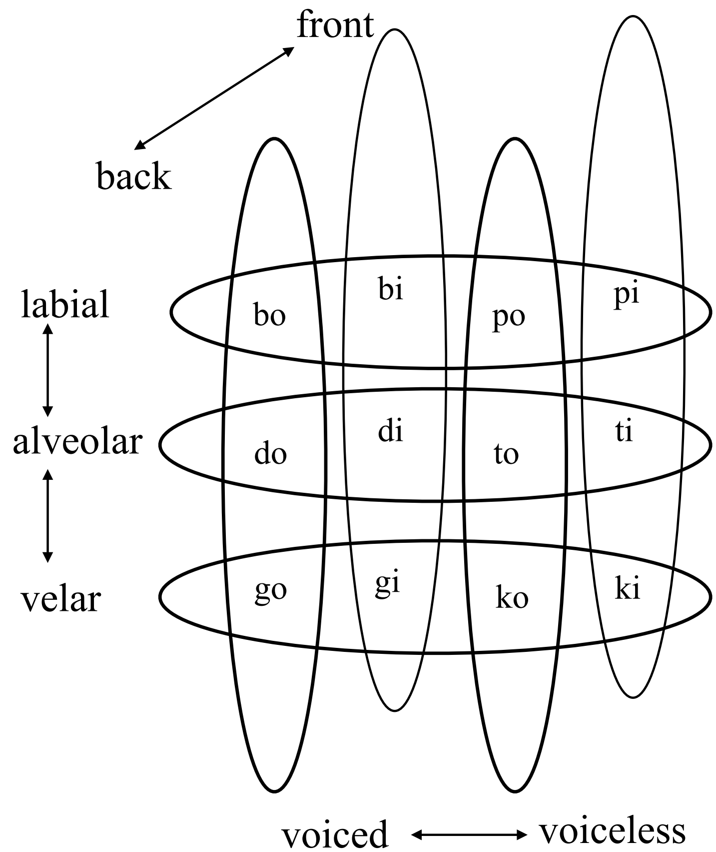

Figure 5: A schematic illustration of 3-dimensional intersecting constraint families.

Since all the factors are fully crossed, there are 12 conditions ($3 \times 2 \times 2$). An interesting question raised by Zuraw & Hayes (2017) is whether given such cases, we can derive these 12 individual patterns from constraints that each regulate only one phonological dimension. To borrow their words, rephrased partially to fit the current three-dimensional case, "[o]n standard scientific grounds, we would hardly want to set up a grammar that stipulates the outcome rate for each cell separately, with [twelve] separate parameters. Rather, we would prefer a system that assigns the right properties (rankings or weights) to each [dimension], and lets the behavior of the individual cells follow from the general theory of constraint interaction" (p. 498).[16] The analysis below shows that we can indeed set up a grammatical system which has this property.

D'Onofrio (2014) generally found that (i) back vowels are associated with round shapes; (ii) voiced stops are more likely to be associated with round shapes than voiceless stops; (iii) alveolars are associated with angular shapes, and labials and velars are associated with round shapes. The participants were 170 native speakers of English, collected via Amazon Turk. The stimuli were

[16]In their original passage, they consider a 3-by-3 two dimensional space; therefore they had "nine" in place of "twelve" and "row and column" in place of "dimension."

presented auditorily. Her results are summarized in Table 1, which shows that voiced consonants, labial and velar consonants, and back vowels tend to increase round responses in a cumulative way.

Table 1: The summary results of D'onofrio (2014) (based on her Table 2). %Round responses are shown in the rightmost column. Voiced consonants, labial and velar consonants, and back vowels tend to increase round responses.

Voicing	Backness	Place	%Round
Voiced	Back	Velar	91
Voiced	Back	Labial	82
Voiceless	Back	Labial	80
Voiced	Front	Labial	76
Voiced	Front	Velar	68
Voiced	Back	Alveolar	67
Voiceless	Front	Labial	64
Voiced	Front	Alveolar	55
Voiceless	Back	Velar	55
Voiceless	Back	Alveolar	41
Voiceless	Front	Velar	15
Voiceless	Front	Alveolar	7

The constraints posited to account for the patterns in Table 1 are listed in (14). Each constraint refers to one phonological dimension and does not refer to interactions between more than one phonological dimension.

(14) Constraints posited for the MaxEnt analysis of D'onofrio's results

a. *VCDANGULAR: Assign a violation mark for each voiced stop in a name for an angular object.
b. *BACKANGULAR: Assign a violation mark for each back vowel in a name for an angular object.
c. *LABANGULAR : Assign a violation mark for each labial in a name for an angular object.
d. *ALVANGULAR: Assign a violation mark for each alveolar in a name for an angular object
e. *VELANGULAR : Assign a violation mark for each velar in a name for an angular object.
f. *ROUND: Assign a violation mark for each round object.

The MaxEnt analysis appears in (15), which is again successful in that the predicted percentages and the observed percentages match very closely. Since these are large tableaux, Figure 6 plots the correlation between the observed percentages and predicted percentages, which shows that the correlation is robust. The Kullback-Leiber divergence is 0.02. To reiterate, each constraint in (14) regulates a sound symbolic mapping for one phonological dimension only, and MaxEnt grammar is able to model how these constraints shape the properties of each three dimensional cell. The model is successful largely because cumulativity is the default result for MaxEnt models, which is the property that the data in Table 1 show.

(15) A MaxEnt analysis of D'onofrio's results

		w = 1.65	w = 1.43	w = 5.24	w = 4.03	w = 4.78	w = 5.95				
Input	Output	*VCD ANGULAR	*BACK ANGULAR	*LAB ANGULAR	*ALV ANGLULAR	*VEL ANGULAR	*ROUND	H-score	eHarmony	Predicted P	Observed P
VcdBackVel	round						1	5.95	0.0026	87.1	91
	angular	1	1			1		7.86	0.00039	12.9	9
VcdBackLab	round						1	5.95	0.0026	91.4	82
	angular	1	1	1				8.32	0.00024	8.6	18
VlsBackLab	round						1	5.95	0.0026	67.1	80
	angular		1	1				6.67	0.0013	32.9	20
VcdFrontLab	round						1	5.95	0.0026	71.8	76
	angular	1		1				6.89	0.001	28.2	24
VcdFrontVelar	round						1	5.95	0.0026	61.8	68
	angular	1				1		6.43	0.0016	38.2	32
VcdBackAlv	round						1	5.95	0.0026	76.1	67
	angular	1	1		1			7.11	0.0008	23.9	33
VcdFrontLab	round						1	5.95	0.0026	71.8	64
	angular	1		1				6.89	0.001	28.2	36
VcdFrontAlv	round						1	5.95	0.0026	43.2	55
	angular	1			1			5.68	0.0034	56.8	45
VlsBackVel	round						1	5.95	0.0026	56.5	55
	angular		1			1		6.21	0.0020	43.5	45
VlsBackAlv	round						1	5.95	0.0026	37.9	41
	angular		1		1			5.46	0.0043	62.1	59
VlsFrontVel	round						1	5.95	0.0026	23.7	15
	angular					1		4.78	0.0084	76.3	85
VlsFrontAlv	round						1	5.95	0.0026	12.8	7
	angular				1			4.03	0.018	87.2	93

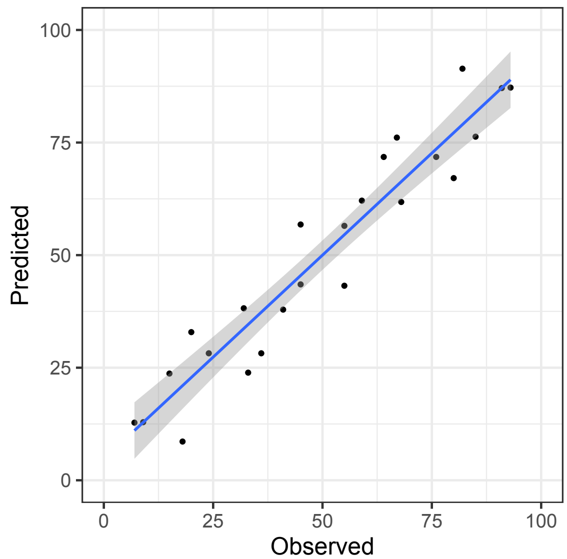

Figure 6: The correlation between observed and predicted values in the MaxEnt analysis of D'onofrio's results.

7 Conclusion

There have been few studies which directly addressed the question of whether sound symbolic patterns in natural languages are cumulative or not, with the notable exceptions being Ahlner & Zlatev (2010) and Thompson & Estes (2011). I argued that this issue is nevertheless important to explore in detail, not only because it may reveal an important nature of sound symbolism, but also because it may reveal an interesting parallel between sound symbolic patterns and phonological patterns. To that end, as one descriptive goal of this paper, I reviewed cases from a variety of languages which seem to support the thesis that sound symbolic patterns are cumulative.

However, it is too premature to conclude that all sound symbolic patterns show a cumulative nature; for example, in affrication found in Japanese baby-talk register, one instance of palatalization may make the whole utterance a baby-talk, so much so that the number of affricated segments does not matter (although this report is only based on intuitions of a few native speakers and not examined quantitatively). Sawada (2013), who presents an extensive analysis of the semantics of this diminutive register, discusses the following pair of examples:

(16)
  a. oiʃii de**tʃ**uka? (one affrication)
  b. oi**tʃ**ii de**tʃ**uka? (two affrication)
  c. (oiʃii desuka? "Is it yummy?": none-diminutive)

In (a), affrication is expressed only on the polite suffix; in (b) on the other hand, affrication appears twice, once on the adjective and once on the polite suffix. In this pair of examples, it seems to be the case that "once diminutive, all diminutive"—both (a) and (b) are equally diminutive (p.c. Osamu Sawada, Dec 2019). If this observation is correct, then this pattern instantiates a case of non-cumulative sound symbolism.

I also note that the set of languages that have been examined so far is also limited: Brazilian Portuguese (Godoy et al. 2019), English (D'Onofrio 2014; Thompson & Estes 2011), Japanese (Kawahara & Kumagai 2019; Kawahara et al. 2019; Kumagai & Kawahara 2019), Swedish (Ahlner & Zlatev 2010), and perhaps Korean (McCarthy 1983). A more extensive cross-linguistic study exploring the (non-)cumulative nature of sound symbolism is thus hoped for.

With these caveats in mind, however, taken together with a growing body of evidence that phonological patterns also show cumulative aspects (as reviewed in section 2), the current results suggest that sound symbolic patterns and phonological patterns are more similar to each other than hitherto assumed in that they both show cumulative patterns. On the theoretical side, I have shown that MaxEnt grammar with Optimality Theoretic constraints, which has been shown to be successful in modeling various phonological patterns, is also successful in accounting for sound-meaning mappings—this is so in the current context, largely because the default outcome of MaxEnt grammar is cumulative, and is thus able to account for this general nature of sound symbolic patterns. In a recent study, Westbury et al. (2018) propose that we should study *weightings* of phonological features/segments in their contribution to their sound symbolic effects, and this is precisely what MaxEnt analyses attempt to do. I understand this to mean that psychologists and phonologists have a shared interest in a non-trivial sense, and we are converging on the same conclusion regarding what is important to study. With this convergence, it is possible that insights offered from recent phonological studies—e.g. we may be able to derive cross-linguistic differences from different weightings of the same set of constraints (Prince & Smolensky 1993/2004) and the set of constraints may be induced from phonetic considerations (Hayes 1999)—may bring in a new perspective for studies of sound symbolism. I hope that these overall results pique theoretical phonologists' interests to study more about sound symbolism, and that researchers who work on sound symbolism find the formalism that theoretical phonologists employ to be useful.

References

Ahlner, Felix & Jordan Zlatev. 2010. Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. *Sign Sytems Studies* 38(1/4). 298–348.

Albright, Adam. 2012. Additive markedness interactions in phonology. Talk presented at UCLA. `https://tinyurl.com/vpdfju9`.

Alderete, John. 1997. Dissimilation as local conjunction. In Kiyomi Kusumoto (ed.), *Proceedings of the North East Linguistics Society 27*, 17–31. Amherst: GLSA.

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

Auracher, J. 2017. Sound iconicity of abstract concepts: Place of articulation is implicitly associated with abstract concepts of size and social dominance. *PLoS ONE* 12(11). e0187196.

Bagemihl, Bruce. 1995. Language games and related areas. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 697–712. Cambridge and Oxford: Blackwell.

Bailey, Todd & Ulrike Hahn. 2001. Determinants of wordlikeliness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 44. 568–591.

Bakovic, Eric. 2000. *Harmony, dominance, and control*: Rutgers University Doctoral dissertation.

Bergen, Benjamin K. 2004. The psychological reality of phonaesthemes. *Language* 80. 290–311.

Berger, Adam L, Stephen A. Della Pietra & Vincent J. Della Pietra. 1996. A maximum entropy approach to natural language processing. *Computational Linguistics* 22. 39–71.

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

Blust, Robert. 2012. One mark per word? Some patterns of dissimilation in Austronesian and Australian languages. *Phonology* 29(3). 355–381.

Boersma, Paul. 1998. Spreading in functional phonology. *Proceedings of the International Congress of Phonetic Sciences* 22. 1–20.

Boersma, Paul & Joe Pater. 2016. Convergence properties of a gradual learning algorithm for harmonic grammar. In John J. McCarthy & Joe Pater (eds.), *Harmonic Grammar and Harmonic Serialism*, 389–434. London: Equinox.

Breiss, Canaan. 2019. Cumulativity by default in phonotactic learning: experimental and computational evidence. Ms. University of California, Los Angeles. `https://ling.auf.net/lingbuzz/004747`.

Breiss, Canaan & Adam Albright. 2020. When is a gang effect more than the sum of its parts. Talk presented at LSA 2020.

Breiss, Canaan & Bruce Hayes. to appear. Phonological markedness effects in sentential formation. *Language* .

Bremner, Andrew J., Serge Caparos, Jules Davidoff, Jan de Fockert, Karina J. Linnell & Charles Spence. 2013. "bouba" and "Kiki" in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition* 126. 165–172.

Chomsky, Noam & Moris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Churma, Don. 1979. *Arguments from external evidence in phonology*: Ohio State University Doctoral dissertation. Published by Garland Press, New York, 1985.

Coetzee, Andries W. 2012. Variation: Where laboratory and theoretical phonology meet. In Abigail Cohn, Cécile Fougeron & Marie K. Huffman (eds.), *Oxford handbook of laboratory phonology*, 62–75. Oxford: Oxford University Press.

Coetzee, Andries W. & Joe Pater. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26. 289–337.

Coetzee, Andries W. & Joe Pater. 2011. The place of variation in phonological theory. In John A. Goldsmith, Jason Riggle & Alan Yu (eds.), *The handbook of phonological theory, 2nd edition*, 401–431. Oxford: Blackwell-Wiley.

Coleman, John & Janet Pierrehumbert. 1997. Stochastic phonological grammars and acceptability. In *Computational phonology: Third meeting of the ACL special interest group in computational* *phonology.*, 49–56. Somerset: Association for Computational Linguistics.

Crowhurst, Megan. 2011. Constraint conjunction. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 1461–1490. Oxford: Blackwell-Wiley.

Daland, Robert. 2015. Long words in maximum entropy phonotactic gammars. *Phonology* 32(3). 353–383.

de Lacy, Paul. 2009. Phonological evidence. In Steve Parker (ed.), *Phonological Argumentation: Essays on Evidence and Motivation*, 43–77. London: Equinox.

Della Pietra, Stephen, Vincent J. Della Pietra & John D. Lafferty. 1997. Inducing features of random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19. 380–393.

Dingemanse, Mark. 2018. Redrawing the margins of language: Lessons from research on ideophones. *Glossa* 3.

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.

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

Featherston, Sam. 2005. The decathlon model of empirical syntax. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 187–208.

Frisch, Stephan, Janet Pierrehumbert & Michael Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22. 179–228.

Fukazawa, Haruka, Shigeto Kawahara, Mafuyu Kitahara & Shin-ichiro Sano. 2015. Two is too much: Geminate devoicing in Japanese. *Phonological Studies* 18. 3–10.

Godoy, Mahayana C., Neemias Silva de Souza Filho, Juliana G. Marques de Souza, Hális Alves & Shigeto Kawahara. 2019. Gotta name'em all: An experimental study on the sound symbolism of Pokémon names in Brazilian Portuguese. *Journal of Psycholinguistic Research* .

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.

Gouskova, Maria. 2013. A review of "The Blackwell Companion to Phonology". *Phonology* 30(1). 173–179.

Green, Christopher R. & Stuart Davis. 2014. Superadditivity and limitations on syllable complexity in Bambara words. In Ashely W. Farris-Trimble & Jessica Barlow (eds.), *Perspectives on phonology, theory, and acquisition: Papers in honor of Daniel A. Dinnsen*, 223–247. Amsterdam: John Benjamins.

Guy, Gregory & C. Boberg. 1997. Inherent variability and the obligatory contour principle. *Language Variation and Change* 9. 149–164.

Hamano, Shoko. 1996. *The sound-symbolic system of Japanese*. Stanford: CSLI Publications.

Hamano, Shoko. 2013. Hoogen-ni okeru giongo-gitaigo-no taikeiteki-kenkyuu-no igi. In Kazuko Shinohara & Ryoko Uno (eds.), *Chikazuku oto-to imi: Onomatope kenkyuu-no shatei*, Tokyo: Hitsuzi Syobo.

Hay, Jennifer, Janet Pierrehumbert & Mary Beckman. 2003. Speech perception, well-formedness, and the statistics of the lexicon. In John Local, Richard Ogden & Rosalind Temple (eds.), *Papers in laboratory phonology VI: Phonetic interpretation*, 58–74. Cambridge: Cambridge University 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. 2017. Varieties of noisy harmonic grammar. *Proocedings of AMP* .

Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.). 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.

Hayes, Bruce & Zsuzsa Londe. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23. 59–104.

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

Hayes, Bruce, Colin Wilson & Ben George. 2009. Maxent grammar tool. Software. `http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/`.

Hayes, Bruce, Colin Wilson & Anne Shisko. 2012. Maxent grammars for the metrics of Shakespeare and Milton. *Language* 88(4). 691–731.

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.

Hyman, Larry M. 2018. Why underlying representations? *Journal of Linguistics* 21. 591–610.

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

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

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

Jakobson, Roman. 1971. Why 'mama' and 'papa'? In *Phonological studies 2nd. vol 1 of Roman Jakobson Selected Writings*, 538–545. The Hague/Paris: Mouton.

Jakobson, Roman. 1978. *Six lectures on sound and meaning*. Cambridge: MIT Press.

Jaynes, Edwin T. 1957. Information theory and statistical mechanics. *Physical Review* 106. 620–630.

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

Jurafsky, Daniel & James H. Martin. 2019. *Speech and language processing (3rd edition, draft)*. `https://web.stanford.edu/~jurafsky/slp3/`.

Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82(3). 536–574.

Kawahara, Shigeto. 2011. Japanese loanword devoicing revisited: A rating study. *Natural Language and Linguistic Theory* 29(3). 705–723.

Kawahara, Shigeto. 2012. Acoustic bases of sound symbolism. Talk presented at Seoul National University. `https://tinyurl.com/yx4kuect`.

Kawahara, Shigeto. 2019. Sound symbolism and theoretical phonology. Ms. Keio University. `https://ling.auf.net/lingbuzz/004885`.

Kawahara, Shigeto, Hironori Katsuda & Gakuji Kumagai. 2019. Accounting for the stochastic nature of sound symbolism using Maximum Entropy model. *Open Linguistics* 5. 109–120.

Kawahara, Shigeto & Gakuji Kumagai. 2019. What voiced obstruents symbolically represent in

Japanese: Evidence from the Pokémon universe. Ms. Keio University and Meikai University. http://user.keio.ac.jp/~kawahara/pdf/pokemonParametersRevised.pdf.

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

Kawahara, Shigeto & Shin-ichiro Sano. 2016. /p/-driven geminate devoicing in Japanese: Corpus and experimental evidence. *Journal of Japanese Linguistics* 32. 57–77.

Kellar, Frank. 2006. Linear Optimality Theory as a model of gradience in grammar. In Gisbert Fanselow, Caroline Féry, Ralf Vogel & Matthias Schlesewsky (eds.), *Gradience in grammar: Generative perspectives*, 270–287. Oxford: Oxford University Press.

Kim, Seoyoung. 2019. Modeling super-gang effects in MaxEnt: Nasal in Rendaku. Talk presented at NELS 50.

Kirchner, Robert. 1996. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27. 341–350.

Köhler, Wolgang. 1947. *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: Liveright.

Kullback, S. & R.A. Leibler. 1951. On information and sufficienty. *Annuals of Mathematical Statistics* 22. 79–86.

Kumagai, Gakuji & Shigeto Kawahara. 2019. Effects of vowels and voiced obstruents on Pokémon names: Experimental and theoretical approaches [in Japanese]. *Journal of the Linguistic Society of Japan* 155. 65–99.

Leben, Will. 1973. *Suprasegmental phonology*: MIT Doctoral dissertation.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. 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*, 388–395. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Lubowicz, Anya. 2002. Derived environment effects in Optimality Theory. *Lingua* 112(4). 243–280.

Lupyan, Gary & Bodo Winter. 2018. Language is more abstract than you think, or, why aren't languages more iconic? *Proceedings of Royal Society B.* 373. 20170137.

Martin, Samuel. 1962. Phonetic symbolism in Korean. In N. Poppe (ed.), *American studies in Uralic and Altaic linguistics*, Indiana University Press.

McCarthy, John J. 1983. Phonological features and morphological structure. In J. Richardson, M. Marks & A. Chukerman (eds.), *Proceedings from the parasesion on the interplay of phonology, morphology and syntax*, 135–161. Chicago: CLS.

McCarthy, John J. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17. 207–263.

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

Mester, Armin & Junko Ito. 1989. Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65. 258–293.

Nuckolls, Janis B. 1999. The case for sound symbolism. *Annual Review of Anthropology* 28. 225–252.

Padgett, Jaye. 2002. Constraint conjunction versus grounded constraint subhierarchies in Optimality Theory. Ms. University of California, Santa Cruz (ROA-530).

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33. 999–1035.

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.

Pierrehumbert, Janet B. 2001. Stochastic phonology. *GLOT* 5. 1–13.

Pizzo, Presley. 2015. *Investigating properties of phonotactic knowledge through web-based experimentation*: University of Massachusetts, Amherst Doctoral dissertation.

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

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

Rose, Sharon & Lisa King. 2007. Speech error elicitation and co-occurrence restrictions in two Ethiopian Semitic languages. *Language and Speech* 50(4). 451–504.

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.

Sawada, Osamu. 2013. The meanings of diminutive shifts in Japanese. *Proceedings of North East Linguistic Society* 42. 163–176.

Shih, Stephanie S. 2017. Constraint conjunction in weighted probabilistic grammar. *Phonology* 243–268.

Shih, Stephanie S. 2019. The contribution of sound symbolic evidence to lexically-conditioned phonology. Talk presented at NELS 50. `https://stephsus.github.io/Shih_NELS2019.pdf`.

Shinohara, Kazuko & Shigeto Kawahara. 2013. The sound symbolic nature of Japanese maid names. *Proceedings of the 13th Annual Meeting of the Japanese Cognitive Linguistics Association* 13. 183–193.

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

Sidhu, David & Penny M. Pexman. 2018. Five mechanisms of sound symbolic association. *Psychonomic Bulletin & Review* 25(5). 1619–1643.

Sidhu, David M., Kristen Deschamps, Joshua S. Bourdage & Penny M. Pexman. 2019. Does the name say it all? Investigating phoneme-personality sound symbolism in first names. *Journal of Experimental Psychology: General* 148(9). 1595–1614.

Smith, Brian W. & 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 D. Rumelhart, J. McClelland & PDPR Group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 1: Foundations, 194–281. Cambridge, MA: Bradford Boooks/MIT Press.

Smolensky, Paul. 1995. On the internal structure of the constraint component CON of UG. Talk

presented at the University of California, Los Angeles (ROA-86).
Sullivan, Lisa. 2018. *Phonology of gender in French and English names*: University of Toronto Master's thesis.
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.
Ultan, Russell. 1978. Size-sound symbolism. In Joseph Greenberg (ed.), *Universals of human language II: Phonology*, 525–568. Stanford: Stanford University Press.
Vance, Timothy. 2015. Rendaku. In Haruo Kubozono (ed.), *The handbook of Japanese language and linguistics: Phonetics and phonology*, 397–441. Berlin: Mouton de Gruyter.
Westbury, Chris, Geoff Hollis, David M. Sidhu & Penny M. Pexman. 2018. Weighting up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language* 99. 122–150.
Wilson, Colin. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30(5). 945–982.
Wilson, Colin. 2014. Maximum entropy models. Tutorial presentation, MIT.
Wong, Kristen Wing Yan & Yoonjung Kang. 2019. Sound symbolism of gender in Cantonese first names. *Proceedings of ICPhS* .
Zuraw, Kie & Bruce Hayes. 2017. Intersecting constraint families: An argument for Harmonic Grammar. *Language* 93. 497–548.