Accounting for the stochastic nature of sound symbolism using Maximum Entropy model

Abstract

Sound symbolism refers to stochastic and systematic associations between sounds and meanings. Sound symbolism has not received much serious attention in the generative phonology literature, perhaps because most if not all sound symbolic patterns are probabilistic. Building on the recent proposal to analyze sound symbolic patterns within a formal phonological framework (Alderete & Kochetov 2017), this paper shows that MaxEnt grammars allow us to model stochastic sound symbolic patterns in a very natural way. The analyses presented in the paper show that sound symbolic relationships can be modeled in the same way that we model phonological patterns. We suggest that there is nothing fundamental that prohibits formal phonologists from analyzing sound symbolic patterns, and that studying sound symbolism using a formal framework may open up a new, interesting research domain. The current study also reports two hitherto unnoticed cases of sound symbolism, thereby expanding the empirical scope of sound symbolic patterns in natural languages.

1   Introduction

In recent linguistic theories, it is almost standard to assume that the relationships between sounds and meanings are arbitrary. This thesis dates back to Hermogenes's view expressed in Plato's *Cratylus*, and it was very clearly articulated by Saussure (1916) as the first organizing principle of natural languages. The thesis of arbitrariness was reiterated as one of the design features of human languages by Hockett (1959). Few modern linguists would disagree to the thesis that language is a system that is able to associate sounds and meanings in arbitrary ways.

On the other hand, many studies have revealed certain systematic relationships between sounds and meanings, an observation which is often referred to as "sound symbolism" (see Dingemanse et al. 2015; Hinton et al. 2006; Lockwood and Dingemanse 2015; Sidhu and Pexman 2017; Spence 2011 for recent reviews). For example, Sapir (1929) and many subsequent studies have shown that low vowels tend to be judged to be larger than high vowels, and front vowels tend to be judged to be smaller than back vowels (Berlin, 2006; Coulter and Coulter, 2010; Jakobson, 1978; Jespersen, 1922; Newman, 1933; Ohala, 1994; Shinohara and Kawahara, 2016; Ultan, 1978). In both English and Japanese, there are stochastic tendencies for sonorants to be associated with female names and for obstruents to be associated with male names (Shinohara and Kawahara, 2013; Uemura, 1965; Wright and Hay, 2002; Wright et al., 2005). It is probably safe to conclude based on these studies that there are tendencies in natural languages that certain sounds are associated with certain meanings.

These sound symbolic patterns are, crucially, stochastic or probabilistic. On the one hand, speakers of many language feel that [a] is larger than [i] or [ɪ] (Shinohara and Kawahara, 2016), and this association is reported to hold in the lexicon of many languages (Ultan, 1978). For instance, diminutive affixes, including English "-y," often contain a high front vowel in a number of languages (Ultan, 1978). However, in no way is this association exception-less. For example the English word *big* contains the "small vowel", [ɪ] (see also Diffloth 1994). There is a sense in which [a] is felt to be bigger than [i], but few would argue that languages cannot use [a] to represent something small or cannot use [i] to represent something big. Similarly, although sonorants are often associated with female names, there are male names that contain sonorants and female names that contain obstruents. If sound symbolic relationships were deterministic, then as Locke (1689) and Saussure (1916) noted, all languages should use the same sound sequences to represent the same objects (or concepts).[1]

While studies of sound symbolism are flourishing both in psychology and pho-

---

[1] However, we also need to take into consideration the fact that different languages use different sets of sounds, and they have different phonotactic restrictions (see Shih et al. 2018 and Styles and Gawne 2017 for the implication of these cross-linguistic differences on sound symbolism). The set of real world attributions being referred to (or more simply, the set of semantic denotations) may differ across languages as well (Shih et al., 2018).

netics (Dingemanse et al., 2015; Hinton et al., 2006; Lockwood and Dingemanse, 2015; Sidhu and Pexman, 2017; Spence, 2011), sound symbolism has not received serious attention in the generative phonology literature. One clear exception is Alderete and Kochetov (2017), who developed a formal theory of sound symbolism, using a set of Optimality Theoretic constraints (Prince and Smolensky, 1993/2004), EXPRESS(X). This set of constraints require certain phonological feature X to be realized to signal certain semantic features (see also Kochetov and Alderete 2011). EXPRESS(X) is suited to account for non-probabilistic patterns driven by sound-symbolic principles such as palatalization in baby talk in Japanese and other languages. However, most if not all sound symbolic relationships are probabilistic, as discussed above. To this end, this paper expands on Alderete and Kochetov's proposal and shows that using Maximum Entropy model (MaxEnt: Goldwater and Johnson 2003) successfully accounts for stochastic aspects of sound symbolic patterns. The current proposal is illustrated with four case studies.[2] In addition to this theoretical contribution, two of our case studies report new empirical data, thereby expanding our knowledge of sound symbolism.

2   Names of Takarazuka Revue actresses

The first case study is based on a new set of empirical data, which comes from the names of Japanese Takarazuka Revue actresses. All Takarazuka actresses are female, but some actresses play a male role, while some others play a female role. Importantly, the actresses' gender in Takarazuka context is fixed; i.e. one actress cannot act both as male and female during her Takarazuka career. In this study, we first explored the effects of sound symbolic relationships between female names and sonorants on the one hand, and male names and obstruents on the other, sound symbolic relationships which have been shown to hold in Japanese (Shinohara and Kawahara, 2013; Uemura, 1965) as well as in English (Wright and Hay, 2002; Wright et al., 2005).

---

[2]A similar MaxEnt analysis of sound symbolism was also developed to model recent experimental results by Authors (to appear).

2.1 Method

All of the 361 names of Takarazuka actresses, who were active as of July 2017, were analyzed. For each name, we coded the number of sonorants and obstruents, as well as whether that name is used for a male role or a female role. Each name contained at most three consonants.
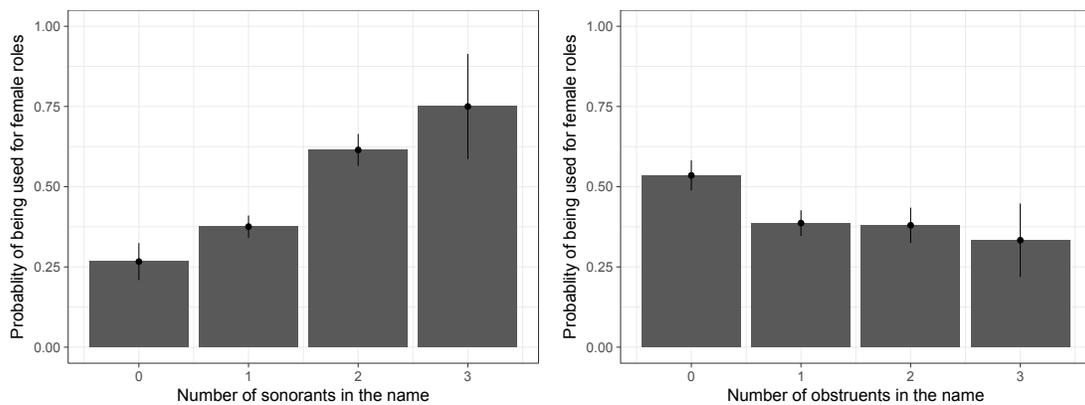
2.2 Result

Figure 1: The effects of sonorants and obstruents on the gender choice in Takarazuka names. The error bars are standard errors.

The results appear in Figure 1, in which the y-axes represent the probability of the names being used for female actresses. The left panel shows that the more sonorants a name contains, the more likely it is used as a female name: the right panel shows that although the tendency is less clear, the more obstruents a name contains, the less likely it is used as a female name. These patterns accord well with the sound symbolic effects previously noted for Japanese and English.

A multiple logistic regression analysis was run with gender as the dependent variable and the number of sonorants and the number of obstruents as independent variables. The result shows that sonorants significantly increase the likelihood of the name being used for a female name ($\beta_1 = 0.822, z = 4.19, p < .001$), while obstruents do not have significant impacts on the gender choice ($\beta_2 = 0.05, z = 0.36, n.s.$).

This analysis thus suggests that there is a stochastic tendency in such a way that sonorants

are associated with female roles. Crucially, though, the effects are not deterministic; it is not the case that the presence of a single sonorant makes the whole name a female name 100% of the time; instead, the presence of a sonorant increases the probability of that name being used as a female name. This sort of stochastic pattern is common-place—in fact, usually the norm—in sound-symbolic patterns in natural languages.[3] While the proposal by Alderete and Kochetov (2017) successfully accounts for deterministic sound symbolic patterns, it does not account for this sort of stochastic nature of sound symbolic patterns; we thus expand on their proposal to account for the gradient nature of sound symbolism, using MaxEnt grammar model.

2.3 MaxEnt analysis

Generative phonology is a function that maps one representation (e.g. underlying representation) to another representation (e.g. surface representation) (McCarthy, 2010); this is true, regardless of whether it is ruled-based or constraint-based, or whether it involves serial derivation or not. The model proposed in this paper is the same: for the case analyzed in this section, our grammar is a function that takes names as inputs and maps them to probability distributions of two candidates, male names or female names. For example, it takes a name like *Yurino*, and calculates the probability of that name being used for a female name and the probability of the name being used for a male name.

The current model uses MaxEnt grammar, as it is suited to account for gradient mapping from one dimension to another (e.g. Goldwater and Johnson 2003; Hayes 2017; Hayes and Wilson

---

[3]This stochasticity may be one of the reasons why sound symbolism did not receive serious attention in the generative phonology literature. In generative phonology (Chomsky and Halle, 1968), it was standard to assume that elements and structures that are *completely* predictable are derived in the phonological component of grammar, although in practice some exceptions are usually tolerated; on the other hand, those that are not completely predictable—whose predictability was lower than 1.00—were assumed to be stored in the lexicon (see Shaw and Kawahara 2018 for a historical overview). This means that phonological patterns should be exception-less, at least at some level of representation or analysis (see Shattuck-Hufnagel 1986 for relevant discussion). This assumption may have precluded generative phonologists from analyzing sound symbolic patterns (although there are likely to be other reasons, like the influence of Saussure 1916).

Zuraw (2000) has shown, however, that those patterns that are not completely predictable can be still systematic ("patterned exceptions"), and subsequent work (Ernestus and Baayen, 2003; Hall, 2009; Hayes and Londe, 2006; Hayes and Wilson, 2008; Hayes et al., 2009; Kumagai and Kawahara, 2018; Pierrehumbert, 2001) has demonstrated that phonological knowledge can be deeply stochastic. There are phonological patterns that are only probabilistically predictable but yet productive.

2008; Wilson 2006; Zuraw and Hayes 2017).[4] One particular attractiveness of this model in the current context is the fact that MaxEnt is able to predict probability distributions of output candidates, which can be compared with the observed probabilities. To capture the sound symbolic effect of sonorants in Figure 1, we posit the following constraint in (1):

(1)   *SON→MALE: For each sonorant contained in name x, assign a violation mark if x is mapped to male names.[5]

This constraint reflects the tendency that is observed in the left panel of Figure 1: sonorants are preferentially associated with female names. This constraint alone, however, cannot account for the fact that even if all the consonants in the name are sonorants, they can be male names 26% of the time (the rightmost bar of Figure 1, left). In order to account for this fact, we posit a *FEMALENAME constraint, a constraint that is akin to *STRUC constraints (see Daland 2015 for the role of *STRUC in MaxEnt grammars). This constraint simply restates the fact that there can be male names.

MaxEnt grammar is similar to Optimality Theory (Prince and Smolensky, 1993/2004) in that a set of candidates is evaluated against a set of constraints. Unlike Optimality Theory, however, constraints are *weighted* rather than *ranked*, as in Harmonic Grammar (Legendre et al., 1990*a*,*b*; Pater, 2009, 2016). Based on the constraint violation profiles, for each candidate x, its Harmony Score (H-Score(x)) is calculated as follows:

$$\text{H-score}(x) = \sum_{i}^{N} w_i C_i(x) \tag{1}$$

---

[4]In addition to phonological input-output mapping, MaxEnt has been used to account for text-setting patterns (Hayes et al., 2012), and even career decisions (Hayes, 2016). Since MaxEnt is a function that maps any sort of input to probability distributions of output candidates, there is no a priori reason to limit MaxEnt to phonological input-output mapping (see Zuraw and Hayes 2017 for actual examples and references).

[5]This constraint is formalized in such a way that it dictates a mapping between sounds and output names. We can alternatively focus on output structures alone (like standard markedness constraints in Optimality Theory: Prince and Smolensky 1993/2004) and posit a constraint that prohibits sonorants in male names. We find that the former formalization captures the sound symbolic relationships more directly, as we are essentially talking about mapping from one modality (sound) to another (meaning) (see also Zuraw 2013). With that said, the latter output-oriented formalization works just as well.

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.

The H-scores are negatively exponentiated (eHarmony, 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. The eHarmony values are relativized against the sum of all the eHarmony values, which we call Z:

$$\text{Z} = \sum_{j}^{N} (e^{-H})_j \tag{2}$$

The probability of each candidate x_i, $p(x_i)$, is $\frac{eHarmony(x_i)}{Z}$.

To implement the analysis of Takarazuka actress names, we used the MaxEnt Grammar tool made available by Bruce Hayes,[6] which calculates the best weights for each constraint given the observed frequencies of each candidate, as well as the predicted probabilities based on these weights. The MaxEnt analysis tableau is shown in (2). The model takes each name as its input and calculates the probability of the name being used for a male name and the probability of the name being used for a female name. The analysis focuses on the number of sonorants that each name contains, as the number of sonorants was the significant factor in the analysis presented in section 2.2.

(2)   The MaxEnt analysis of Takarazuka names

---

[6]`http://linguistics.ucla.edu/people/hayes/MaxentGrammarTool/`

Name	Gender	*SON→MALE $w = 0.74$	*FEMALENAME $w = 1.11$	H-score	e^{-H}	Observed	Expected
0 son	Male			0	1	74%	75%
	Female		1	1.11	0.33	26%	25%
1 son	Male	1		0.74	0.47	62%	59%
	Female		1	1.11	0.33	28%	41%
2 son	Male	2		1.48	0.22	39%	41%
	Female		1	1.11	0.33	61%	59%
3 son	Male	3		2.22	0.11	25%	25%
	Female		1	1.11	0.33	75%	75%

As shown, the MaxEnt tool yielded the weights of the two constraints (0.74 and 1.11), which produce expected percentages that are very close to the observed percentages (the right two columns).

3 Nicknames of Japanese AKB idols

The sound symbolic association between sonorants and femaleness which we observed in the previous section manifests itself in another pattern in Japanese. Japanese idols often use nicknames that involve reduplication; those instances usually show reduplication of syllables with sonorant onsets (e.g. *mayu-yu*, *kayo-yo-n*, *miru-ru-n*), and reduplication of syllables with obstruent onsets is rare. This asymmetry can be understood as an instance of the sound symbolic association between sonorants and femaleness.

3.1 Method

To qualitatively assess this hypothesis, the nicknames of all of the AKB idol group members (AKB48, SKE48, NMB48, HKT48, NGT48, STU48, SDN48) were coded based on the wikipedia website specifically devoted to these idol groups (`https://48pedia.org/`, accessed September 2017). The numbers of obstruents and sonorants that were targeted by reduplication were coded. As a baseline, all obstruents and sonorants that appear in their names were also coded. The

current study focused on onset consonants and did not count coda consonants (Japanese coda consonants are limited to so-called "coda nasals"). There were no geminates in the names analyzed in the current study.

3.2 Result and analysis

Table 1: The numbers of sonorants and obstruents targeted by reduplication in AKB idol nicknames.

	Sonorants	Obstruents
Reduplicated	45 (5%)	7 (1%)
Total	900	643

Table 1 shows the results. In total, there were 900 sonorants in the corpus of AKB idol names, and 45 of them were targeted by reduplication (5%). On the other hand, only 7 out of 643 obstruents (1%) were targeted by reduplication. This difference between sonorants and obstruents is statistically significant ($\chi^2(1) = 16.44, p < .001$). Viewed from a different perspective, only 58% of the consonants in the whole AKB names are sonorants, but 87% of the reduplicated consonants are sonorants. We can conclude there that there is a clear bias to target sonorants in reduplication.

Accounting for these reduplication patterns requires only two constraints, one of which is very similar to (1):

(3)   a.   *OBS→FEMALE: For each obstruent contained in name x, assign a violation mark if x is a female name.

        b.   *REDUP: Assign a violation mark for each syllable that is reduplicated.

The first constraint militates against reduplicating syllables with obstruent onsets, which accounts for the very low probability of obstruents being reduplicated in the current dataset. The second constraint is necessary, because not all sonorants are reduplicated (this constraint is formalized as

INTEGRITY "no breaking" by McCarthy and Prince 1995). The MaxEnt analysis of AKB nicknames using these constraints are provided in (4). In this tableau, constraint violations of the first syllables are not shown, as they are shared by all the candidates considered (i.e. cancellation lemma: Prince and Smolensky 1993/2004).

(4)   The MaxEnt analysis of reduplicative patterns in AKB idol names

Name	Nickname	*REDUP W=2.94	*OBS→FEMALE W=1.65	H-score	e^{-H}	Observed	Expected
moka	moka		1	1.65	0.19	99%	99%
	mokaka	1	2	6.24	0.002	1%	1%
mayu	mayu			0	1	95%	95%
	mayuyu	1		2.94	0.05	5%	5%

As observed, the sound-symbolic constraint *OBS→FEMALE, together with a constraint that militates against reduplication, accounts for the difference between sonorants and obstruents in a very straightforward manner. One notable aspect of this analysis is that even when a candidate violates no constraints (i.e. [mayu]), it is not the case that it is assigned the probability of 1.00. This is because MaxEnt calculates the probability distribution over all candidates considered; less than optimum candidates are assigned some probabilities, and hence even the perfect candidate does not get "all the share."

4   Voiced obstruents and evolution levels in Pokémon names

As another case study, this section examines the role of voiced obstruents in determining the evolution levels of Pokémon characters. In the Pokémon game series, Pokémon characters undergo evolution, at most twice, and when they do so, they are called by a different name. Kawahara et al. (2018), based on all the Pokémon character names available as of October 2016, show that there is a positive correlation between the numbers of voiced obstruents contained in the names and the

evolution levels of the Pokémon characters.[7] Based on the results of Kawahara et al. (2018), the crucial aspect of their data is reproduced in Table 2. The probabilities of the names being used for non-evolved characters (Evol level 0) decrease as the names contain more voiced obstruents (the leftmost column). The probabilities of names being used for the most-evolved characters (Evol level 2) increase as the names contain more voiced obstruents (the rightmost column).

Table 2: The effects of voiced obstruents on the evolution levels in Pokémon names.

	Evol level 0	Evol level 1	Evol level 2
0 voiced obs	63%	30%	7%
1 voiced obs	50%	39%	11%
2 voiced obs	33%	26%	21%

To account for these distributional skews, we posit two constraints in (5):

(5)   a.   *VOIOBS→NONEVOL: For each voiced obstruent contained in name x, assign a violation mark if x is mapped to Evol level 1 names; assign two violation marks if x is mapped to Evol level 0 names.

      b.   *EVOLNAME: Assign two violation marks to Evol level 2 names; assign one violation mark to Evol level 1 names.

The first constraint is formalization of the sound symbolic correspondence between voiced obstruents and evolved characters—ideally, names with voiced obstruents are mapped onto names for the most evolved characters. In other words, it prohibits names with voiced obstruents from being used for non-evolved characters. *EVOLNAME is akin to *STRUC constraint (Daland, 2015), which represents the fact that evolved characters are less frequent than non-evolved characters. The MaxEnt analysis of Pokémon names using these two constraints appears in (6). As was the case for the two cases analyzed above, the model takes names as input and yields probability distributions

---

[7]In Japanese, voiced obstruents are associated with various types of images such as largeness, heaviness and strength parameters (Hamano, 1986; Kawahara, 2017; Kubozono, 1999; Suzuki, 1962). The sound symbolic patterns in Pokémon names are most likely rooted in this sound symbolic association that is present in the phonology and lexicon of Japanese. The "Pokémon corpus" allows us to quantitatively assess this sound symbolic association.

of each candidate; for this analysis, each candidate corresponds to each evolution level.

(6)   The MaxEnt analysis of Pokémon names

Name	Evol level	*Voi Obs→ NonEvol $w=0.40$	*EvolName $w=0.99$	H-score	e^{-H}	Observed	Expected
0 voi obs	0			0	1	63%	66%
	1		1	0.99	0.37	30%	25%
	2		2	1.98	0.14	7%	9%
1 voi obs	0	2		0.80	0.45	50%	54%
	1	1	1	1.39	0.25	39%	30%
	2		2	1.98	0.14	11%	17%
2 voi obs	0	2*2 = 4		1.60	0.20	33%	40%
	1	2*1 = 2	1	1.79	0.17	26%	33%
	2		2	1.98	0.14	21%	27%

We observe that there is a close correlation between the observed probabilities of each name type and their predicted values.

5   Associations between sounds and shapes

The last three case studies were based on corpus data. However, sound symbolic patterns are not only observed in existing names, but also in experimental settings. This section models the results of a naming experiment reported in Kawahara and Shinohara (2012). This experiment examined the association between sonorants and round figures on the one hand, and the association between obstruents (in this experiment, oral stops) and angular shapes on the other. These associations were first noted by the influential work of Köhler (1947), who showed that a nonce word like *takete* is more likely to be associated with an angular object, whereas a nonce word like *maluma* is more likely to be associated with a round object.[8] In the experiment, 20 disyllabic stimuli containing stop onsets (e.g. [bakə]) and 20 disyllabic stimuli containing sonorant onsets (e.g. [wejə]) were

---

[8]This effect is also known as the *bouba-kiki* effect (Ramachandran and Hubbard, 2001), and is extensively studied in psychology and phonetics (D'Onofrio, 2014; Fort et al., 2015; Maurer et al., 2006).

presented together with various pairs of an angular shape and a round shape. Within each trial, the participants were presented with an auditory stimulus together with a pair of an angular and a round object, and were asked to choose which object would be a better match for the auditory prompt. The participants were 17 native speaker of English. The results show that nonce names with sonorant onsets were associated with round shapes about 70% of the time, whereas nonce names with obstruents were associated with angular shapes 68% of the time (their Figure 6).

To account for these results, we posit two constraints, each of which directly corresponds to the observed sound-shape association:

(7)   a.   *STOP→ROUND: For each oral stop contained in name x, assign a violation mark if x is used for a name of a round shape.

        b.   *SONORANT→ANGULAR: For each sonorant contained in name x, assign a violation mark if x is used for a name of an angular shape.

The MaxEnt analysis appears in (8). In fact, this analysis achieves a perfect match between observed and predicted values. This is because the structure of violation profiles are simple: each constraint takes care of one mapping (either from oral stops to angular shapes, or from sonorants to round shapes), and the number of candidates within each mapping is limited to two.

(8)   The MaxEnt analysis of sound-shape mapping

Name	Shape	*STOP→ ROUND $w=0.38$	*SON→ ANGULAR $w=0.42$	H- score	e^{-H}	Observed	Expected
2 stops	Angular			0	1	68%	68%
	Round	2		0.76	0.47	32%	32%
2 son	Angular		2	0.84	0.43	70%	70%
	Round			0	1	30%	30%

6 Conclusion

This paper has shown that the MaxEnt grammars can model sound symbolic patterns, especially with regard to its gradient aspects. The way we analyzed the sound symbolic patterns is not different from how phonologists analyze phonological patterns using MaxEnt. We believe that this is an important success, because sound symbolism has not been extensively studied in the theoretical literature (modulo Alderete and Kochetov 2017), and one of the reasons may be that sound symbolic patterns are almost always gradient. A formal tool like MaxEnt can naturally account for the gradient nature of sound symbolic association patterns. To the extent that MaxEnt offers a natural way to account for probabilistic generalizations in phonology (Hayes and Wilson, 2008; Zuraw and Hayes, 2017), formal phonologists can analyze sound symbolic patterns in a way that may provide a new, interesting general domain for future phonological research.

References

Alderete, J., and Kochetov, A. (2017), "Integrating sound symbolism with core grammar: The case of expressive palatalization," *Language*, 93, 731–766.

Berlin, B. (2006), "The first congress of ethonozoological nomenclature," *Journal of Royal Anthropological Institution*, 12, 23–44.

Chomsky, N., and Halle, M. (1968), *The Sound Pattern of English*, New York: Harper and Row.

Coulter, K., and Coulter, R. A. (2010), "Small sounds, big deals: Phonetic symbolism effects in pricing," *Journal of Consumer Research*, 37(2), 315–328.

Daland, R. (2015), "Long words in maximum entropy phonotactic gammars," *Phonology*, 323, 353–383.

Diffloth, G. (1994), "i: *big*, a: *small*," in *Sound Symbolism*, eds. L. Hinton, J. Nichols, and J. J. Ohala, Cambridge: Cambridge University Press, pp. 107–114.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., and Monaghan, P. (2015), "Arbitrariness, iconicity and systematicity in language," *Trends in Cognitive Sciences*, 19(10), 603–615.

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

Ernestus, M., and Baayen, H. (2003), "Predicting the unpredictable: Interpreting neutralized segments in Dutch," *Language*, 79(1), 5–38.

Fort, M., Martin, A., and Peperkamp, S. (2015), "Consonants are more important than vowels in the bouba-kiki effect," *Language and Speech*, 58, 247–266.

Goldwater, S., and Johnson, M. (2003), "Learning OT constraint rankings using a maximum entropy model," *Proceedings of the Workshop on Variation within Optimality Theory*, pp. 111–120.

Hall, K. C. (2009), A Probabilistic Model of Phonological Relationships from Contrast to Allophony, Doctoral dissertation, Ohio State University.

Hamano, S. (1986), The Sound-Symbolic System of Japanese, Doctoral dissertation, University of Florida.

Hayes, B. (2016), "Some remarks on linguistics, textsetting, and life,", Commencement ceremony lecture, Boston University, May 14.

Hayes, B. (2017), "Varieties of Noisy Harmonic Grammar," *Proocedings of AMP*, .

Hayes, B., and Londe, Z. (2006), "Stochastic phonological knowledge: The case of Hungarian vowel harmony," *Phonology*, 23, 59–104.

Hayes, B., and Wilson, C. (2008), "A Maximum Entropy model of phonotactics and phonotactic learning," *Linguistic inquiry*, 39, 379–440.

Hayes, B., Wilson, C., and Shisko, A. (2012), "Maxent grammars for the metrics of Shakespeare and Milton," *Language*, 88(4).

Hayes, B., Zuraw, K., Siptár, P., and Londe, Z. (2009), "Natural and unnatural constraints in Hungarian vowel harmony," *Language*, 85(4), 822–863.

Hinton, L., Nichols, J., and Ohala, J. (2006), *Sound Symbolism, 2nd Edition*, Cambridge: Cambridge University Press.

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

Jakobson, R. (1978), *Six Lectures on Sound and Meaning*, Cambridge: MIT Press.

Jespersen, O. (1922), "Symbolic value of the vowel *i*," in *Linguistica: Selected Papers in English, French and German*, Vol. 1, Copenhagen: Levin and Munksgaard, pp. 283–30.

Kawahara, S. (2017), *Introducing phonetics through sound symbolism*, Tokyo: Hitsuji Syobo.

Kawahara, S., Noto, A., and Kumagai, G. (2018), "Sound symbolic patterns in Pokémon names," *Phonetica*, 75(3), 219–244.

Kawahara, S., and Shinohara, K. (2012), "A tripartite trans-modal relationship between sounds, shapes and emotions: A case of abrupt modulation," *Procedings of CogSci*, 2012, 569–574.

Kochetov, A., and Alderete, J. (2011), "Scales and patterns of expressive palatalization: Experimental evidence from Japanese," *Canadian Journal of LInguistics*, 56(3), 345–376.

Köhler, W. (1947), *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*, New York: Liveright.

Kubozono, H. (1999), *Nihongo-no Onsei: Gendai Gengogaku Nyuumon 2 [Japanese Phonetics: An Introduction to Modern Linguisitcs 2]*, Tokyo: Iwanami Shoten.

Kumagai, G., and Kawahara, S. (2018), "Stochastic phonological knowledge and word formation in Japanese," *Journal of the Linguistic Society of Japan*, 153, 57–83.

Legendre, G., Miyata, Y., and Smolensky, P. (1990*a*), "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.

Legendre, G., Miyata, Y., and Smolensky, P. (1990*b*), "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.

Locke, J. (1689), *An Essay concerning Human Understanding*, London: MDCC.

Lockwood, G., and Dingemanse, M. (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.

Maurer, D., Pathman, T., and Mondloch, C. J. (2006), "The shape of boubas: Sound-shape correspondences in toddlers and adults," *Developmental Science*, 9, 316–322.

McCarthy, J. J. (2010), "An introduction to Harmonic Serialism," *Language and Linguistic Compass*, 4(10), 1001–1018.

McCarthy, J. J., and Prince, A. (1995), "Faithfulness and reduplicative identity," in *University of Massachusetts Occasional Papers in Linguistics 18*, eds. J. Beckman, L. Walsh Dickey, and S. Urbanczyk, Amherst: GLSA, pp. 249–384.

Newman, S. (1933), "Further experiments on phonetic symbolism," *American Journal of Psychology*, 45, 53–75.

Ohala, J. J. (1994), "The frequency code underlies the sound symbolic use of voice pitch," in *Sound Symbolism*, eds. L. Hinton, J. Nichols, and J. J. Ohala, Cambridge: Cambridge University Press, pp. 325–347.

Pater, J. (2009), "Weighted constraints in generative linguistics," *Cognitive Science*, 33, 999–1035.

Pater, J. (2016), "Universal grammar with weighted constraints," in *Harmonic Grammar and Harmonic Serialism*, eds. J. J. McCarthy, and J. Pater, London: Equinox, pp. 1–46.

Pierrehumbert, J. B. (2001), "Stochastic phonology," *GLOT*, 5, 1–13.

Prince, A., and Smolensky, P. (1993/2004), *Optimality Theory: Constraint Interaction in Generative Grammar*, Malden and Oxford: Blackwell.

Ramachandran, V., and Hubbard, E. M. (2001), "Synesthesia–A window into perception, thought, and language," *Journal of Consciousness Studies*, 8(12), 3–34.

Sapir, E. (1929), "A study in phonetic symbolism," *Journal of Experimental Psychology*, 12, 225–239.

Saussure, F. (1916), *Cours de linguistique générale*, Paris: Payot.

Shattuck-Hufnagel, S. (1986), "The representation of phonological information during speech production planning: Evidence from vowel errors in spontaneous speech," *Phonology Yearbook*, 3, 117–149.

Shaw, J., and Kawahara, S. (2018), "Predictability and phonology: Past, present & future," *Linguistics Vanguard*, 4(S2).

Shih, S. S., Ackerman, J., Hermalin, N., Inkelas, S., and Kavitskaya, D. (2018), "Pokémonikers: A study of sound symbolism and Pokémon names," *Proceedings of LSA*, 2018.

Shinohara, K., and Kawahara, S. (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, K., and Kawahara, S. (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.

Sidhu, D., and Pexman, P. M. (2017), "Five mechanisms of sound symbolic association," *Psychonomic Bulletin & Review*, pp. 1–25.

Spence, C. (2011), "Crossmodal correspondences: A tutorial review," *Attention, Perception & Psychophysics*, 73(4), 971–995.

Styles, S. J., and Gawne, L. (2017), "When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter," *i-Perception*, pp. 1–17.

Suzuki, T. (1962), "Oninkookan to igibunka no kankei ni tsuite–iwayuru seidakuon tairitsu-o chuushin toshite," *Gengo Kenkyu [Journal of the Linguistic Society of Japan]*, 42, 23–30.

Uemura, Y. (1965), "Onsei-no hyoushousei-ni tsuite [On the symbolic aspects of sounds]," in *Gengo Seikatsu*, Tokyo: Honami Shuppan, pp. 66–70.

Ultan, R. (1978), "Size-sound symbolism," in *Universals of Human Language II: Phonology*, ed. J. Greenberg, Stanford: Stanford University Press, pp. 525–568.

Wilson, C. (2006), "Learning phonology with substantive bias: An experimental and computational study of velar palatalization," *Cognitive Science*, 30(5), 945–982.

Wilson, C. (2014), "Maximum entropy models,", Tutorial presentation, MIT.

Wright, S., and Hay, J. (2002), "Fred and Trema: A phonological conspiracy," in *Gendered Practices in Language*, eds. S. Benor, M. Rose, D. Sharma, J. Sweetland, and Q. Zhang, Stanford: CSLI Publications, pp. 175–191.

Wright, S., Hay, J., and Tessa, B. (2005), "Ladies first? Phonology, frequency, and the naming conspiracy," *Linguistics*, 43(3), 531–561.

Zuraw, K. (2000), Patterned Exceptions in Phonology, Doctoral dissertation, University of California, Los Angeles.

Zuraw, K. (2013), "*MAP constraints,", ms. UCLA.

Zuraw, K., and Hayes, B. (2017), "Intersecting constraint families: an argument for Harmonic Grammar," *Language*, 93, 497–548.