A received wisdom in modern linguistic theories is that the relationships between sounds and meanings are generally arbitrary. However, there is a growing body of evidence suggesting that in some cases sounds and meanings have systematic relationships—patterns known as “sound symbolism.” Yet most of these studies are conducted by psychologists, cognitive scientists or cognitive linguists, and currently, only a few theoretical phonologists pay serious attention to sound symbolism. This paper reviews major studies on sound symbolism and argues that sound symbolism can be an interesting topic of exploration for theoretical phonologists. This paper also demonstrates that insights gained by phonological research can shed light on some important issues in the studies of sound symbolism, and vice versa. Overall, I hope that this paper is informative for both theoretical phonologists and researchers who work on sound symbolism.
One well-known example is the observation that speakers of many languages feel nonce words containing [a] (e.g. [mal]) to be larger than those containing [i] (e.g. [mil]) (Berlin 2006; Newman 1933; Sapir 1929; Shinohara & Kawahara 2016; Ulan 1978). Another well-known case is the *takete-maluma* effect (Köhler 1947), in which names with voiceless obstruents tend to be associated with angular shapes, while names with sonorants tend to be associated with round shapes (Sidhu et al. 2019). These sound symbolic effects are observed in experimental settings as well as in the form of statistical skews in the lexicon (see the overview papers cited below).

The rise of interests in sound symbolism is partly evidenced by the number of recent overview articles on sound symbolism, each written from a slightly different perspective (Akita 2015; Cuskley & Kirby 2013; Dingemanse et al. 2015; Lockwood & Dingemanse 2015; Nuckolls 1999; Perniss et al. 2010; Schmidtke et al. 2014; Sidhu & Pexman 2018; Spence 2011; Svantesson 2017).

Given that so many overview papers already exist, one may wonder if there is a need for another overview paper. I contend that the answer is positive, because none of these papers are directed at theoretical phonologists. Nevertheless, I believe that studying sound symbolism can offer important insights into the architecture of phonological knowledge, and also that phonological studies have much to offer for the studies of sound symbolism.

It is helpful to start this discussion by considering why sound symbolism has generally been considered as residing outside the realm of theoretical phonology. To the extent that phonological knowledge is about what speakers know about the sound structure of their native language, if systematic connections between sounds and meanings exist, then, exclusion of sound symbolism from a topic of phonological inquiry does not seem to be clearly motivated.\(^1\) But then why is sound symbolism not actively studied in theoretical phonology?\(^2\) Of course the influences by Saussure and Hockett must have been non-negligible. Another possible reason may be that generative linguistic theories do not, or did not, accept probabilistic tendencies as belonging to competence—it is, or was, believed that grammars should only make a dichotomous, grammatical vs. ungrammatical distinction (Schütze 1996). However, we now have an extended body of evidence showing that phonological knowledge can be stochastic (Boersma & Hayes 2001; Hayes & Londe 2006;

---

\(^1\)I acknowledge that there are general disagreements regarding what the domain of phonological inquiry should include. de Lacy (2009), for example, excludes loanword adaptation patterns and diachronic changes from topics of phonological inquiry, while both of these receive extensive attention from other phonologists. Some researchers are more willing to accept “external evidence” (Churma 1979)—e.g. language game patterns or rhyming patterns—for phonological argumentation than others (Ohala 1986).

\(^2\)I do not have quantitative evidence for this claim, but Alderete & Kochetov (2017: 731) state that “it is fair to say that sound symbolism has never found a natural place in generative grammar.” I note that there have been phonological analyses of alternations that are demonstrably driven by sound symbolic principles (see §2.1). There are also a few monumental studies in the 90s that examined sound symbolic connections from a linguistic point of view, which include Hinton et al. (1994) and Hamano (1998). Ideophones have long been considered as exceptional to the Saussurean theorem of arbitrariness; their linguistic behaviors have attracted attention from theoretical linguists in the 70s and 80s (Dingemanse 2018: §5.2), and recently, there have been renewed interests in the linguistic analyses of ideophones (Akita 2019; Akita & Dingemanse 2019).
Pierrehumbert 2001; Zuraw & Hayes 2017), and several grammatical models have made it possible to model stochastic phonological knowledge (Coetze & Pater 2011). This rise of new stochastic grammatical models removed one obstacle that prevented us from exploring stochastic sound symbolic patterns (§2.3).

2 Phonology and sound symbolism

Some theoretical phonologists have started analyzing sound symbolism—in particular, alternation patterns that are caused by sound symbolic principles—as a part of their phonological inquiry. This section reviews some of these studies and presents other arguments that studies of sound symbolism are not as irrelevant as broadly assumed in theoretical phonology.

2.1 Some alternations are motivated by sound symbolic principles

The first reason to study sound symbolism from the perspective of theoretical phonology is the emerging observation that some alternations appear to be motivated by sound symbolic considerations. In particular, Alderete & Kochetov (2017) have shown that for example, palatalization in Japanese baby talk (what they call “expressive palatalization”) shows a number of properties that are different from purely phonological palatalization processes. First, phonological palatalization is assimilatory in nature, usually caused by (high) front vowels or palatal glides; expressive palatalization in Japanese, however, does not require such a trigger, changing all /s/ in a target word into either [ʃ] or [tʃ] (e.g. /osakana-san/ → [ofakana-tʃan] or [otʃakana-tʃan] ‘fish(-y)’). Second, affrication of fricatives can occur in expressive palatalization, as in the previous example, but never occurs in purely phonological palatalization (Bhat 1978). Third, expressive palatalization shows place and manner asymmetries that are not shared by purely phonological palatalization; for instance, the majority of expressive palatalization patterns targets coronal consonants, and moreover, when non-coronals are targeted, so are coronals. Fourth, expressive palatalization characteristically targets all relevant consonants within a word, which is rare at best for purely phonological processes. Finally, expressive palatalization can create an otherwise phonotactically illegal sequence; e.g. /sense:/ ‘teacher’ can become [tʃentʃe:], but [tʃe] is not a legal CV combination (Ito & Mester 1995). Expressive palatalization thus does not seem to be motivated by purely phonological considerations; instead, they appear to occur to express sound symbolic meanings.

These alternations driven by sound symbolic principles can interact with other phonological considerations within a single grammatical system. To illustrate, Kumagai (2019) shows that in Japanese nickname formation, /h/ can be turned into [p] (e.g. /haruka/ → [paruru]). It is hard to consider this alternation to be caused by a phonological constraint, since [p] is at best a marked
segment in the native phonology of Japanese (Ito & Mester 1995). Kumagai argues that this alteration is instead caused by a sound symbolic principle to express cuteness; indeed his experiment shows that [p] is judged to be “cuter” than any other consonants. Moreover, he has shown that this alteration interacts with an independently motivated phonotactic constraint that prohibits the [p...[+voice]] configuration (Kawahara 2018)—nicknames created by the [h]-[p] alternation are judged to be less natural when they violate *[p...[+voice]]. See Jang (2019) for a similar case found in Korean baby talk, Aegyo. These observations have led these authors to propose that sound symbolic principles should be integrated with “core phonological grammar.”

2.2 The role of distinctive features in sound symbolism

Another aspect of sound symbolism that makes it interesting for theoretical phonologists is the observation that some sound symbolic patterns operate at the level of distinctive features rather than individual segments. For example, secondary palatalization in Japanese mimetic words denotes, among other expressive meanings, “uncontrolledness” (Hamano 1998). This pattern should be characterized using a distinctive feature (e.g. [palatal] or [-back]), because consonants at all places of articulation can be affected by it (although there is a preference toward targeting coronal consonants). Indeed, this palatalization pattern, despite its sound symbolic nature, has been analyzed by a number of theoretical phonologists (see Alderete & Kochetov 2017: 732).

To take an example from a sound-meaning association, Kumagai & Kawahara (2020) point out that almost all diaper names in Japanese contain [p] and/or [m] (e.g. [mamiipoko]); when asked to produce new diaper names, Japanese speakers tend to use all types of labial consonants ([p], [b], [f], [m], [w]) more often than when they are asked to come up with new names for adult cosmetics. This result suggests that the sound symbolism at issue operates at the level of a distinctive feature—[labial]—rather than at the level of individual segments. Considering that the existing diaper names contain only [p] and [m] but not other labial consonants, the participants of their experiment seem to have shown feature-based generalizations, just as “normal” phonological patterns do (Albright 2009; Finley & Bedecker 2009). This observation raises the possibility that phonological patterns and sound symbolic patterns can use the same set of vocabularies, i.e. distinctive features. See also D’Onofrio (2014), Jakobson (1960), McCarthy (1983) and Nobile (2015) for related discussion.

2.3 Modeling sound symbolism with formal grammatical frameworks

Mainstream generative models of grammar have long considered sound symbolism as residing outside the purview of linguistic knowledge proper. However, the existence of sound symbolism

---

3 This is not to say that all sound symbolic patterns can be defined based on distinctive features (Akita 2015). However, not all phonological alternations can be neatly captured in terms of distinctive features either (Mielke 2008). This complexity may instantiate a parallel between phonological alternations and sound symbolic patterns (see §3).
may require us to posit a mechanism that directly connects sounds and meanings, and some studies propose to use a formal grammatical model to capture such connections. For instance, as discussed above, Alderete & Kochetov (2017) demonstrate that expressive palatalization patterns are motivated by sound symbolic principles. They propose an analysis of expressive palatalization using Optimality Theory (Prince & Smolensky 2004), with a set of violable constraints (EXPRESS(X)) specifying which sounds should be realized to express which meanings. Kawahara et al. (2019) analyze sound symbolic connections themselves (rather than alternations triggered by sound symbolic considerations), and argue that as long as generative phonology is a function that maps one representation (e.g. underlying forms) to another (e.g. surface forms)—as it in fact has been—there is nothing that prevents us from using the same formalism to model the mapping from representation in one modality (i.e. sound) to representation in another modality (i.e. meaning). In other words, the grammatical device that phonologists have been using for decades can be applied to formalize sound symbolic connections at no additional costs.

As stated in §1, sound symbolic connections are almost always stochastic—certain sounds tend to be associated with certain meanings, but such associations are never deterministic (Dingemanse 2018). Likewise, recent studies have demonstrated that phonological knowledge can be stochastic; e.g., some structures tend to be preferred over others, and some alternations are more likely to occur in one environment than in other environments (Boersma & Hayes 2001; Hayes & Londe 2006; Pierrehumbert 2001; Zuraw & Hayes 2017). Focusing on this parallel between phonological patterns and sound symbolic connections, Kawahara et al. (2019) propose to use MaxEnt model with Optimality Theoretic constraints (Goldwater & Johnson 2003) to account for stochastic aspects of sound symbolic patterns. For example, in Japanese Takarazuka actress names, the more sonorants are contained in their name, the more likely that name is used for a female name. MaxEnt grammar with a constraint like *SONORANTINMALENAME accounts for the stochastic nature of this sound symbolic pattern—this analysis is fundamentally the same as a MaxEnt analysis of “normal” phonological patterns (e.g. Zuraw & Hayes 2017). In short, phonological patterns and sound symbolic patterns share a non-trivial nature (stochasticity), and an analytical tool like MaxEnt is able to capture that parallel.

3 Common issues and shared interests

There are common issues that are addressed both by theoretical phonologists and those who study sound symbolism. My intention in this section is not to solve any of these issues, but to show that we have shared interests. By making these shared interests clear, it is hoped that insights gained in one domain of inquiry can shed light on questions that are addressed in the other. I also hope that these parallels pique phonologists’ interests to study more about sound symbolism.
3.1 Phonetic naturalness

One continuing debate in phonological theory is to what extent phonological patterns are natural with respect to phonetic considerations. On the one hand, there is a group of proposals arguing that most if not all phonological patterns are phonetically motivated, and/or phonological representations contain detailed phonetic information (Hayes et al. 2004). On the other hand, some researchers argue that synchronic phonological systems should be completely void of phonetic substance (Reiss 2018). See Kingston (2019) for the most recent review on this debate. At the observational level, some phonological generalizations do seem to be motivated by phonetic considerations. For instance, voiced stops are generally considered marked compared to voiceless stops, because voiced stops present an aerodynamic challenge (Hayes & Steriade 2004). In order to maintain glottal vibration, there must be a sufficient transglottal airpressure drop; however, stop closure raises intraoral airpressure, and speakers need to expand their oral cavity to accommodate this aerodynamic challenge (Ohala 1983a). Many languages thus avoid voiced stops in favor of voiceless stops, which seems to have its roots in the aerodynamic challenge. Several questions still remain actively debated: (1) whether the influence of phonetics on phonology should be directly encoded in synchronic grammar, or it should be treated as the results of phonetically motivated diachronic changes; (2) whether those phonological patterns that seem “crazy” from the perspective of phonetic naturalness (Bach & Harms 1972) can be productive; (3) and if so, how we should model natural and unnatural aspects of phonological knowledge.

Just like some phonological generalizations, some sound symbolic connections seem to be grounded in the articulatory or acoustic properties of the sounds at issue. This intuition was already expressed by Sapir (1929: 233), a pioneering work on modern studies of sound symbolism: “the symbolic discriminations run encouragingly parallel to the objective ones based on phonetic considerations.” Jespersen (1922) and Sapir (1929) found that speakers generally judge [a] to be larger than [i], and attributed this observation to either the differences in the degrees of oral aperture, or the differences in their resonance frequencies (most likely their F2). To provide another example, voiced obstruents are often considered to be larger than voiceless obstruents by speakers of different languages (Newman 1933; Hamano 1998; Shinohara & Kawahara 2016), and it is not hard to imagine that this image of largeness is grounded in the expansion of the supralaryngeal cavity that is necessitated by the aerodynamic requirement that voiced stops present (Ohala 1983a). D’Onofrio (2014) shows that segments that involve lip gestures—such as [b] and [u]—tend to be associated with round figures, in the context of studies of what is known as the *bouba-kiki* effect (Ramachandran & Hubbard 2001). It is again likely that the lip rounding gesture of these sounds leads to the image of roundedness. These sound symbolic patterns do not seem to arise from statis-

---

4This observation actually goes back to Socrates, who in *Cratylus* (427) discusses relationships between sound symbolic meanings and phonetic properties of the sounds at issue.
tical skews in the lexicon; neither is there a plausible diachronic story in which these relationships arose—they instead appear to emerge in experimental settings using nonce words, suggesting that the influence of phonetic considerations in sound symbolism is synchronically active.

On the other hand, some sound symbolic patterns do not appear to have such clear phonetic motivations. For example, English gl- sequences occur in many words that are related to the notion of “light” (e.g. glitter, glow, gloaming). However, there are no phonetic reasons to expect that the sequence of gl- should be connected to the meaning of “light.” It has been shown moreover that this sound-meaning connection is psychologically real; i.e., this sound-meaning correspondence cannot be relegated as an accidental connection in the English lexicon (Bergen 2004).

We also observe cases that are akin to phonetically “crazy rules” (Bach & Harms 1972) in sound symbolism. For example, in Korean, [a] and [o] are symbolically smaller than [u] and [ʌ] (Garrigues 1995; Kim 1977). This pattern flouts an otherwise cross-linguistically common observation that high vowels are generally judged to be smaller than non-high vowels (Sapir 1929 et seq), and not only that, to the extent that this sound symbolism has its roots in the different degrees of oral aperture, it runs counter to what we expect from phonetic considerations. In other words, these sound symbolic connections in Korean are phonetically crazy (see also Diffloth 1994).

In the phonology literature, there is a debate regarding whether phonetically crazy rules can be productive or not (Hayes et al. 2009; Kawahara 2008; Sanders 2003). Sometimes crazy rules turn out to be non-productive (Sanders 2003), or there is a learning bias against them (Hayes et al. 2009; Hayes & White 2013; White 2014; Wilson 2006). In this connection it is interesting that Shinohara & Kawahara (2016) found that given nonce words, Korean speakers judge high vowels to be smaller than low vowels, contrary to what we expect from the lexical patterns. Similar results in which grammatical naturalness, which is demonstrably grounded in phonetic considerations, triumphs unnatural lexical skews are reported in some recent phonological studies (Jarosz 2017; Guilherme 2019).

3.2 Bases of representations—articulation or acoustics?

Another major debate in phonological theory is the phonetic bases of distinctive features. Jakobson et al. (1952) first formalized distinctive features in terms of acoustic characteristics. On the other hand, the feature set deployed by Chomsky & Halle (1968) was primarily based on articulation. Since then, there has been a heated debate as to whether distinctive features—or phonological representations in general—should be defined based on articulation or acoustics/perception (Kingston 2007). A similar debate exists in the phonetics literature regarding whether phonetic targets should be defined articulatorily or acoustically, and also regarding whether the object of speech perception is articulation or acoustics/audition (Kingston 2007).

6 Setting aside the view that phonological representations should completely lack phonetic substance (Reiss 2018).
A similar debate arises in the studies of sound symbolism. Take the case of [a] being perceived as larger than [i]. Jespersen (1922: 558-559) entertains two hypotheses regarding why [i] is considered to be small, one based on acoustics and one based on articulation: “[t]he reason why the sound [i] comes to be easily associated with small, and [u, o, a] with bigger things, may be to some extent the high pitch of the vowel...; the perception of the small lip aperture in one case and the more open mouth in the other may have also its share in the rise of the idea.” Ohala (1994) advocated a general theory of sound symbolic patterns, which is now known as the Frequency Code Hypothesis. In this theory, sounds with high frequency energy (either f0 or F2) evoke images of smallness, since these sounds are generated by a set of small vocal folds (in case of high f0) or a small resonating chamber (in case of high F2).

One argument for the acoustics-based explanation is that in some languages such as Twi, H-tone can represent something small (e.g. [kákrà] ‘small’ vs. [kàkrá] ‘large’), and there is nothing plausible in the articulation of H-tone that can be connected to the image of smallness (Ohala 1983b). Another argument for the acoustics-based explanation is that speakers of different languages can order their vowels in terms of how big they sound (Newman 1933), and the inverse of F2 is almost a perfect predictor of this judgment pattern (Shinohara & Kawahara 2016).

On the other hand, there are arguments in favor of the articulation-based explanation as well. First, deaf children can detect the sound symbolic values of their own speech and generally, these patterns are similar to the patterns observed for hearing people (Eberhardt 1940). Second, the acoustics-based explanation leaves it unclear why F1 does not affect the images of size, at least not as much as F2—[a] has higher F1 than [i], and therefore, if listeners deduce the sound symbolic values of vowels from their F1, [a] should be judged to be smaller than [i]. Third, the connection between labial segments and round shapes (D’Onofrio 2014) seems to be most straightforwardly explained in terms of articulation, not in terms of acoustics—nothing in the acoustics of labials appears to be “round.”

3.3 Universality and language-specificity

If sound symbolic patterns have their roots in their phonetic characteristics, one can imagine that sound symbolism is universal, shared across all languages, as we use the same articulatory and perceptual systems. Therefore, the universality of sound symbolic patterns is one topic that is actively discussed. A parallel question—how universal is the phonological system—is one of the central questions in modern linguistics. In both research domains, to the extent that there are universals, a deeper question is what level of abstraction is necessary to establish those universals (Akita 2015).

7Knoeferle et al. (2017) show that stimuli with higher F1 can lead to higher size rating.
The universality of sound symbolism has been addressed in two ways. One is a cross-linguistic comparison, often in the form of experimentation. Shinohara & Kawahara (2016), for example, used nonce word stimuli to explore the judgment of size associated with five different vowels, [i], [e], [a], [o], and [u], targeting speakers of Chinese, English, Japanese, and Korean. They found that speakers of all languages judged [a] to be larger than [i], hinting at the universality of this pattern (though see Diffloth 1994); on the other hand, Japanese speakers judged [o] to be larger than [a], whereas Chinese and Korean speakers showed the opposite pattern, and English speakers did not show a substantial difference between the two vowels. To take another example, the takete-maluma effect (Köhler 1947) has been shown to hold across many languages (Styles & Gawne 2017), but it fails in Songe (Rogers & Ross 1975) and Syuba (Styles & Gawne 2017). Based on a meta-analysis of the previous studies on the takete-maluma effect, Styles & Gawne tentatively propose that it fails to hold if the stimuli violate phonotactic restrictions of the target language. See Bremner et al. (2013), Dingemanse et al. (2016), Saji et al. (2019) and Shih et al. (2019) for related discussion.

Another approach is to examine the behavior of pre-verbal infants. As reviewed in §4.2, there is now a growing body of work showing that pre-verbal infants are sensitive to cross-linguistically prevalent sound symbolic patterns (Imai et al. 2008; Kantartzis et al. 2011; Maurer et al. 2006; Ozturk et al. 2013; Peña et al. 2011). On the other hand, some acquisition studies show that not all sound symbolic patterns may be universal. Fort et al. (2013), for example, failed to find an otherwise cross-linguistically robust bouba-kiki effect (Ramachandran & Hubbard 2001) in French infants. Iwasaki et al. (2007) found that native speakers of English without any L2 background on Japanese were able to guess the meanings of some sound symbolic, onomatopoetic words in Japanese, but not others, arguing that some sound symbolic patterns are universal while others are not.

Building on these results, Imai & Kita (2014) hypothesize that “young children are sensitive to all possible sound symbolic correspondences that could appear in any language of the world, but only a subset of these correspondences are compatible with the phonological inventory and the existing words in the language the children are learning. As they grow up, the sensitivity to the incompatible correspondences wanes, and adults maintain only the sensitivity to the compatible correspondences.” This hypothesis may remind phonologists of Stamp’s (1973) proposal that babies are born with a set of universal, phonetically-motivated processes, and as they grow up, they unlearn some of them and acquire language-specific rules.

3.4 Cumulativity

Another common issue is the question of cumulativity. In phonology, this issue is most actively discussed in the context of comparing Optimality Theory (Prince & Smolensky 2004) with other
related constraint-based theories (Pater 2009; Zuraw & Hayes 2017): when a structure violates two
constraints, do the effects of these two constraints add up? Optimality Theory posits that only the
higher-ranked constraint matters, whereas other theories such as Harmonic Grammar, suggest that
the expected outcome should be cumulative.

A similar question arises when studying sound symbolism. For instance, some studies report
that two instances of the same segment can evoke a stronger image than one instance (Hamano
2013; Kawahara et al. 2019; Kawahara & Kumagai to appear; Martin 1962), which is akin to what
Jäger & Rosenbach (2006) refer to as *counting cumulativity*. Moreover, D’Onofrio (2014) shows
that in the *bouba*-kiki effect, several phonetic/phonological features matter in determining the per-
ceived roundness/angularity of visual images (e.g. vowel backness, voicing, and place of artic-
ulation), and that these effects are cumulative, which would remind phonologists of *ganging-up
cumulativity* (Jäger & Rosenbach 2006). Kumagai & Kawahara (2019) show that Japanese speak-
ers prefer to use low vowels and voiced obstruents for the names of evolved Pokémon characters;
while each of these effects is observed independently, when combined in one stimulus, their ef-
facts are cumulative. Based on these observations, they present an analysis using MaxEnt Grammar
(Goldwater & Johnson 2003), analogous to the phonological analyses presented in Zuraw & Hayes
(2017). See Ahlner & Zlatev (2010) and Thompson & Estes (2011) for other potential cases of
cumulative effects in sound symbolism. Kawahara (2020) offers MaxEnt analyses of these cumu-
late patterns.

## 3.5 Positional asymmetry

Yet another common issue is positional asymmetry. It is well-known in the phonology literature
that some positions—e.g. onsets, stressed syllables, word-initial segments—are privileged in that
they can, for example, trigger phonological processes and license a wider range of segments (e.g.
Beckman 1998). Similarly, word-initial voiced obstruents cause stronger sound symbolic image
than word-internal voiced obstruents in Japanese (Kawahara et al. 2008) (see also Haynie et al.
2014; McGregor 1996).

Compared to the extensive body of research on positional effects in phonology, this issue is less
well examined in the studies of sound symbolism. Exploring positional effects in sound symbolism
may offer a novel perspective on this issue, since why and how some positions are privileged is
related to the issue of phonetic grounding of phonological patterns (§3.1)—positional privileges
are often considered as arising from phonetic/psycholinguistic prominence of these positions, but
whether this influence should be encoded synchronically or diachronically is much debated (e.g.
Barnes 2002 vs. Smith 2002). Also, for some positions, it is not clear whether that position is
phonologically privileged or not (e.g. word final position). Exploring the positional effects in
sound symbolism may help phonologists address this question as well; see e.g. McGregor (1996)
who explored the sound symbolic values of root-final consonants in Gooniyandi.

4 Additional potential benefits

There are additional benefits that phonologists can gain from studying sound symbolism. Due to space limitation, the discussion in this section needs to be brief; see the other overview papers cited in §1 as well as the reference cited in this section for further details.

4.1 Addressing the origin of human languages

First, some researchers propose that mimicking real world attributes with different types of vocalization contributed to the origin of human languages (Berlin 2006; Cabrera 2012; Haiman 2018; Perlman & Lupyan 2018; Perniss & Vigiliocco 2014; Perniss et al. 2010; Ramachandran & Hubbard 2001). A recent study by Perlman & Lupyan (2018), for example, shows that speakers are able to communicate as many as 30 different meanings by way of iconic vocalizations, well above the chance level; moreover, vocalizations which were judged to be more iconic were learned faster. They conclude that “[t]his newly emerging understanding of iconicity as a widespread property of spoken languages suggests iconicity may also have played an important role in their origin. An intriguing possibility is that many of the now arbitrary words in modern spoken languages may have originated from the innovation of iconic vocalizations.” If this hypothesis is on the right track, analyzing sound symbolic patterns might shed light on how human languages may have emerged and evolved.

4.2 The role of sound symbolism in language acquisition

Some researchers argue that sound symbolism plays a non-trivial role in language acquisition (Asano et al. 2015; Imai et al. 2008; Kantartzis 2011; Perry et al. 2018). In the context of first language acquisition, Maurer et al. (2006) demonstrate that 2.5 year old children are sensitive to sound symbolic associations. Even more strikingly, Peña et al. (2011) show that this sensitivity is detectable in 4 month old neonates. These observations have led to the general hypothesis that sound symbolism guides the first language acquisition process (Imai & Kita 2014; see §3.3).

A partial support of this theory comes from the observation that Japanese caretakers use more sound symbolic, onomatopoetic words to infants and children than to adults (Fernald & Morikawa 1993). Sound symbolism may thus provide a partial answer to the question of why human children acquire languages so quickly—a fundamental question that theoretical linguists attempt to answer. In addition, it has been observed that, in the context of L2 acquisition too, those items that follow sound symbolic principles are easier to learn and more frequently used by language learners
4.3 Tighter connections with cognitive science

Sound symbolism is now considered as a specific instance of general cross-modal, synesthetic correspondences, in which sensation in one modality has correspondences with sensation in another modality (Bankieris & Simner 2015; Cuskley & Kirby 2013; Sidhu & Pexman 2018; Spence 2011). There has been a growing body of interest in these cross-modal perception patterns in cognitive science, exploring not only the relationship between sounds and meanings, but also the relationships between sounds and cognitive patterns in other modalities (such as vision and taste). Therefore, engaging with linguistic analyses of sound symbolism has the potential to facilitate more extensive interdisciplinary communication between phonologists and cognitive scientists. Two questions addressed in this research that are of particular interest to theoretical linguists are: (1) are sound symbolism/synesthetic connections innate or acquired, and (2) are sound symbolic connections deducible to a domain-general synesthetic mechanism? (Akita 2015: §4.3)

4.4 Sound symbolism and branding

There is an increasing body of work that seeks to deploy sound symbolism in the context of marketing. Their general finding is that there are sounds that are “suitable” to convey particular images for brand products; such names that make sound symbolic sense are judged to be better, and possibly better remembered, by potential customers (Bolts et al. 2016; Coulter & Coulter 2010; Jurafsky 2014; Klink 2000; Peterson & Ross 1972; Yorkston & Menon 2004). This line of research has opened up a new domain of interdisciplinary research.

4.5 Popularizing linguistics and application to pedagogy

Finally, studying sound symbolism may help us popularize linguistics, and relatedly, can be useful for teaching. One example that instantiates these points is a study of sound symbolic patterns in Pokémon names (Kawahara et al. 2018), which has shown, for example, that Pokémon characters’ weight positively correlates with the number of voiced obstruents contained in their names. This result was featured in various popular magazines in Japan. Stephanie Shih, who followed up the original study with a much wider range of languages (Shih et al. 2019), was featured in a radio show to talk about this project. Being able to show to the general public that analytical tools that phonologists use—such as voiced obstruents—help in the analysis of Pokémon names can be appealing. This feature of sound symbolism also often attracts undergraduate students’ interests as well, since the targets of the analyses include “fun materials” that they are already interested in (Shih et al. 2019).
in (e.g. Pokémon). See Kawahara (2019) and MacKenzie (2018) for reports of the usefulness of using onomastics, including sound symbolism, in undergraduate education.

5 Conclusion

Studying sound symbolic patterns can be interesting for theoretical phonologists. To recap, there are some alternation patterns that seem to be motivated by sound symbolic principles, which can interact with other phonological considerations. There are many common interests shared by theoretical phonologists and those who work on sound symbolism, and we can mutually inform one another. There are various additional benefits that phonologists can gain from working on sound symbolism. Exploration of sound symbolism may contribute to the further development of the field by attracting interests from students and researchers in other fields.

References


Perlman, Marcus & Gary Lupyan. 2018. People can create iconic vocalizations to communicate various meanings to naïve listeners. *Scientific Reports* 26–34.

Perniss, Pamela, Robin L. Thompson & Gabriella Vigliocco. 2010. Iconicity as a general
Shih, Stephanie S., Jordan Ackerman, Noah Hermalin, Sharon Inkelas, Hayeun Jang, Jessica Johnson, Darya Kavitskaya, Shigeto Kawahara, Miran Oh, Rebecca L Starr & Alan Yu. 2019. Cross-linguistic and language-specific sound symbolism: Pokémonastics. Ms. University of Southern California, University of California, Merced, University of California, Berkeley, Keio University, National University of Singapore and University of Chicago.
Smith, Jennifer. 2002. *Phonological augmentation in prominent positions*: University of Mas-
Schartau, Amherst Doctoral dissertation.


