Automatic Detection of Phonestheme-like Features in Japanese: Insights for Cognitive Sound Symbolism Research

Thomas "Raz" Parker
Department of English, San Francisco State University

Abstract

It is traditionally believed that words do not have innate meaning. Sound symbolism is a controversial feature of many languages that may suggest a non-arbitrary link between sound and semantic content. Often dismissed as a fringe phenomenon, sound symbolism has recently been shown to play a role in language acquisition, and potentially language evolution. One sub-morphemic sound-symbolic structure known as *phonesthemes* has been identified in many languages. Phonesthemes are sub-morphemic patterns of sounds that are associated with a specific meaning more frequently than accounted for by chance. Examples include *gl-*, observed in many words related to light phenomena such as *gleam*, *glow*, or *glimmer*. Cognitive approaches to sound-symbolism and phonesthemes typically rely on a known or hypothesized example. Phonesthemes are not well attested in Japanese, so this option is unavailable. This paper is intended to provide a route for further research into the cognitive effects – in particular language acquisition – of Japanese sound symbolism. With this goal in mind, I present a model for automatically identifying phonestheme-like features in Japanese adapted from a model originally used with English. Using this model, I identify two candidate phonestheme-like clusters in Japanese. I discuss potential cognitive routes of conventionalization for one of these clusters, and outline methods in which the psychological reality of these clusters may be empirically tested based on existing sound symbolism literature in both English and Japanese.

Keywords: language acquisition; language evolution; sound symbolism; phonesthemes; Japanese; natural language processing; conceptual metaphor; cognitive linguistics

Introduction

Arbitrariness is a traditional core tenet of linguistics: the assumption that the semantic content and the phonetic expression of a word are unrelated (Saussure, 1916; Chomsky, 1957; Hockett, 1959). Sound symbolism is the idea that there is in fact a direct relationship between semantic content and phonetic expression. Research has persistently shown that sound symbolism does, at a minimum, have a measurable and manipulable effect on language users: sound symbolism has psychological reality. One sound symbolic feature on which recent work has been conducted is *phonesthemes*. Phonesthemes are sequences of sounds that appear in words with a specific semantic relationship at frequencies that exceed chance. Phonesthemes are not meaningful in isolation, but may add non-compositional semantic information to words in which they appear. Examples include *gl-*, which is observed in many words related to light phenomena, such as *gleam*, *glow*, or *glimmer*, or *bl-*, observed in words relating to apathy or disinterest such as *blasé*, *blithe* and *blank* (Hutchins, 1998; Bergen, 2004). Much work on sound symbolism tests a known or hypothesized example of sound symbolism. Given that phonesthemes often go unrecognized, some researchers have proposed seeking out patterns that may be explained by sound symbolism first, before testing them (Liu, Levow, & Smith, 2018). Research into word-initial sound symbolism similar to phonesthemes is less common in Japanese. Candidate phonestheme-like features are accordingly less well documented, posing an issue for establishing both the existence of these features and their psychological effects on Japanese speakers. This paper presents an adaptation of a model originally used in English to predict phonesthemes to Japanese. The goal is to identify phonestheme-like sequences in Japanese that may be used in future experiments with Japanese speakers s to establish or refute their psychological reality for Japanese speakers.

Conventionalized and Iconic Sound Symbolism

Some researchers differentiate between iconic and conventionalized sound symbolic features. *Iconic* features can be shown to be inspired by articulatory characteristics, acoustic characteristics, or both. One of the most common examples of an iconic sound symbolic relationship is /ɪ/–/æ/ magnitude symbolism, which is suggested to be motivated by the shape of the mouth, where /æ/ suggests a large sound source and therefore naturally has an association with largeness (Ohala, 1994). *Conventionalized* sound symbolism are instances where there is no identifiable link between the production of the sound and the semantic association. Most if not all phonesthemes are assumed to be conventionalized rather than iconic[1]. Researchers have demonstrated that both iconic and conventionalized sound symbolism have an identifiable psychological effect on speakers. In iconic instances, this can be seen as a speaker's cognitive process influencing production (e.g. an impression of largeness promoting a corresponding mouth shape resulting in a specific vowel) that, in turn, has an effect on a speaker.

Conventionalized instances can be seen as an utterance that contains a specific non-iconic sound having a cognitive effect on the speaker that eventually gains an association which

[1]Bolinger (1950), for instance, suggests this when he illustrates the lack of a stable relationship between a sound and its meaning (i.e. the cluster *gl-* in glow and gleam can be assumed to be the same phonestheme, but it is not the same as the cluster in glass).

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

is then reproduced in additional, nominally unrelated words (Hinton, Nichols, & Ohala, 1994). This is visualized schematically in Figure 1.

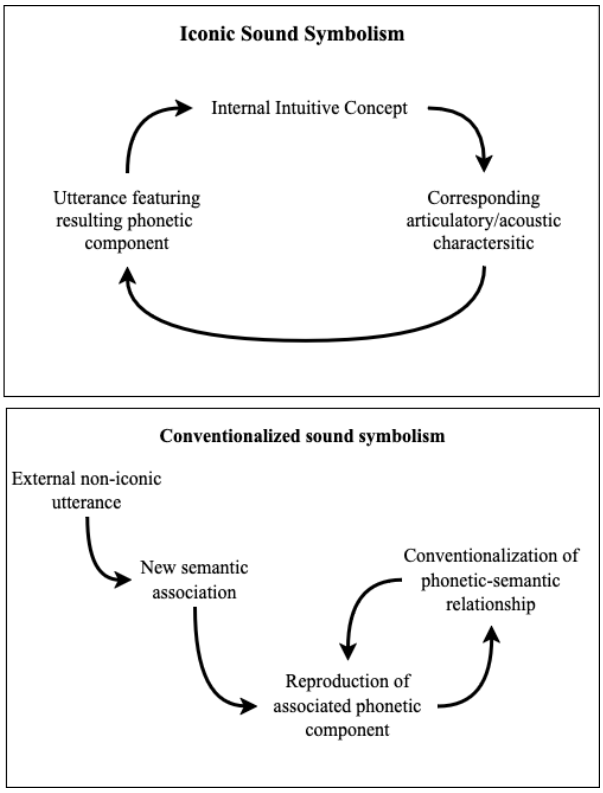

Figure 1: A schematic representation of iconic vs conventionalized sound symbolism, based on Hinton's (1994) explanation. Diagram is the author's own.

Given that this paper investigates phonestheme-like features, they will likely also be conventionalized. Their route of conventionalization may or may not be identifiable to a Japanese speaker.

Cross-linguistic Effectiveness of Japanese Sound Symbolism

Sound symbolism has been investigated many languages – from languages like English with more than 380 million speakers, to critically endangered languages such as Nez Perce. Japanese has a rich lexicon of reduplicative sound-symbolic words on which much and varied research has been undertaken (Hamano, 1998; Iwasaki & Yoshida, 2019)[2]. However, research on other forms of sound symbolism in Japanese such as phonesthemes is less readily available.

Researchers have demonstrated that at least some sound-symbolic features have a cognitive effect cross-linguistically. Kantartzis, Imai, and Kita (2011) tested L1 English-speaking children with a verb generalization test. Participating children were shown a video of an action being performed, and prompted with a training phrase "Look, he is doing X", where X was a novel word. In one condition, the novel word was a sound-symbolic word derived from a Japanese word with sound-symbolic content matching the action. In the other, the novel word was not intended to be sound-symbolic. They found that even with no knowledge of Japanese, English-speaking children performed better with the Japanese-derived sound symbolic words, indicating sound-symbolism not only plays a part in language acquisition, but also that some features of it may be accessible regardless of language of origin.

Iwasaki, Vinson, and Vigliocco (2007) found similarly with adult speakers. In their experiment, native speakers of Japanese and participants with no prior Japanese learning experience were asked to rate how well Japanese ideophones related to walking and laughing matched with depictions of these actions for a variety of dimensions. For many of these dimensions, they found that the associations between the word and the depictions correlated for the Japanese-speaking and non-Japanese-speaking participants.

Sound symbolism is evidently robustly attested and researched in Japanese ideophones. Individual consonant and vowel associations are identified in Japanese and research is ongoing. However, research into word-initial sounds that may function similarly to consonant-string phonesthemes like those in English and other languages is less common.

Yoshida et al. (2023) assert that phonesthemes like the word-initial strings of consonants identified in the previous literature on English are not possible in Japanese due to the phonotactic restrictions of its mora system[3]. It is certainly true that the orthography and phonotactics of Japanese requires vowels between consonants except in a handful of environments therefore apparently precluding consonant strings. However, Japanese words are not necessarily pronounced rigidly according to this CV structure[4]. Instead, Yoshida et al. (2023) suggest the existence of "phonesthemic clusters", which they tentatively identify in the first two consonants of kanji readings, using an automatic labeler based on WordNet semantic labeling.

Given that the effects of some Japanese sound symbolic features apply cross linguistically, the question remains as to whether phonesthemes (which are known to exist in other languages) exist in Japanese and have psychological reality for Japanese speakers. The following section discusses some of the approaches researchers have taken to establish various cognitive effects of phonesthemes in English.

[2]Asano Tsuruku's Onomatopoeia and Mimetic Words: A dictionary (1978) identifies around 1,500, while Chang's Thesaurus of Japanese mimesis and onomatopoeia (1990) identifies more than 1800.

[3]The phonotactics of Japanese permit a limited set of syllable structures, most frequently V or CV, with bare C only occurring in syllable-final nasal /n/ and in germinate clusters.

[4]The Japanese for deer, *shika*, composed of mora *shi* and *ka* is commonly pronounced with [ɕi̥ka̠], with a devoiced /i/ (Fujimoto, 2015).

The Cognitive Reality of Phonesthemes in English

In English, Sereno (1994) for instance observed in the Brown Corpus that nouns were statistically more likely to contain back vowels, while verbs were more likely to have front vowels. To test the significance of this finding with English speakers empirically, Sereno conducted an experiment in which participants were asked to identify words they were shown on a computer screen as a noun or a verb. The results indicated that nouns that contained back vowels and verbs that contained front vowels were categorized by participants faster than the reverse, demonstrating that the association was not merely statistically significant, but cognitively significant as well.

A similar approach has also been applied to phonesthemes. Bergen (Bergen, 2004) used priming experiments to demonstrate cognitive effects of phonesthemes. In the experiments, participants are flashed a pair of words sequentially and asked to identify whether the second word was a real English word or not. Five conditions were used in the experiment: (i) the words share a phonestheme (e.g. glitter and glow), (ii) the words share an onset (e.g. druid and drip), (iii) the words share semantic content (e.g. cord and rope), (iv) the words share an onset and semantic content, but not a statistically significant phonestheme (e.g. crony and crook), and (v) words that are unrelated (e.g. frill and barn). Bergen finds that the phonestheme condition (i) resulted in the quickest response rate and concludes that phonesthemes therefore have a detectable priming effect.

Parualt (Parault, 2006), an educational psychology researcher, conducted a series of experiments with a different approach to phonesthemes that explored the benefit of phonesthetic content to word learning. In the study, participants were asked to complete three tasks related to word recognition: (i) rating their recognition of groups of obsolete sound-symbolic and non-sound-symbolic English words, (ii) providing written definitions for each word, and (ii) matching words with definitions from multiple options. The intention was to demonstrate whether sound symbolic content was informative to word learners about the meaning of the word. The experiments in the study found that participants were more likely to display partial or full knowledge of the meaning of a word when asked to generate definitions for the obsolete words that had sound-symbolic content than when completing the same task for obsolete words that did not. Participants were also more likely to select the correct definition for words with sound-symbolic content than for those without.

Shih, Ackerman, Hermalin, Inkelas, and Kavitskaya (2018) demonstrated that sound symbolism is also productive by showing that the names of Pokémon in the popular Nintendo–Game Freak franchise take advantage of sound symbolism[5]. By applying statistical methodologies, the team of researchers show that the English names of heavier Pokémon (per in-game descriptions) were more likely to have names with a higher number of segments, and more likely to include voiced obstruents, and low vowels. Kawahara and Moore (2021) built on these findings by creating nonce names, coupled with nonce Pokémon illustrations. In a series of experiments, participants were presented with pairs of pre-evolution and post-evolution nonce Pokémon, a pair of nonce names, and asked to choose which name suited the post-evolution Pokémon best[6]. In each experiment the conditions were crafted to test the findings by Shih and colleagues (2018). Although the level of correlation varied for the different conditions, the results generally verified the previous findings. This conclusion shows that sound symbolism can be productive and indicates the value of identifying sound symbolic features to later research and to naming applications.

Similar research on the sound symbolic content of Japanese Pokémon names has also been undertaken and found similar patterns. (Shih et al., 2018; Kawahara & Kumagai, 2019). Indeed, much of the recent cognitive research into Japanese sound symbolism has been cross linguistic.

The cognitive research above illustrates that speakers of a language generally lack the meta-awareness of phonesthemes that speakers have for other components of words. NLP models that work with large linguistic datasets to identify patterns that otherwise go unnoticed are extremely valuable for features of language like this. Liu, Levow and Smith's 2018 study utilized a word embedding model to automatically detect phonesthemes in English. Such models build on the distributional hypothesis that similar words will occur in similar contexts by representing the meaning of a word (a word embedding) as a number corresponding to its location in a multi-dimensional space (a vector). Words that have similar vectors are therefore assumed to be similar semantically.

In this paper, I present an adaptation of Liu's approach for Japanese, discuss potential routes of conventionalization of the model-predicted sound symbolic clusters, and propose methods based on previous literature for concretely establishing their semantic component and empirically testing their psychological reality.

The Model

What follows is an overview of the version of Liu's model for automatic identification of phonesthemes adapted for Japanese.

Method

For this experiment, I used word2vec to build a word embedding dictionary trained on a corpus of all Japanese Wikipedia articles as of August 8, 2022, comprising approximately 1.3 million articles – roughly 64.6 million sentences and 1.2 billion words. word2vec relies on tokenization of individual

[5]Examining Pokémon names has in fact been a relatively rich vein for sound symbolism research, such that it has even been referred to as a field of its own, coined as Pokémonstatics by Shigeto Kawahara.

[6]In the Pokémon franchise, Pokémon change form after gaining a certain amount of experience. The form change is usually correlated with a size and power increase. This change is referred to as evolution. Different forms in an evolution line have different, though usually related names: e.g. Pikachu evolves into Raichu.

Table 1: A random selection of translated words from the clusters generated by K-means clustering, and an associated semantic domain.

Semantic domain	Terms	Cluster no.
Sports	sports meet, overall victory, match, participation (in a match), rank (esp. martial arts), baseball, seconds, championship title, female competitor, completion, victory	0
Politics	the world, society, counter for people, culture, city, era, history, deity, politics, science	1
Industry	plan, manufacture, winning (a prize), English language, United States of America, neighborhood, onset, division (of a larger group), discovery, research	2
Music	speech, role (character), (writing) song lyrics, volume (of a text), musical arrangement, teaching, script, starring (in a film, play etc.), director, performance	3
Television	TV program, musical theme, recording (audio or visual), stock company, work (e.g. a film, book etc.), release (e.g. publication), commemoration, song, synthesis, theatre	4
Military	battle, conflict, unclear, space, condition, forces (e.g. troops), continuation, protagonist, adversary, mobilization	5
Engineering	machine, vehicle, body, technology, surface, type, thing, use, both, place (location)	6
Municipal	Osaka, primary school, department (suffix), number (in a series), district of a town, municipal, store (shop), four (4), middle school, public building	7
Identity, life	voice, self, woman, man, daughter, name, life (esp. day-to-day life), wife, marriage, child	8
Bureaucracy	education, business, yen, incident, fiscal year, money, issue (problem), locality, side, project (plan)	9

words. As Japanese lacks word spacing, tokenization is a non-trivial task. I applied the *janome* text segmentation library (Uchida, 2015), which uses a dictionary look-up approach to tokenize Japanese content[7]

Words that occurred fewer than 1000 times were removed per Liu and colleagues' approach, motivated by Hutchins (1998) findings that phonesthemes are unlikely to be found in uncommon or very common types. To eliminate very common words, Liu removed types that appeared in more than half of the documents in their corpus. For this dataset, the same measure did not result in a satisfactory lexicon of types to eliminate (33 types, some of which appeared in more than 97% of articles) from the trained vectors. As such, the limit was gradually increased, inspecting the list each time until types that were reported in other literature as being of interest to sound symbolism research appeared[8]. Numbers, stopwords and types containing anything other than kanji and hiragana were removed from the trained vectors. Katakana types were removed assuming they were unlikely to contribute to phonesthetic sequences[9]. The resulting vector dictionary contained 484,497 types.

10 clusters of vectors located in close proximity in the 900-dimension were identified using k-means clustering via the kmeans module from the scikit-learn library. The resulting clusters were examined for semantic grouping. Table 1 contains a random sample of terms within each cluster, translated into English, and with an associated domain based on my own intuition.

The *Pykakasi* library (Miura, 2011) was applied to transliterate kanji types to hiragana to facilitate phonetic analysis of the vectors[10].

Inspired by the findings by Yoshida et al. (2023) that sound-symbolic content may occur within the first two consonants of kanji readings, I focus on the first two sounds (represented by morae) in Japanese words in general. Accordingly, I extracted each unique word-initial kana bigram for each word in a cluster.

Word-initial bigrams that are unlikely to be sound symbolic were systematically excluded. The bigrams had to meet the following criteria to be considered: (i) Found in words that occurred more than 1000 times in the corpus, (ii) Found in words

[7]Tokens processed by *janome* are matched against a precompiled dictionary that contains information about the linguistic properties of words in Japanese. For many tokens, there will be multiple options for segmentation included in the dictionary. To select the most likely option, *janome* relies on the Viterbi algorithm. The Viterbi algorithm examines a sequence of words and predicts the most likely outcome of the current state. This is not a perfect solution for Japanese and incorrectly segmented words and fragments were evident from an examination of the corpus. Unfortunately, word boundary identification is a difficult and sometimes subjective task for Japanese, and solving this problem is not within the scope of this study.

[8]In this case, 父, chichi, father. The English term "papa" was part of a study into cross-linguistic parallels in familial terms by Murdock (1959) that was later taken up and discussed in terms of iconic sound symbolism by Jakobson (1962).

[9]Katakana is primarily used for loanwords, neologisms and sounds (i.e. directly mimetic onamatopes) but there are exceptions. For this experiment, I reasoned that these exceptions are likely to be represented in the corpus and resulting vectors in other forms. For instance, animal names are frequently rendered in katakana, but inspection of the corpus showed that the animal names tested were universally included in the corpus as kanji and hiragana as well.

[10]Like *janome*, *Pykakasi* relies on a dictionary lookup approach, which entails the same issues. Using a human-transliterated excerpt of the corpus as the gold standard for comparison with the same excerpt transliterated by *Pykakasi* resulted in 95% accuracy. This is certainly low compared to comparable tasks such as modern part-of-speech tagging, which achieves accuracy of 97% or higher (Manning, 2011), but must be accepted for the scope of this study.

that occurred in fewer than 8% of articles in the corpus, (iii) Not found in homophones, (iv) Occurred more than 5 times in a cluster, (v) Found in a cluster statistically more frequently than in the corpus, (vi) Did not occur in more than one cluster considered. This assumes that a bigram cannot be phonesthetic for more than one meaning (vii) Associated with more than five kanji (e.g. all words beginning *kachi* were removed from consideration, as every example had been transliterated from words beginning with 勝, meaning victory). This measure is intended to mitigate interference due words with similar meanings and sounds likely being derived from the same word root.

Results and Analysis

The output of the model, filtered by the above constraints, resulted in nine candidate phonestheme-like bigrams in four semantic domains. The results are shown in Table 2.

Table 2: Transliterated word-initial bigrams that occur in only one cluster with a frequency that is statistically significantly higher than in the source corpus.

Semantic domain	Word-initial bigrams	Cluster no.
Politics	kou	1
Engineering	hei, tan, setsu, sui	6
Municipal, addresses	tou	7
Bureaucracy	kaka, kan, sou	9

I examined the list of words in which each bigram appeared within a cluster for any semantic similarities. I selected and compared only words that began with different kanji. In most cases, a close semantic grouping was not obvious. It was possible to identify two bigrams of interest.

The bigram *tou* from Cluster 7 (municipal, addresses) appears to have an association with height.

Table 3: A list of words featuring the initial sound tou, illustrating a potential semantic relationship.

Word	Reading	Meaning
塔	tou	Tower
峠	touge	Highest point on a mountain trail
東宮	touguu	Crown prince
杜氏	touji	Chief brewer at a sake brewery
棟梁	touryou	Central figure, boss
桃源	tougen	Paradise

Tower and *highest point on a mountain road*, relate to physical height. The next three terms are related to status or superiority. To borrow from conceptual metaphors known in English and Japanese, the inclusion of these examples in this group may be a result of (HIGHER) STATUS IS UP. Paradise also has an association with height in English, with heaven conceived as up, or above, potentially derived from VIRTUE IS UP (Kövecses, 2010). The sound *tou* may have an association with physical and metaphorical height[11]. Accordingly, this suggests *tou* is conventionalized, relying on conceptual metaphor as its route of conventionalization. More specifically, this phonesthetic-cluster may be derived from 東, "east", and its association with Tokyo symbolizing importance. Certainly, this appears to be true crown prince (*touguu*) which combines the character 東, "east" and 宮, "imperial residence". Alternatively, this may be motivated by an association with 桃, "peaches". Peaches feature prominently in Japanese folklore in the story of *Momotarō*, a boy born from a giant peach who becomes a hero. Paradise (*tougen*) is composed of the character 桃, "peach" and 源, "origin".

There are two potential associations for the bigram *kan* from Cluster 9 (business, bureaucracy).

Table 4: A list of words featuring the initial sound kan, illustrating two potential semantic relationships.

Word	Reading	Meaning
官庁	kanchou	government office
管区	kanku	jurisdiction
幹事	kanji	executive secretary
監理	kanri	supervision (administration)
関税	kanzei	tariff
患者	kanja	patient
感染	kansen	infection
看護	kango	nursing
肝炎	kanen	hepatitis

The first grouping of words in Table 4 has an official or bureaucratic association, which matches the semantic domain assigned to cluster in which this bigram was found. The second group however has a medical connotation.

Since multiple phonesthetic meanings of the *kan* bigram are unlikely, further research would be required to identify which association is salient. It is worth noting that the *bureaucracy* association of *kan* is very similar to Yoshida and colleagues' (2023) findings that words beginning with /k/ in Japanese have a high affinity with the WordNet-based semantic label of *control*.

Discussion

Future Directions

There are multiple ways to test the meaning of these bigrams to language users. Liu et al. (2018) surveyed participants about how well a real word containing the candidate bigrams matched its definitions versus words that do not contain the

[11] It must be acknowledged that this association is not fully consonant with the semantic domain of addresses and municipal features that its cluster appears to relate to. Its inclusion in this cluster may be related to physical location, and the over-representation of 東, "east", in words associated with Tokyo as a city name.

bigrams. Preparing nonce words containing the bigrams (after Sapir (1929); Kawahara and Moore (2021)) and asking participants to generate definitions or match the nonce words with definitions (after Parault (2006)) would be a valuable direction to explore. Tanno (2003) asked participants to describe how individual morae made them feel: it may be possible to adapt this approach by asking participants exposed to the candidate sounds to pick between semantic domains and identifying trends. This design would however be affected by the limitations of forced-choice tasks, which research has shown to potentially exaggerate, or bias findings made this way: even with the familiar *bouba-kiki* paradigm, presenting participants with more than a binary choice has been found to result in weaker correlation between target stimuli and expected associations (Dingemanse, 2015).

A procedure based on Bergen's use of primer experiments would help establish meanings associated with these clusters, as well as provide useful data about whether the clusters effect a participant's processing of the word (after Bergen (2004)). In this instance, as sound-symbolic structures in Japanese like those found in this study are not otherwise widely reported on or researched, establishing their psychological reality in this way would be crucial.

Methodological Insights

The bigrams identified as potentially phonesthetic by this model have not previously been reported in the literature to the best of my knowledge. Phonestheme-like structures like the ones tentatively identified in this study are not yet recognized as a part of the sound-symbolic system of Japanese. The finding that there are under-reported or unreported patterns like these in Japanese is a strong indicator that the approach of using word embeddings to automatically identify trends in linguistic data that speakers lack awareness of is one with tremendous potential.

Using word embedding algorithms like word2vec in isolation is increasingly regarded as obsolescent or obsolete (Kokab, Asghar, & Naz, 2022). However, similar word embedding algorithms are foundational components of transformers such as ELMo (Peters et al., 2018) and BERT (Devlin, Chang, Lee, & Toutanova, 2019) which are able to account for polysemy and homonymy. Adapting this paper's approach to use a transformer-based model would be a logical next step.

Sound Symbolism and Language Evolution

Many researchers have suggested that sound symbolism may have been crucial to the development of language. Imai and Kita (2014) provide an overview of the theory that *iconic* sound symbolism may have played a role in the evolution of human language. The theory that sound symbolism played a part in language evolution has been substantiated via studies of language acquisition by children. Kantartzis et al. (2011) explicitly state that the performance of children exposed to sound-symbolic words in their word-learning task may be due to the role of sound symbolism in language evolution.

Monaghan, Shillcock, Christiansen, and Kirby (2014) demonstrated that the distribution of sound symbolism is even across nouns and Verbs in English. Imai and Kita (2014) build on this by illustrating that the same is not true of Japanese, where object names are less likely to be sound symbolic. Imai et al. therefore suggest that distribution of sound symbolic words may vary by language. In the case of Japanese, they suggest that this may be due to the presence of a specific category of sound-symbolic words (the ideographs discussed earlier). The phonestheme-like features identified by this model may represent a facet of sound symbolism in Japanese that is otherwise unaccounted for in the imbalance identified by (Imai & Kita, 2014). These features require more research to understand whether they represent a pattern in Japanese. If they do, they may be a remnant of iconicity in Japanese that is no longer recognized as such. Testing words containing the features identified by this model with children in a similar way to Kantartzis et al. (2011) might establish one way or another whether they similarly facilitate learning – a positive result would suggest that these are indeed an unrecognized remnant of iconicity and provide insight into the development of Japanese, and mechanisms involved in the development of language at large.

Conclusion

This paper addressed the issue of sound symbolism in Japanese – namely, it is concerned with Japanese phonesthemes. To begin to identify phonesthemes in Japanese, or groupings of sounds that have a similar effect on language users in Japanese, I adapt a model based on word embeddings originally used by Liu and colleagues Liu et al. (2018) in English. The model identifies a small number of candidate phonestheme-like sounds in Japanese in the form of word-initial bigrams. Two of these warrant further investigation: *tou* and *kan*. Both *tou* and *kan* can be categorized as conventionalized phonestheme-like clusters. *tou* is associated with both physical and metaphorical height, with the latter possibly motivated by conceptual metaphor. *kan* likely has an association with bureaucracy, but also frequently appeared associated with a medical context; previous research has suggested a link between words beginning with the phoneme /k/ and a meaning of *control*, but more testing would be required to establish the salient meaning.

Much of the previous research has illustrated that sound symbolism including phonesthemes may facilitate language acquisition, even cross-linguistically (Parault, 2006; Kantartzis et al., 2011; Iwasaki et al., 2007). Applying methods used by previous researchers such as priming tests and definition assignment tasks to establish the psychological reality of the model-identified candidates would in turn make it possible to empirically test whether they similarly facilitate word acquisition. Crucially, studying language acquisition offers insights into language evolution. If the features identified here, like other sound-symbolic traits, aid acquisition, they may help explain how Japanese – and language in general – developed.

Acknowledgments

I would like to thank Drs. Anastasia Smirnova and Jenny Lederer at San Francisco State University for their advice, feedback and patience during the process of writing this paper. This paper also would not have been possible without the coding assistance and emotional support of my partner Jared Allen. I am grateful as well to the anonymous reviewers of Cognitive Science for their indispensable feedback. Any errors or deficiencies are my own.

References

Asano, T. (Ed.). (1978). Giongo/gitaigo-jiten *[onomatopoeia and mimetic words: A dictionary]*. Tokyo, Japan: Kadokawa.

Bergen, B. (2004). The psychological reality of phonesthemes. *Language*, *80*(2), 290–311.

Bolinger, D. L. (1950). Rime, assonance, and morpheme analysis. *WORD*, *6*(2), 117–136.

Chang, A. C. (1990). *A thesaurus of japanese mimesis and onomatopoeia: usage by categories*. Tokyo, Japan: Taishukan.

Chomsky, N. (1957). *Syntactic structures*. The Hague, Netherlands: Mouton de Gruyter.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the north american chapter of the association for computational linguistics: Human language technologies* (pp. 4171–4186). Minneapolis, MN: Association for Computational Linguistics.

Dingemanse, G. L. . M. (2015). Iconicity in the lab: a review of behavioural, developmental, and neuroimaging research into sound-symbolism. *Frontiers in psyhcology*, *6*, 1–14.

Fujimoto, M. (2015). Vowel devoicing. In H. Kubozono (Ed.), *Handbook of japanese phonetics and phonology.* Boston, MA: De Gruyter.

Hamano, S. (1998). *The sound-symbolic system of japanese*. Stanford, CA & Tokyo, Japan: Center for the Study of Language and Information.

Hinton, L., Nichols, J., & Ohala, J. (Eds.). (1994). *Sound symbolism*. New York, NY: Cambridge University Press.

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

Hutchins, S. (1998). *The psychological reality, variability and compositionality of english phonesthemes*. Doctoral dissertation, Emory University.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B*, *369*.

Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007). What do english speakers know about gera-gera and yota-yota?: A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese-Language Education around the Glove*, *17*, 53–78.

Iwasaki, N., & Yoshida, K. (2019). Iconicity in l2 japanese speakers' multi-modal language use. In K. Akita & P. Parudeshi (Eds.), *Ideophones, mimetics and expressives.* Amsterdam, Netherlands: John Benjamins Publishing Company.

Jakobson, R. (1962). Why "mama" and "papa"? In R. Jakobson (Ed.), *Roman Jakobson selected writings I, phonological studies.* Gravenhage, Netherlands: Mouton.

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in english-speaking children. *Cognitive Science*, *35*, 575–586.

Kawahara, S., & Kumagai, G. (2019). Expressing evolution in pokémon names: Experimental explorations. *Journal of Japanese Linguistics*, *35*(1), 3–38.

Kawahara, S., & Moore, J. (2021). How to express evolution in english pokémon names. *Linguistics*, *59*(3), 577–607.

Kokab, S. T., Asghar, S., & Naz, S. (2022). Transformer-based deep learning models for the sentiment analysis of social media data. *Array*, *14*.

Kövecses, Z. (2010). *Metaphor* (Second ed.). New York, NY: Oxford University Press.

Liu, N. F., Levow, G., & Smith, N. A. (2018). Discovering phonesthemes with sparse regularization. In *Proceedings of the second workshop on subword/character level models* (pp. 49–54). New Orleans, LA: Association for Computational Linguistics.

Manning, C. D. (2011). Part-of-speech tagging from 97% to 100%: Is it time for some linguistics? In *Conference on intelligent text processing and computational linguistics* (pp. 171–189). Heidelberg, Germany: Springer Berlin.

Miura, H. (2011). *Pykakasi.* Github. Retrieved from `https://github.com/miurahr/pykakasi`

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B*, *369*.

Murdock, G. P. (1959). Cross-language parallels in parental kin terms. *Anthopological Literature*, *1*(9), 1-5.

Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In J. N. L. Hinton & J. Ohala (Eds.), *Sound symbolism.* New York, NY: Cambridge University Press.

Parault, S. J. (2006). Sound-symbolism: a piece in the puzzle of word learning. *Journal of Psycholinguistic Research*, *35*, 329–351.

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). *Deep contextualized word representations.* Retrieved from `https://arxiv.org/abs/1802.05365`

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*(3), 225–239.

Saussure, F. (1916). *Course in general linguistics*. London, England: Duckworth.

Sereno, J. A. (1994). Phonosyntactics. In J. N. L. Hinton & J. Ohala (Eds.), *Sound symbolism.* New York, NY: Cambridge University Press.

Shih, S. S., Ackerman, J., Hermalin, N., Inkelas, S., & Kavitskaya, D. (2018). Pokémonikers: A study of sound symbolism and pokémon names. In *Proceedings of the linguistic society of america vol. 3* (pp. 1–6). Washington DC: Linguistic Society of America.

Tanno, M. (2003). *Nihongo onin ni okeru onshōchō no sonzai* [the existence of sound symbolism in japanese phonology]. *Studies in childhood education*, *22*(3), 1-10.

Uchida, T. (2015). *janome.* Github. Retrieved from `https://github.com/mocobeta/janome`

Yoshida, A., Matsuhira, C., Kato, H., Hirayama, T., Komamizu, T., & Ide, I. (2023). Discovering phonesthemic clusters in readings of kanji characters toward exploring phonestheme in japanese. In *Proceedings of the 37th pacific asia conference on language, information and computation* (pp. 322–337). Hong Kong, China: Association for Computational Linguistics.