

PROSODIC SYSTEMATICITY MEETS LANGUAGE DEVELOPMENT: SKEWED TONAL INVENTORIES IN NON-ARBITRARY LANGUAGE

Thompson, Arthur Lewis^a

^aDepartment of Linguistics, University of Hong Kong, Pok Fu Lam Road, Hong Kong

Abstract

Human language is not entirely arbitrary. Recent advances in the literature have shown sound symbolism to facilitate word-learning and acquisition cross-linguistically. Most research focuses on the segment level and not much is known about the prosody or the role it plays in the sound symbolism of tonal and non-tonal languages alike. By comparing sound-symbolic corpora with general lexicon corpora, this study shows that the sound-symbolic lexicons of three Sinitic varieties (Mandarin, Taiwanese Southern Min, Hong Kong Cantonese) are skewed so that the majority of sound-symbolic syllables are in high tone while the general lexicon is not. Building on the idea that gesture and eye movement facilitate word-learning, it is proposed that prosodic demarcation of descriptive and non-arbitrary words is analogous to a gesture in the speech stream which helps the listener (child) pinpoint its referent. This systematic mapping may be derived from the lexical embodiment of expressive voice like that found in Infant Directed Speech.

Keywords: sound symbolism, tone, iconicity, language acquisition, language evolution, Chinese

total page count: 20

1.0 Introduction¹

Sound symbolism defies de Saussure's (1916) commonly accepted dogma of language as arbitrary (where words bear no resemblance to their physical referent, e.g. 'dog'), demonstrating that features of the outside world are indeed replicated in the development and structure of certain linguistic features. How the features of the outside world are replicated in speech has been further organized according to (often overlapping) feature/function-based classifications: onomatopoeia, ideophones, size sound symbolism, etc. Recent studies have carried out experimental research into these classifications with phonological, language evolution, and acquisition-related claims in mind (see Lockwood & Dingemanse 2015 for a review). Most of these studies focus on the segment level, leaving much to be said about the prosodic level. By using corpora methods to examine Sinitic tonal inventories, this paper aims to explore the prosodic role in systematic organization and communicative strategies of sound symbolism in language acquisition and evolution.

1.1 Sound Symbolism

There are two major subcategories under the umbrella of sound symbolism. One fairly straightforward category is onomatopoeia. Hinton et al. (1994) classify this as imitative sound symbolism because onomatopoeia mimic sounds from the physical world. Take the following onomatopoeia for dogs barking: *woof woof* (English), *wàng wàng* (Mandarin), *wōng wōng* (Cantonese), *wan wan* (Japanese). These instances of sound symbolism are directly based on the linguistic interpretation of a physical referent (dog bark). The other major category is the ideophone, otherwise known as 'mimetic' or 'expressive'. Ideophones are based on a description of physical or emotional means (Dingemanse 2012). That is to say, they do not necessarily imitate a physical sound but a psychological one. The Cantonese ideophone *kàuh kàuh kèih kèih* describes an action done in a careless manner, it evokes a sensory or emotional depiction of its referent to both the listener and speaker, despite the lack of physical sound for acting in a careless way. Many languages, such as Japanese (Hamano 1998), Korean (Kim 1977), Mandarin (Li 2007), and Cantonese (Bodomo 2006), categorize ideophones and onomatopoeia into one sound-symbolic word class. Both onomatopoeia and ideophones are structured by how language encodes iconicity and systematicity.

1.2 Iconic Properties

Iconicity is the use of perceptuomotor analogy to connect a word to its referent (Dingemanse et al. 2015). In this way, iconicity is not arbitrary because inherent properties of the referent play a major role in determining the word's form and can be expressed through a number of modes, most of which are language-specific. Vowel quality has been shown, notably in kiki-bouba tests, to depict referent size (Bremner et al. 2013; Thompson & Estes 2011; Maurer et al. 2006) through analogy by widened or contracted dimensions of the oral cavity (Oda 2000) or pitch association (Ohtake & Haryu 2013). Repetition and reduplication has been analogically used to refer to the duration or iterative state of an action (Li 2007; Elders 2002). These are but a few examples showing that iconicity partnered

¹ Abbreviations: HKC = Hong Kong Cantonese; TSM = Taiwan Southern Min; SS = Sound Symbolism, sound-symbolic; GL = General Lexicon; HRS = High Register Systematicity.

with phonology determines the structure of sound-symbolic words but the application and extension thereof is guided by systematic means.

1.3 Systematic Properties

While iconicity comprises the content of sound-symbolic words, systematicity is a guiding principle or method which demarcates a class of words from others in the lexicon (Perniss & Vigliocco 2014). Though applicable to iconicity, systematicity is grounded in arbitrariness since its function and form are not necessarily determined by perceptumotor analogies based on the referent. Systematicity aids language acquisition by providing a regular pattern for learners to create categorical correspondences between sound and meaning (Dingemanse et al. 2015: 608). How systematicity is applied seems largely language dependent. It is apparent in morphological phenomena such as verbal suffixes (*-ić*, *-yć* Polish verbs), sub-morphemic phenomena such as phonaesthemes (English: *gl-* gleam, glow, glisten, and *sn-* snigger, snuffle, snore) (Bergen 2004), and perhaps even regularities in phonological patterning (Monaghan et al. 2007). Non-segmental, prosodic, distinctions have also been shown to distinguish grammatical class (Kelly 1992; Monaghan et al. 2005). As explained later, prosody also plays a major role in the systematicity of Sinitic sound symbolism.

1.4 Recent Studies

Establishing iconic relationships between sound and different physical referents has been carried out in previous studies cross-linguistically (Oda 2000; Ramachadran & Hubbard 2001; Hirata et al. 2011; Nielsen & Rendall 2012, 2013; Ohtake & Haryu 2013). Recent studies take a more evolutionary and acquisitional approach by proposing how sound symbolism facilitates language acquisition and communication (Kantartzis et al. 2011; Yoshida 2012; Laing 2014; Imai & Kita 2014; Imai et al. 2015). In Sinitic, sound symbolic research is largely Mandarin-centric, descriptive, and restricted to orthographical means with emphasis on multi-syllabic phrasal structures (Wang 1992; Chan 1996; Li 2007; Chu 2012; Lai 2015; Thompson 2016). Prosody is thus a new direction for both general and Sinitic sound-symbolic research.

2.0 Methods

The methods in this study are based on existing corpora. One corpus of sound-symbolic words was consulted per variety of Sinitic: Mandarin, Hong Kong Cantonese (HKC), Taiwan Southern Min (TSM). See Table 1 for a description of tones in each of the Sinitic varieties examined. Corpus data for each variety's general lexicon was used for comparative purposes. In each corpus, tokens were counted per tone, separately for the sound-symbolic class and the general lexicon. The goal of counting tokens was to examine whether the tonal distribution is different between the sound-symbolic class and general lexicon per Sinitic variety. A chi-squared test was used to compare tonal distribution between sound-symbolic and general lexicon datasets. Statistical inference was corrected for multiple comparisons using Bonferroni correction across 20 tests (1 test of general distribution per variety, as well as 4, 6, and 7 single-tone tests for Mandarin, HKC, and TSM respectively) at a threshold $\alpha = 0.05$.

Table 1: Descriptions of Tonal Inventory per Variety of Sinitic (Bauer & Benedict 1997: 49; Wu 2000: 2691)

	T1	T2	T3	T4	T5	T6	T7	T8
Mandarin	High level	High rising	Dipping	High falling				
Hong Kong Cantonese	High level	Mid rising	Mid level	Low falling	Low rising	Low level		
Taiwan Southern Min (Taipei)	High level	High falling ²	Low falling	Mid stopped	Rising	High falling	Mid level	High stopped

The Mandarin onomatopoeic corpus is Gong's (1991) dictionary of Mandarin Onomatopoeia which consists of 188 different syllable types. The Mandarin general lexicon data is from van de Weijer and Sloos's (2014) corpus-based survey of the 500 most common words, totalling to 759 syllables. The HKC sound-symbolic corpus was Bodomo's (2008) *A Corpus of Cantonese Ideophones* which consists of 174 different syllable types³. The HKC general lexicon data is taken from the Hong Kong Cantonese Adult Language Corpus cited in Leung et al. (2004: 503), consisting of 128,345 syllables. The TSM onomatopoeic corpus was Hung's (2007) survey of TSM onomatopoeia which consists of 824 different syllable types. TSM general lexicon data is from the Taiwan Southern Min (TSM) Corpus 1.0, an open access source, compiled by Ching Chu Sun and John Newman at the University of Alberta, which consists of 9,131 syllables.

The sound-symbolic corpora cited above were designed only to show sound-symbolic utterances available per variety of Sinitic (i.e. the number of contexts in which a single given sound-symbolic word might be used). In this study, sound-symbolic tokens were counted according to syllable type instead of syllable frequency in order to avoid redundancy. For example, no matter how many times *ma1* appears in a sound-symbolic corpus, it was only counted as one point in the Sound-symbolic Tone 1 category tally. In this way, the sound-symbolic corpora have been counted according to the tonal content of a variety's sound-symbolic syllable inventory.

Conversely, counting the general lexicon per syllable type would lead to a uniform tonal distribution, since almost all tonal categories can map onto every syllable type⁴. In this study, general lexicon tokens were counted according to frequency (i.e. how many times a syllable appeared throughout a corpus). For example, if *ma1* occurs only twice in a general lexicon corpus, then only two points are awarded to the General Lexicon Tone 1 category tally. Hence, two different counting methods (sound-symbolic tokens = syllable type; general lexicon = syllable frequency) were implemented in

² Due to historical reasons T2 and T6 of Taiwan Southern Min (Taipei) have merged.

³ 73 additional syllable types were excluded. These are not sound symbolic words but adjectives and/or adverbs which exhibit a similar multi-syllabic structure to sound symbolic utterances.

⁴ Since Bodomo (2008) cites Cantonese checked tones according to their underlying form, T7 – T9 were not taken into account for this study.

this study.

As a control analysis, counting syllable frequency in Bodomo's (2008) HKC ideophone corpus as well as Hung's (2007) TSM onomatopoeia corpus was carried out in order to test whether the difference in counting methods skews the tonal distribution between datasets. The control analysis consisted of 538 syllables for HKC and 4,197 syllables for TSM. The same method of correcting for multiple comparisons was used in the control as well as main analysis. In the control analysis 15 tests have been corrected for (tonal distribution in HKC and TSM; additionally 6 and 7 single tone tests for HKC and TSM respectively).

3.0 Results

For Mandarin, the overall distribution of tones within the sound-symbolic (SS) corpus is significantly different from the distribution of tones within the general lexicon (GL) corpus (Figure 1). The relative frequency of Mandarin T1, T3, and T4 is significantly different between the SS and GL corpora (see Table 2 for χ^2 test results) after correcting for multiple comparisons. Asterisks mark tests surviving correction for multiple comparisons (see also results for HKC and TSM).

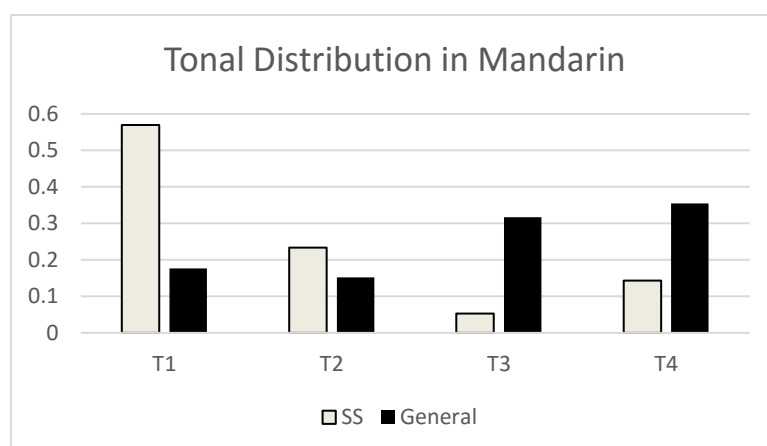


Figure 1, Tonal Distribution in Mandarin: the overall tonal distribution is significantly different between corpora, with the majority of the SS corpus skewed to T1.

Table 2: χ^2 test result for Mandarin datasets

	SS tone count	GL tone count	χ^2	DF	Significance
All Tones	188	759	147.4	3	<.001*
T1	107	106	111.8	1	<.001*
T2	44	91	6.84	1	n.s.
T3	10	190	52.47	1	<.001*
T4	27	213	30.2	1	<.001*

For HKC, the overall distribution of tones within the SS corpus is significantly different from the distribution of tones within the GL corpus (Figure 2). The relative frequency of HKC T1, T2, T4, T5, and T6 is significantly different between the SS and GL corpora (see Table 3 for χ^2 test results). After correcting for multiple comparisons, the relative frequency of T1, T2, and T4 remains

significantly different between the SS and GL corpora (see Table 3).

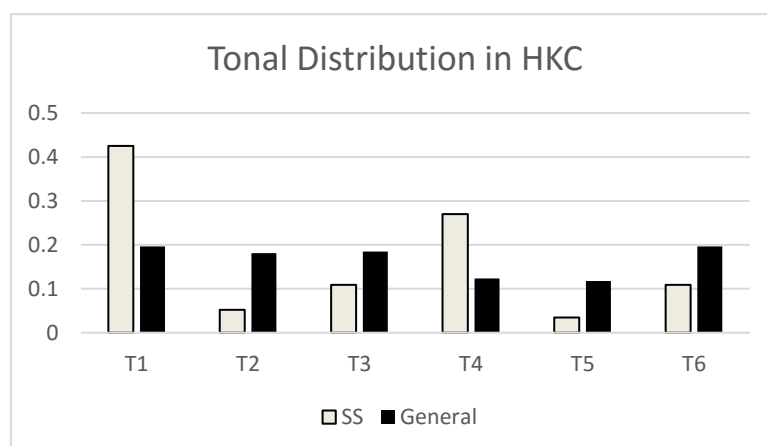


Figure 2, Tonal Distribution in HKC: the overall tonal distribution is significantly different between corpora, with the majority of the sound-symbolic corpus skewed to T1 and T4.

Table 3: χ^2 test result for HKC datasets

	SS tone count	GL tone count	χ^2	DF	Significance
All Tones	174	128345	114.89	5	<0.001*
T1	74	25193	57.69	1	<0.001*
T2	9	23279	19.69	1	<0.001*
T3	19	23693	6.57	1	n.s.
T4	47	15869	34.36	1	<0.001*
T5	6	15093	11.58	1	<0.01
T6	19	25218	8.39	1	<0.05

For TSM, the overall distribution of tones within the SS corpus is significantly different from the distribution of tones within the GL corpus (Figure 3). The relative frequency of TSM T2, T4, T5, and T8 is significantly different between the SS and GL corpora (see Table 4). All these tests survive correction for multiple comparisons.

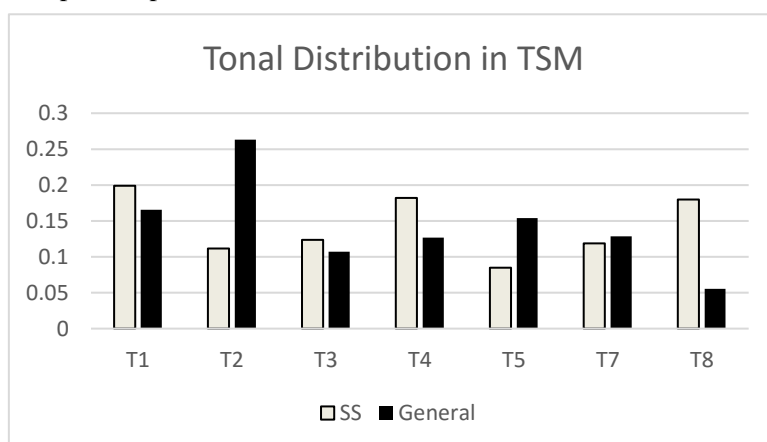


Figure 3, Tonal Distribution in TSM: the overall tonal distribution is significantly different between corpora.

Table 4: χ^2 test result for TSM datasets

	SS tone count	GL tone count	χ^2	DF	Significance
All Tones	824	9131	296.67	6	<0.001*
T1	164	1512	6.04	1	n.s.
T2	92	2402	92.28	1	<0.001*
T3	102	977	2.2	1	n.s.
T4	150	1156	20.38	1	<0.001*
T5	70	1405	28.44	1	<0.001*
T7	98	1174	0.63	1	n.s.
T8	148	505	190.54	1	<0.001*

3.1 Control Analyses

For the HKC control analysis, tonal frequency plus syllable frequency was counted for SS syllables rather than tonal frequency according to syllable type. The overall distribution of tones within the SS corpus is significantly different from the distribution of tones within the GL corpus (Figure 4). The relative frequency of HKC T1, T2, T3, T4, T5, and T6 is significantly different between the SS and GL corpora (see Table 5). After correcting for multiple comparisons, the relative frequency of T1, T2, T4, T5 and T6 remains significantly different between the SS and GL corpora (see Table 5). T3 is only significant at an uncorrected threshold of $\alpha=.05$. Asterisks mark tests surviving correction for multiple comparisons.

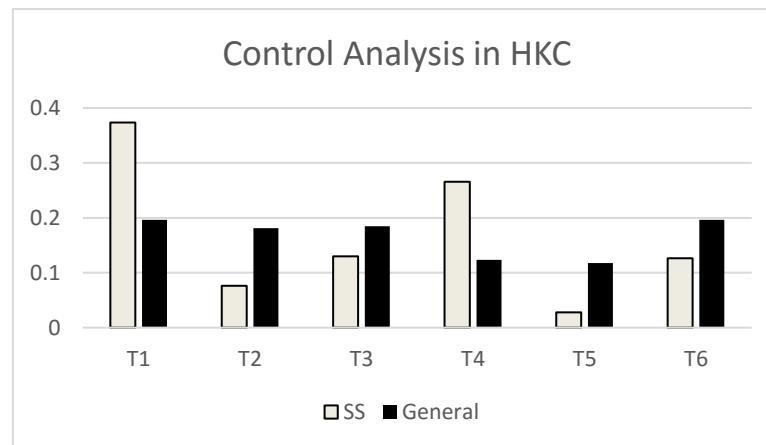


Figure 4, Control Analysis in HKC: results show that the two different counting methods implemented do not affect the significance of results. As with Figure 2, the overall tonal distribution here is significantly different between corpora, with the majority of the SS corpus skewed to T1 and T4.

Table 5: χ^2 test result for HKC control analysis

	SS tone frequency count	GL tone count	χ^2	DF	Significance
All Tones	538	128345	264.22	5	<0.001*
T1	201	25193	106.47	1	<0.001*
T2	41	23279	39.99	1	<0.001*

T3	70	23693	10.58	1	<0.05
T4	143	15869	99.51	1	<0.001*
T5	15	15093	41.67	1	<0.001*
T6	68	25218	16.69	1	<0.001*

Similarly, for the TSM control analysis, tonal frequency plus syllable frequency was counted for SS syllables rather than tonal frequency according to syllable type. The overall distribution of tones within the SS corpus is significantly different from the distribution of tones within the GL corpus (Figure 5). The relative frequency of TSM T1, T2, T4, T5, T7, and T8 is significantly different between the SS and GL corpora (see Table 6). Thus, the significance of T1 and T7 is dependent on the counting method implemented. After correcting for multiple comparisons, the relative frequency of T2, T4, T5, and T8 remains significantly different between the SS and GL corpora (see Table 6). Asterisks mark tests surviving correction for multiple comparisons.

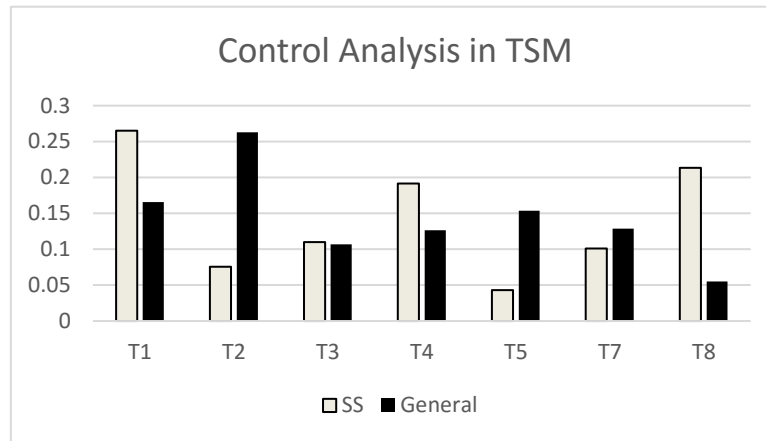


Figure 5, Control Analysis in TSM: results show that the two different counting methods implemented do affect the significance of results. As with Figure 3, the overall tonal distribution here is significantly different between corpora, however T1 and T7 are now significant after counting the SS corpus according to syllable frequency.

Table 6: χ^2 test result for TSM control analysis

	SS tone frequency count	GL tone count	χ^2	DF	Significance
All Tones	4197	9131	1720.33	6	<0.001*
T1	1113	1512	180.34	1	<0.001*
T2	317	2402	622.7	1	<0.001*
T3	462	977	0.28	1	n.s.
T4	804	1156	96.74	1	<0.001*
T5	181	1405	336.38	1	<0.001*
T7	425	1174	20.31	1	<0.001*
T8	895	505	762.99	1	<0.001*

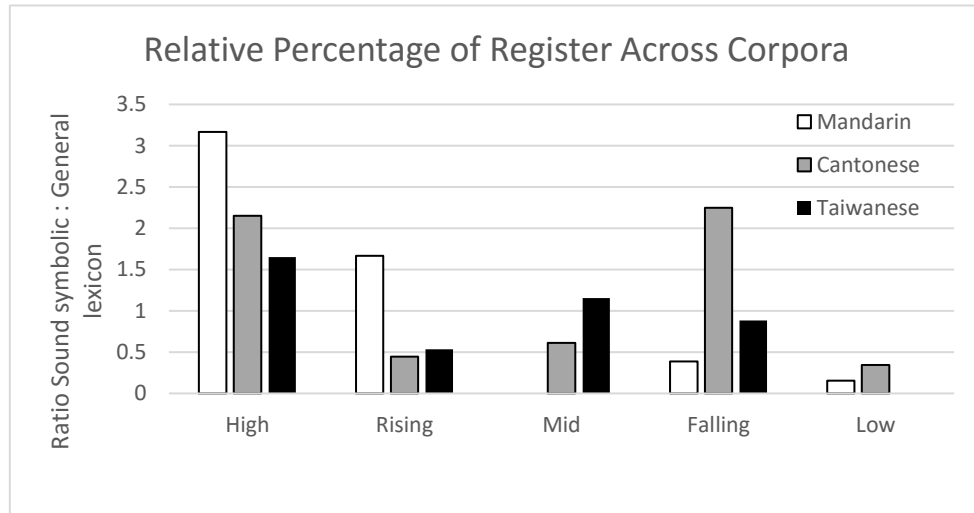


Figure 6, Relative Percentage Across Corpora: Tone registers scoring at 1 indicate that the percentage of said tone register is equal for both SS and GL corpora. Tone registers scoring above 1 indicate that the percentage of a given tone register is higher in the SS corpus than GL corpus. Scoring below 1 indicates that said tone register has a lower percentage in the SS corpus than GL corpus. The high tone register is the only register which is consistently more common among SS corpora than GL corpora cross-linguistically.

Table 7: Percentage of Tonal Distribution across Corpora According to Register

Tone (register)	General Lexicon			Sound-symbolic		
	Mandarin	HKC	TSM	Mandarin	HKC	TSM
High	18%	20%	23%	57%	43%	38%
Rising	15%	18%	15%	25%	8%	8%
Mid	N/A	18%	13%	N/A	11%	30%
Falling	36%	12%	39%	14%	27%	23%
Low	32%	32%	N/A	5%	11%	N/A

4.0 Discussion

Across three major Sinitic varieties, the majority of SS syllables are expressed in high register tones (Figure 6 and Table 7). This skewed tonal distribution is a product of systematicity. That is to say, the high register acts as identifying marker, or a pattern, systematically demarcating SS words from non-sound-symbolic words. Through iterated learning, this pattern thus signals the SS status of the utterance. This is not to say that systematicity is absolute or exclusive to phonology. This will also be addressed in the following sections. In addition, the origins of this high register systematicity (HRS) will be posited from both a phonological and evolutionary perspective. This prosodic systematicity, separate from segment-level systematicity, has a strong referential potential, much like gesture and eye movement, in the Bootstrapping Hypothesis for language acquisition and evolution (Imai & Kita 2014). The evolution of HRS itself reveals what linguistic stages exist between iconic and arbitrary on the continuum of arbitrariness. Finally, how prosodic systematicity, and what I call the Orthographic Opacity problem, could influence results of (past and future) SS research will be also discussed.

4.1 Tiers of Sound-symbolic Structure

Sinitic SS structure is not limited to phonology because it generally conforms to the syllabic structure of the variety in which it is found (Bauer 1985; Li 2007; Chu 2012). The segmental level supplies most of the iconic information associated with the referent while the prosodic level further conveys its SS status through HRS (Figure 7). Potential for expressivity is maximal here, in that Sinitic is using all available tiers (segmental, suprasegmental) to convey meaning.

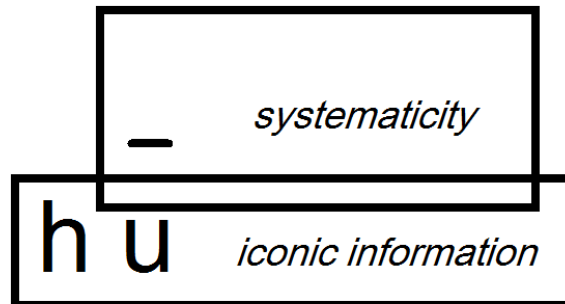


Figure 7, Sound-symbolic Tiers: The upper rectangle represents the prosodic level, which conveys systematicity, and the lower rectangle represents the segmental level, which conveys the iconic information determined by the referent. In non-SS words, both tiers are used for lexical purposes only.

Prosody is ideal for systematic purposes as it neither interferes with nor compromises iconic information at the segmental level. Further subdivision of systematic domains is left free to occur at the segmental level. For example, SS words to do with unobstructed airflow (wind, breathing, laughter) are often /x/ initial in Mandarin (Li 2007). Size SS, like that of kiki-bouba tests in the literature, is also free to be extrapolated⁵. Hence, this dual tier structure (Figure 7) allows for efficiency in signal which alerts the listener to the status of an utterance as SS through two simultaneous linguistic channels of processing.

There are, of course, exceptions. Systematicity is not absolute and not all SS words show HRS. The falling tone makes up the other majority of SS words in HKC. There are two potential reasons why this is the case: either 1) these prosodic exceptions possess iconic properties that somehow exempt them from HRS, or 2) the prosodic exception is not iconic but lexical and used to distinguish one word from another which is segmentally similar but semantically different.

⁵The author proposes that cross-linguistic size SS is not derived from iconic relationships expressed through oral cavity size but a systematic extrapolation of acoustical (vowel) associations previously acquired in onomatopoeia. Often smaller animals are assigned onomatopoeia with higher vowels (mice, cats, birds) than larger animals (dogs, cows, lions). Ohtake & Haryu (2013) findings support this as pitch was shown to be a stronger correlate than oral cavity size for size SS. Moreover Fernández-Prieto et al. (2015) found that six-month-olds make this pitch size association while four-month-olds do not.

4.2 Bootstrapping

Based on a number of acquisitional and perceptual SS studies, Imai and Kita (2014: 4) outline five claims in their Bootstrapping Hypothesis, the fourth and fifth of which are most relevant here: "(4) Sound symbolism helps infants associate speech sounds and their referents and establish a lexical representation. (5) Sound symbolism helps toddlers identify referents embedded in a complex scene." The present study shows that three varieties of Sinitic have maximalized said referential bootstrapping process by systematically lexicalizing the majority of SS words in high tone, known here as HRS, which is not attested in the general lexicon of these Sinitic varieties. This cross-linguistic trend shows that certain prosodic features can be systematically lexicalized to aid language acquisition. The consistent application of high tone may help the infant connect (SS) segments of the speech stream to the referent. Similar comprehension-oriented prosodic phenomena, though not lexicalized, have been attested in studies examining Infant Directed Speech (Nygaard et al. 2009; Herold 2011a, 2011b) as well as adult word learning (Reinisch et al. 2013). It seems unlikely that HRS alone would allow infants to make the referential connection of the Bootstrapping Hypothesis. Thus bootstrapping is a multi-modal approach: gesture (pointing, motion) and eye movement work together with HRS to aid in the process of language acquisition. HRS can be likened to a gesture presented in the speech signal which, by prosodically highlighting an utterance, aids in the connection the learner (i.e. child) must make to the referent. But where could HRS have originated from?

4.3 HRS and Language Evolution

SS and the use of SS cross-modal mapping has been proposed as a first step in the process of language evolution (Kita 2008; Kita et al. 2010). Studies have also shown that iconicity undergoes systematic patterning after acquisition and subsequent transmission to new speakers (Verhoeef et al. 2015). Therefore one might assume that the iconic information, at the segment level, came before the prosodic systematicity of HRS. Even so, how did HRS come into play? There are at least two potential answers to this question: 1) HRS was initially iconic (e.g. high register depicts smallness) before eventually being regularized to its current systematic status, or 2) HRS was once a feature of expressive voice, with no iconic connotations, before eventually fossilizing into its current systematic status.

Given that HRS is present in diachronically divergent and distinct varieties of Sinitic, it seems unlikely that HRS evolved from some (universal) association between SS and tone. The absence of SS at the current prosodic level also fails to support this hypothesis. Moreover, this assumes that at one point the prosodic tier detached from the iconicity of the whole SS word and evolved into its current systematic tier. Though the systematic application of iconic features in language is well attested (Thiesen et al. 2010; Imai and Kita 2014; Imai et al. 2015; Dingemanse et al. 2015; Verhoeef et al. 2015), it is worth noting that this split would have had to occur very early on, before Sinitic would have branched into fully-fledged and independent varieties, to remain diachronically consistent. Otherwise how could one explain how such a split might have occurred across three varieties separately without minimal interaction?

A more likely answer to the development of HRS lies in our second choice, that which is driven by an emphatic or expressive voice. Indeed, Dingemanse (2013) notes that most ideophones are prosodically emphasized in natural speech. Considering that SS is found in mother-to-child speech (Suzuki 2013; Saji & Imai 2013; Akita 2009; Ogura 2006; Fernald & Morikawa 1993), let us consider Infant Directed Speech (IDS) as catalyst for HRS. IDS is known to consist of higher pitch (Herold et al. 2011a). In order to discuss how HRS evolved from IDS, the hypothetical first generation of speakers to develop HRS will be dubbed proto-speakers. Likewise, the SS utterances used prior to the development of HRS will be proto-SS. For the first proto-speakers to develop HRS, proto-SS words were already sufficiently iconic in their segmental features alone, requiring no further prosodic distinction. In this way, their tonal composition was secondary to their segmental composition. Proto-SS tone did not referentially associate or contribute to meaning in the way that the segmental level did. However, because Sinitic languages are tonal, over time infants assigned a stable (or default) tone to this flexible SS prosodic tier. This process was gradual, happening as said proto-infants grew attuned to the phonological properties of their first language. Supposing that frequency has a hand in language acquisition, the assigned default tone would be one frequently associated with SS utterances. IDS is the vehicle that propelled this frequent association to SS words since IDS consists of higher pitch. High register assignment became regularized as the infant continues to grow and acquire language. Eventually this high register fossilises as more non-SS words enter the child's lexicon and SS words fall out of regular use. Due to the once tonally ambiguous nature of these proto-SS utterances, adult speakers would find no reason to correct the systematic application of high register to proto-SS words in proto-child speech. As the language was transmitted, HRS is fossilized and applied as a prosodically gestural tool used to facilitate the bootstrapping process.

IDS seems more likely the catalyst (as opposed to an iconicity-based trigger) behind this systematic and uniform patterning across three varieties of Sinitic given that Sinitic is tonal and IDS is integral to language acquisition (Herold et al. 2011a, 2011b). Following on from the IDS-driven hypothesis above, it looks as though prosody and IDS go hand in hand when it comes to bootstrapping and SS. As SS in language processing is an emerging field of research, it remains to be seen whether the prosody of IDS could influence conclusions drawn from experimental findings in the literature. Future studies should take into account IDS prosody when conducting SS tests. Finally, given that prosodic patterning is evidenced in Sinitic SS, future studies should also investigate whether or not such patterning exists in other languages, as such phenomena could influence perceptual responses in future experiment-driven papers.

Discrepancies between SS judgement results cited in recent studies (Forte et al. 2013; Ozturk et al. 2012) might be down how participants perceive the prosodic differences in these trials. How would judgements change if pitch were altered? The prosodic presentation of stimuli could sway participants' judgements as to whether or not stimuli are SS (Mitterer et al. 2012). With speakers of Sinitic, stimuli presented in high register seem likely to influence SS judgements. Future studies should investigate this as well as the potential for prosodic systematicity in other languages. To do this, the effect prosodic variation has on SS judgements in other tonal and non-tonal languages should be investigated, making sure to provide detailed descriptions of prosodic stimuli used. This will deepen the current understanding of iconicity, systematicity, and SS alike.

As other studies have shown, prosodic phenomena such as expressive voice (Herold et al. 2011b; Jesse & Johnson 2012; Reinisch et al. 2013) as well as IDS (Herold et al. 2011a) are used to facilitate language acquisition and comprehension. As SS has recently been shown to be an important feature of language acquisition (Imai et al. 2014; Asano et al. 2015; Lockwood et al. 2016), future studies should investigate the formative or acquisitional relationship between IDS, expressive voice, and SS, as the findings of the present study implicate. Not only will this advance the subfield of SS, but also open up new directions for research into language evolution, as IDS, expressive voice, and SS support an interface for language development.

4.4 Note on Orthographic Opacity

Studies indicate that prosodic phenomena, not accounted for by orthography, facilitate word learning (Nygaard et al. 2009; Reinisch et al. 2013; Herold et al. 2011a, 2011b). Other studies (Kelly 1992; Monaghan et al. 2005) show that prosody is used to systematically class grammatical categories. Thus, linguistic research cannot draw firm conclusions from orthography alone. Sinitic orthography, likewise, is opaque as the segmental and prosodic levels are not apparent. Moreover, certain Sinitic SS have no orthographic equivalent (Bauer 1985). The present study has worked around this obstacle through less opaque (Romanization) systems which reveal both segmental and prosodic information. The HRS phenomenon reported here is, however, not the end for systematicity in Sinitic SS. While Romanization systems (Yale, Pinyin, Peh-oe-ji) gloss high tone as orthographically identical to non-sound-symbolic high tones, it is possible that the SS instances exhibit marked differences in the natural speech stream. These differences could further contribute to acquisitional and bootstrapping process (e.g. HRS is louder or articulated at a higher frequency than non-sound-symbolic high tones). Subsequent investigation is needed to bypass opacity by referring directly to acoustic measurements of the speech stream.

5.0 Conclusion

The results of the present corpora analyses show that high register tone systematically marks a majority of SS words in Mandarin, Taiwan Southern Min, and Cantonese. This high register systematicity (HRS) is thought to facilitate language processing by prosodically alerting the listener to a word's SS status while simultaneously allowing iconic information to be conveyed at the segmental level. HRS thus divides a SS word into two tiers where the systematic information is supplied at the prosodic level (systematic tier) and the iconic or SS information at the segment level (iconic tier). Moreover, through the Sound Symbolism Bootstrapping Hypothesis (Imai and Kita 2014), HRS has been posited as a prosodically gestural tool which helps learners (children) make the referential connection between a SS word and its meaning. Infant Directed Speech (IDS) is posited here as the evolutionary catalyst for such a uniform patterning, calling into question the influence of IDS on SS development cross-linguistically.

SS research is a key point of departure in language cognition for several reasons: it is prevalent in the acquisition process, it facilitates word-learning, and is known to have cross-linguistic tendencies. Given its groundings in language development, empirical SS research has the potential to lead the

discussion of how language evolves. Further research into the iconic relationships between language and referents will shed light on how SS shapes fundamental language development. Moreover, by harnessing the systematic relationships present in SS, we should come closer to developing an understanding of the underlying mechanisms which drive communication towards arbitrariness.

Acknowledgements:

Many thanks to Dr Kirsty Rowan, Dr Ryszard Auksztulewicz, Emily Lake, and Yang Li.

References

- Akita, K. (2009). *A grammar of sound-symbolic words in Japanese: theoretical approaches to iconic and lexical properties of mimetics*. PhD Dissertation, Kobe University, Kobe, Japan.
- Asano, M., Imai, M., Sotaro, K., Kitajo, K., Okada, H., Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex*, 63, 196-205.
- Bauer, R.S. (1985). The expanding syllabary of Hong Kong Cantonese. *Cahiers de linguistique – Asie Oriental*, 14, 99-111.
- Bauer R.S., & Benedict, P. K. (1997). *Modern Cantonese Phonology*. Walter de Gruyter: Berlin. 49.
- Bergen, B.K. (2004). The psychological reality of phonaesthemes. *Language*, 80, 290-311.
- Bodomo, A. (2006). The structure of ideophones in African and Asian languages: the case of Dagaare and Cantonese. In *Selected Proceedings of 35th Annual Conference in African Linguistics, Harvard University, Cambridge, MA 2 - 4 April 2006* J Mungane, JP Hutchinson, & DA Worman (eds). 203-213.
- Bodomo, A. (2008). A Corpus of Cantonese Ideophones. Unpublished; Matthews, S. personal communication.
- Bremner, A.J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K.J., and Spence, C. (2013). 'Bouba' and 'kiki' in Namibia? A remote culture make similar shape-sound matches but different shape-taste matches to Westerners. *Cognition*, 126, 165-172.
- Chan, M. K. M. (1996). Sound Symbolism and the Chinese language. In T. F. Cheng, Y. Li & H. Zhang (eds.), *Proceedings of the 7th North American conference on Chinese linguistics and the 4th international conference on Chinese linguistics*, 17–34. Los Angeles: GSIL Publications, University of Southern California.
- Chu, T.-S. (儲泰松). (2012). Putonghua Nishengci de Yuyin Guilü jiqi Liwai 普通話擬聲詞的語音規律及其例外 [Phonological Rules of Onomatopoeias in Mandarin and Analysis of Its Exceptions]. *Journal of Anhui Normal University*, 40, 107-112.
- De Saussure, F. (1916). *Cours de Linguistique Generale*. Paris: Payot.
- Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. *Language and Linguistics Compass* 6, Oxford: Blackwell Publishing, 654-672.
- Dingemanse, M. (2013). Ideophones and gesture in everyday speech. *Gesture*, 13, 143-165.
- Dingemanse, M., Damián, B. E., Lupyan, G., Christiansen, M. H., Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences*, 10, 603-615.

- Elders, S. (2002). [Review of *Ideophones and Sound Symbolism in Atlantic Creoles*]. *Language in Society*, 31, 134–137.
- Ferinald, A., Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, 64, 637-656.
- Fernández-Prieto, I., Navarra, J., and Pons, F. (2015). How big is this sound? Crossmodal association between pitch and size in infants. *Infant Behavior and Development*, 38, 77-81.
- Forte, M. Weiss, A., Martin, A., Peperkamp, S. (2013). Looking for the bouba-kiki effect in prelexical infants. In *Proceedings of the 12th International Conference on Auditory-Visual Speech Processing, Annecy, France 29 August - 1 September 2013*. S Ouni, F Berthomrer, A Jesse (eds). Le Chesnay Cedex, France: Inria. 71-76.
- Gong, Liangyu (龔良玉). (1991). *Xiàngshēngcí cídiǎn 象聲詞詞典* [Onomatopoeia Dictionary]. Guizhou: Guizhou Education Publishing.
- Hamano, S. (1998). *The sound-symbolic system of Japanese*. Stanford, California: CSLI and Kurocio.
- Herold, D. S., Nygaard, L. C., Namy, L. L. (2011a). Say it like you mean it: mothers' use of prosody to convey word meaning. *Language and Speech*, 55, 423-436.
- Herold, D. S., Nygaard, L. C., Chicos, K. A., Namy, L. L. (2011b). The developing role of prosody in novel word interpretation. *Journal of Experimental Psychology*, 108, 229-241.
- Hinton, L., Nichols, J., and Ohala, J. J., eds., (1994). *SS*. Cambridge: Cambridge University Press.
- Hirata, S., Ukita, J., and Kita, S. (2011). Implicit phonetic symbolism in voicing of consonants and visual lightness using Garner's speeded classification task. *Percept Mot Skills*, 113, 929-940.
- Hung, H.-Y. (洪慧鈺). (2007). *Taiwan Minnanyu Nishengci Yanjiu 台灣閩南語擬聲詞研究* [Taiwan Southern Min Onomatopoeic Research]. Soochow University 私立東吳大學: Master's thesis.
- Imai, M. and Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B* 369: 20130298.
- Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., and Kita, S. (2015). Sound symbolism facilitates word learning in 14-month-olds. *PLoS One*, 10, e0116494
- Jesse, A. and Johnson, E. K. (2012). Prosodic temporal alignment of co-speech gestures to speech facilitates referent resolution. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 1567-1581.
- Kantartzis, K., Imai, M., and Kita, S. (2011). Japanese sound-symbolism facilitates word learning

- in English-speaking children. *Cognitive Science*, 35, 575-586.
- Kelly, M. H. (1992). Using sound to solve syntactic problems. *Psychological Review*, 99, 349-364.
- Kim, K.O. (1977). Sound symbolism in Korean. *Journal of Linguistics*, 13, 67-75.
- Kita, S. (2009). World-view of protolanguage speakers as inferred from semantics of sound symbolic words: a case of Japanese mimetics. N Masataka (ed). *Origins of Language*. Tokyo, Japan: Springer. 25-38.
- Kita, S., Kantartzis, K., Imai, M. (2010). Children learn sound symbolic words better: evolutionary vestige of sound symbolic protolanguage. *The Evolution of Language, Proceedings of 8th International Conference, Utrecht, the Netherlands 14 – 17 April 2010*. 206 – 213.
- Lai, Y.-P. (黎奕葆). (2015). Xianggang Yueyu Shuangyinjie Zhuangmao Houzhui de Yinyun Tese 香港粵語雙音節狀貌後綴的音韻特色 [Phonology of Disyllabic Ideophonic Suffixes in Hong Kong Cantonese]. *Language and Linguistics*, 16 (5), 691-729.
- Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. *First Language*, 34, 387-405.
- Leung, M.-T., Law, S.-P., and Fung, S.-Y. (2004). Type and token frequencies of phonological units in Hong Kong Cantonese. *Behavior Research Methods, Instruments, & Computers*, 36, 500-505.
- Li, Jing'er (李鏡兒). (2007). Xiàndài hànyǔ nǐshēngcí yánjiū 現代漢語擬聲詞研究 [Onomatopoeias in Modern Chinese]. Shanghai: Xuelin chubanshe.
- Lockwood, G., and Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound symbolism. *Frontiers in Psychology*, 6, 1246. doi: 10.3389/fpsyg.2015.01246
- Lockwood, G., Dingemanse, M., and Hagoort, P. (2016). Sound symbolism boosts novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication: <http://dx.doi.org/10.1037/xlm0000235>.
- Maurer, D., Pathman, T., Mondloch, and C. J. (2006). The shape of boubas: sound-shape correspondences in toddlers and adults. *Developmental Science*, 9, 316-322.
- Mitterer, H., Schuerman, W., Reinisch, E., Tufvesson, S., and Dingemanse, M. (2012). The limited power of sound symbolism. *Proceedings of the 16th Annual Conference on Architectures and Mechanisms for Language Processing*. Riva des Garda, Italy, 27.
- Monaghan, P., Charter, N., and Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, 96, 143-182.
- Monaghan, P., Christiansen, M. H., Chater, N. (2007). The phonological distributional coherence hypothesis: cross-linguistic evidence in language acquisition. *Cognitive Psychology*, 55, 259-305.

- Newman, J., and Sun, C. C. (2009). Taiwanese Southern Min Corpus 1.0. University of Alberta, Canada. DOI:10.7939/DVN/10485
- Nielsen, A., and Rendall, D. (2012). The source and magnitude of sound-symbolic biases in processing artificial word material and their implications for language learning and transmission. *Language and Cognition*, 4, 11-125.
- Nielsen, A., and Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology*, 67, 153-163.
- Nygaard, L. C., Herold, D. S., and Namy, L. L. (2009). The semantics of prosody: acoustic and perceptual evidence of prosodic correlates to word meaning. *Cognitive Science*, 33, 127-146.
- Oda, H. (2000). *An Embodied Semantic Mechanism for Mimetic Words in Japanese*. Bloomington: Indiana University.
- Ogura, T. (2006). How the use of 'non-adult words' varies as function of context and children's linguistic development. In *Studies in language science (5): papers from the 5th annual conference of the Japanese society for language science, Kobe, Japan, 5-6 July 2003*. M Nakayama, M Minami, H Morikawa, K Nakamura, & H Sirai (eds). Tokyo, Japan: Kurosio Publishers. 103-120.
- Ohtake, Y. and Haryu, E. (2013). Investigation of the process underpinning vowel-size correspondence. *Japanese Psychological Research*, 55, 390-399.
- Ozturk, O., Krehm, M., and Vouloumanos, A. (2012). Sound symbolism in infancy: evidence for sound-shape cross modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, 114, 173-186.
- Perniss, P., and Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B* 369, 20130300.
- Ramachadran, V. S., and Hubbard, E.M. (2001). Synaesthesia -a window into perception, thought and language. *Journal of Conscious Studies*, 8, 3-34.
- Reinisch, E., Jesse, A., and Nygaard, L. C. (2013). Tone of voice guides word learning in informative referential contexts. *The Quarterly Journal of Experimental Psychology*, 66, 1227-1240.
- Saji, N., Imai, M. (2013). Goshitoku ni okeru ruizousei no kouka no kentou [the role of iconicity in lexical development]. In *Onomatope kenkyu no shatei - chikadzuku oto to imi (sound symbolism and mimetics)*. K Shinohara & R Uno (eds). Tokyo, Japan: Hituji Syobo. 151-166.
- Suzuki, Y. (2013). Intarakushon no naka de tsukawareru 'onomatope + suru' doushi [the mimetic + light verb (suru) construction in the father - child interaction. In *Onomatope kenkyu no shatei - chikadzuku oto to imi (sound symbolism and mimetics)*. K Shinohara & R Uno (eds). Tokyo, Japan: Hituji Syobo. 167-181.
- Thiesen, C., Oberlander, J., and Kirby, S. (2010) Systematicity and arbitrariness in novel communication systems. *Interactional Studies*, 11, 14-32.
- Thompson, A. L. (2016). Who moo-ved my cow? The lexicalization of onomatopoeia and imitative shift in Mandarin. *Cambridge Occasional Papers in Linguistics*, 9, 246-253.
- Thompson, P. D., Estes, Z.E. (2011). Sound symbolic naming of novel objects is a graded function. *The Quarterly Journal of Experimental Psychology*, 64, 2392-2404.

van de Weijer, J. & Sloos, M. (2014). The four tones of Mandarin Chinese: representation and acquisition. Anita Auer & Björn Köhnlein (eds). *Linguistics in the Netherlands*, 31, 180-191.

Verhoef, T., Roberts, S. G., Dingemanse, M. (2015). Emergence of systematic iconicity: transmission, interaction and analogy. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Magio, P., et al. (eds). 2481-2486.

Wang, M. R. (王明仁). (1992). Guanyu Nishengci Wenti 關於擬聲詞問題 [Questions about Onomatopoeia]. Beijing Shifan Xueyuan Xuebao 北京師範學院學報 [Journal of Capital Normal University], 5, 93-97.

Wu, Soli (吳守禮). (2000). *Guo Tai dui zhao huo yong ci dian: Ci xing fen xi, xiang zhu Xia Zhang Quan yin* 國臺對照活用辭典 [Mandarin-Taiwanese Comparative Living Dictionary]. Taipei Shi: Yuan liu chu ban gong si. 2691

Yoshida, H. (2012). A cross-linguistic study of sound symbolism in children's verb learning. *Journal of Cognitive Development*, 13, 232-265.