International Journal of Applied Linguistics

WILEY

ORIGINAL ARTICLE

The Acquisition of Japanese Sound-Symbolic Words by Chinese JFL (Japanese as a Foreign Language) Learners

Yajing Feng[1] | Katsuo Tamaoka[2,3]

[1]School of International Studies, Hangzhou Normal University, Hangzhou, Zhejiang Province, China | [2]Professor at School of Foreign Languages, Shanghai University, Shanghai, China | [3]Graduate School of Humanities, Nagoya University, Nagoya-shi, Aichi-ken, Japan

Correspondence: Katsuo Tamaoka (ktamaoka@gc4.so-net.ne.jp)

Funding: This study was supported by funding from the National Social Science Fund of China (#22CYY051; PI: Yajing Feng at the School of International Studies, Hangzhou Normal University, China) and a Grant-in-Aid for the Japan Society for the Promotion of Science (#23K00628; PI: Emeritus Professor Katsuo Tamaoka at the Graduate School of Humanities, Nagoya University, Nagoya, Japan).

Keywords: acquisition patterns | Japanese as a Foreign Language (JFL) | phenomimes | phonomimes | sound-symbolic words

关键词: 习得模型 | 日语象声词 | 拟声词 | 拟态词 | 日语二语习得

ABSTRACT

This study examined how Chinese learners of Japanese as a Foreign Language (JFL) acquire sound-symbolic words, with a particular focus on phonomimes (mimetic sounds) and phenomimes (mimetic states). A multiple-choice test was developed using 12 sound-symbolic words that can function as both phonomimes and phenomimes, yielding 24 test items. In addition, a lexical knowledge test was administered. A total of 141 Chinese JFL learners participated. Structural equation modeling (SEM) was used to evaluate competing models of acquisition. While earlier studies on native Japanese children suggest a sequential acquisition from phonomimes to phenomimes, our SEM analysis supported a parallel acquisition model, in which the two types of sound-symbolic words were learned independently. This finding suggests that learners do not rely on semantic extension from phonomimes to phenomimes but instead treat each as a distinct lexical category.

摘要

本研究探讨了汉语母语者将日语作为外语学习时如何习得日语象声词, 尤其聚焦于拟声词 (模拟声音的象声词) 和拟态词 (模拟状态的象声词) 。开发了日语象声词词汇测试, 选用12个兼具拟声词和拟态词功能的日语象声词作为实验刺激词, 编制24道单项选择题, 并同步实施日语词汇知识测试。共计141名中国日语学习者参与实验。进一步采用结构方程模型 (SEM) 对不同习得模型进行评估。结果显示, 尽管以往针对日语为母语的儿童的研究结果表明日语母语者遵循从拟声词到拟态词的顺序习得模式, 但本研究的SEM分析结果支持平行习得模型, 即中国日语学习者在二语习得过程中将拟声词和拟态词作为独立范畴分别习得。这一发现表明, 二语习得者并不依赖从拟声词到拟态词的语义扩展推理, 而是将二者视为不同的词汇类别进行习得。

1 | Introduction

In many languages, sound-symbolic words hold a special place in conveying sensory experiences and emotions. In Japanese, these words are particularly significant, uniquely imitating sounds, actions, or states of being to vividly express meaning. They are integral not only to communication but also to the cultural and cognitive development of native speakers. Research on language acquisition has shown that Japanese infants exhibit an innate sensitivity to sound-symbolic words, which gradually evolves into a comprehensive understanding of their diverse meanings and usages (Akita 2009; Asano et al. 2015; Haryu

2010; Imai and Kita 2014; Imai et al. 2015). This progression reflects the interplay between sensory experiences and cognitive development. However, non-native learners of Japanese, particularly Chinese learners of Japanese as a Foreign Language (JFL), often face notable challenges in acquiring these words (e.g., Cao 2016; Feng and Tamaoka 2018a, 2018b; Haryu and Zhao 2007; Jin 1989). These difficulties raise broader questions about the universality of sound-symbolic word acquisition and whether sequential patterns observed in native speakers are unique to their linguistic and cultural context. Thus, the present study seeks to contribute to this discussion by examining the acquisition of Japanese sound-symbolic words among Chinese JFL learners. Specifically, it investigates how these learners develop an understanding of sound-symbolic words and whether their acquisition process parallels that of native Japanese speakers.

2 | Background Literature

2.1 | Acquisition of Sound-Symbolic Words by Native Japanese Speakers

Japanese sound-symbolic words are broadly categorized into three main types: sound mimetics or *phonomimes* (擬音語, *giongo*), manner mimetics or *phenomimes* (擬態語, *gitaigo*), and emotional mimetics or *psychomimes* (擬情語, *gijôgo*). Phonomimes involve direct sensory input and mimic sounds from the environment, making them more intuitive and easier to acquire. In contrast, phenomimes require a higher level of abstraction, representing qualities or states such as texture or motion. Psychomimes, on the other hand, pertain to internal emotional or psychological experiences, making them the most abstract and subjective.

Previous studies (Asano et al. 2015; Haryu 2010; Imai and Kita 2014; Imai et al. 2015) have shown that Japanese infants possess an innate sensitivity to sound-symbolic words. As they develop, they progress from acquiring simple sound mimetics (phonomimes) to manner mimetics (phenomimes) and eventually to emotional mimetics (psychomimes). In Japanese, the acquisition of sound-symbolic words reflects the interaction between children and their environment, as they gradually expand their understanding of meaning through sounds associated with their sensory experiences. Phonomimes, which directly mimic sounds from the environment, are typically acquired first due to their intuitive connection to real-world stimuli. As infants continue to interact with their surroundings, their ability to process more abstract qualities, such as texture or motion, facilitates the acquisition of phenomimes. Finally, psychomimes, which describe internal emotional or psychological states, are the most abstract and are acquired as children develop a more complex understanding of their emotions. As illustrated in Figure 1, this developmental progression depicts how infants' interactions with the physical world shape their cognitive processing of sound and meaning, starting with concrete auditory experiences and gradually expanding to more abstract concepts.

Unlike general vocabulary, sound-symbolic words are deeply rooted in the sensory experiences of native Japanese speakers from infancy. Akita (2009) analyzed data from infant speech records collected by Noji (1973-1977). Japanese children are reported to acquire sound-symbolic words in the following order: superexpressive > phonomimes > phenomimes > psychomimes > nonmimetics. This progression is referred to as the *lexical-iconicity hierarchy*, which is based on children's sensitivity to mimetic sounds. Exposure to different forms of sound-symbolic words in media, interactions with caregivers, and social contexts influences how children acquire these words. The lexical-iconicity hierarchy suggests that the cognitive processing involved in understanding sound-symbolic words mirrors the complexity and abstraction of the experiences they describe. As a result, sound-symbolic words are learned from concrete to abstract meaning; phonomimes are typically acquired first, followed by phenomimes, and eventually psychomimes.

In the prelinguistic stage, which serves as a preparatory phase for language acquisition, Japanese infants already begin to develop a sense of sound symbols as early as 11 months of age (Asano et al. 2015). This early sensitivity to sound-symbolic words gradually expands as they acquire a deeper understanding of sound symbols and begin to grasp the meanings of new general vocabulary (Imai et al. 2008; Imai and Kita 2014). As they grow older, Japanese-speaking infants typically become aware of the relationship between sound-symbolic words and the objects they refer to by around the age of four. At this stage, they begin to apply phonological rules, such as the principle that voiced sounds correspond to loud sounds made by larger or louder objects, while voiceless sounds correspond to quieter sounds produced by smaller or quieter objects (Tamori 2002; Haryu 2010). Japanese infants typically learn to interpret phonomimes based on phonological rules (Tamori 2002), where voiced sounds correspond to larger sounds produced by larger objects, and unvoiced sounds represent smaller sounds produced by smaller objects.

In line with this developmental progression, Osaka (1999) analyzed phonomimes and phenomimes from infant speech records collected by Maeda and Maeda (1983), showing that infants began to produce phonomimes by imitating physical sounds as early as 1 year and 1 month. By 1 year and 3 months, these sound-symbolic words appeared more consistently in their speech, with both phonomimes and phenomimes being produced. By around 3 years of age, phonomimes were slightly more frequently produced than phenomimes, reflecting the natural progression in the acquisition of sound-symbolic words.

To examine the order of acquisition in the lexical-iconicity hierarchy proposed by Akita (2009), Saji et al. (2011) conducted an experiment with Japanese children. In this study, novel verbs (e.g., the progressive form "*zabi-tteiru*" from the sound-symbolic word "*zabu*") were created based on videos representing the lexical conditions of phonomimes, phenomimes, and psychomimes. These novel verbs were randomly presented to the children, who were asked to select the corresponding video. The results indicated that the percentage of correct responses to phonomimes was unrelated to the children's age. However, the percentage of correct responses to phenomimes and psychomimes increased with age. Saji et al. (2011) concluded that phonomimes are understood from an early stage, while phenomimes and psychomimes are acquired as children grow older, supporting the sequential acquisition of these categories.

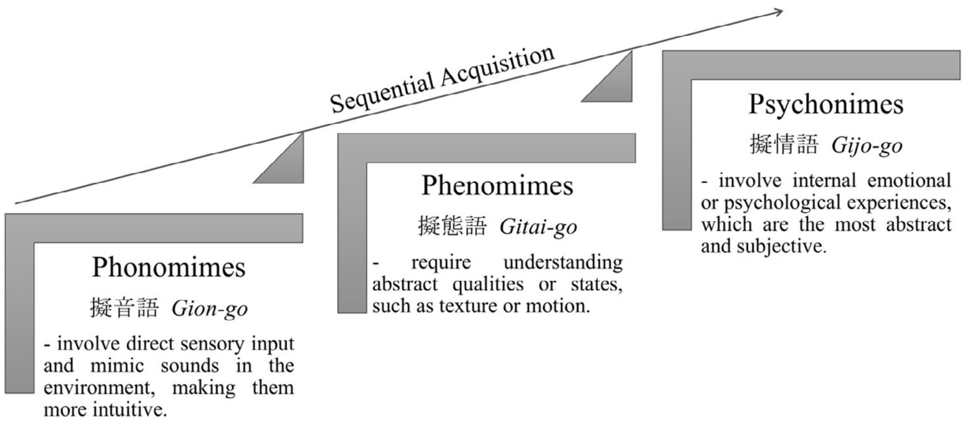

FIGURE 1 | Sequential acquisition of Japanese sound-symbolic words by native Japanese speakers.

For native Japanese speakers, the acquisition of sound-symbolic words follows a developmental route that begins with phonomimes and then expands semantically to include phenomimes and psychomimes. As these words are acquired, they become integrated into the broader framework of language development, eventually associating with syntactic structures and general vocabulary.

2.2 | Acquisition of Sound-Symbolic Words by Chinese JFL Learners

Previous research suggests that sensitivity to sound-symbolic words is not limited to native speakers of Japanese. Even individuals unfamiliar with Japanese sound-symbolic expressions often exhibit comparable intuitions. A growing body of cross-linguistic evidence supports the idea that certain types of sound symbolism reflect perceptual associations shared across cultures and languages.

One of the most well-known examples is the "Bouba–Kiki" effect (Köhler 1929; Ramachandran and Hubbard 2001), where people consistently match round shapes with softer-sounding words like *bouba* and angular shapes with sharper-sounding words like *kiki*. This effect has been replicated across diverse linguistic and cultural groups, including among non-literate populations such as the Himba of Namibia (Bremner et al. 2013; see also Davis 1961). Likewise, Sapir (1929) demonstrated that words like *mal* tend to be associated with larger objects, whereas *mil* is judged to refer to smaller ones. This finding echoed across multiple language backgrounds.

Building on this line of research, Wong et al. (2022) conducted controlled experiments using novel, genetically algorithm-generated Japanese-style sound-symbolic words that described surface textures. Participants from Japan, Singapore, and the United States rated these stimuli consistently: words containing consonants like /g/ and /k/ were reliably judged as representing harder and rougher textures. These patterns held regardless of participants' language backgrounds, confirming robust phonetic-iconic associations. The study demonstrated that sound-symbolic associations in Japanese are not solely the result of native language acquisition but rather reflect more universal mechanisms of cross-modal perception.

Furthermore, Wang (2011) investigated how both native and non-native speakers of Japanese perceive voiced and voiceless consonants in Japanese sound-symbolic words using the semantic differential (SD) method (Osgood et al. 1957). This method assessed conceptual responses across 10 semantic dimensions: "slow—fast," "big—small," "dark—bright," "strong—weak," "not pleasant—pleasant," "heavy—light," "dull—sharp," "hard—soft," "noisy—quiet," and "bad—good." Participants included 57 native Japanese speakers, 25 Indonesian JFL learners, and 92 Chinese JFL learners, the latter divided into two groups depending on their prior knowledge of the target sound-symbolic words (known vs. unknown). The results showed that native Japanese speakers perceived clear differences between voiced [g, d, b, z] and voiceless [k, t, p, s, h] consonants in eight of the ten semantic dimensions (excluding "dull—sharp" and "hard—soft"). Indonesian learners showed similar evaluation patterns in three dimensions: "big—small," "dark—bright," and "heavy—light." Chinese JFL learners who already knew the target words (known group) exhibited native-like perceptual judgments in six dimensions, including "heavy—light," "strong—weak," "not pleasant—pleasant," "noisy—quiet," "bad—good," and "dark—bright." Even the unknown group—those unfamiliar with the lexical meanings of the words—matched native speaker patterns in five of these dimensions.

Wang (2011) concluded that despite the absence of a voiced–voiceless contrast in the Chinese phonological system, Chinese JFL learners were nonetheless able to intuitively perceive the sound-symbolic meanings through sensory-based evaluation. This suggests that cross-modal correspondences in sound symbolism can be accessed even without formal instruction or lexical familiarity, reinforcing the idea that perceptual mechanisms underlying sound symbolism may operate independently of specific linguistic experience.

In contrast to studies supporting the perceptual universality of sound symbolism, other research suggests that JFL learners may not acquire Japanese sound-symbolic words solely through sensory perception as native speakers do. Haryu and Zhao (2007), for example, investigated whether the association between voiced sounds and large objects, and voiceless sounds and small objects, an iconic mapping pattern proposed by Tamori (2002), holds across different language backgrounds. The study involved three groups: (1) native Japanese university students, (2) Chinese university students with no prior experience studying Japanese, and (3) Chinese students majoring in Japanese (i.e., Chinese JFL learners). Participants were asked to associate voiced and voiceless sound-symbolic stimuli with pictures of objects differing in size under two conditions: existing real Japanese sound-symbolic words (*tonton* vs. *dondon*) and novel nonwords (*zonzon* vs. *sonson*).

The results showed that native Japanese speakers strongly adhered to the expected mapping pattern, selecting voiced–large and voiceless–small pairings at rates of 96.40% for real words and 92.90% for nonwords. In contrast, Chinese university students without Japanese language training performed only slightly above chance, at 57.10% for real words and 54.00% for nonwords, suggesting no strong inherent sensitivity to the sound-size mapping. Interestingly, Chinese JFL learners showed significantly above-chance performance: 71.40% for real words and 89.30% for nonwords, indicating a pattern more aligned with native speakers. These findings suggest that the ability to associate phonetic features with size-related meanings is not purely universal but rather is influenced by language-specific learning experiences. Unlike studies showing perceptual access to sound-symbolic associations in the absence of prior knowledge, this study emphasizes the role of explicit Japanese language learning in developing such associations. Hence, the voiced/voiceless–size mapping may be better understood as an acquired linguistic sensitivity rather than a universally intuitive phenomenon.

Furthermore, Feng and Tamaoka (2018a) investigated whether Chinese JFL learners acquire Japanese onomatopoeic (i.e., sound-symbolic) words through a process of semantic extension, specifically from concrete tactile meanings to more abstract, non-tactile meanings. The rationale was that tactile-related sound-symbolic expressions are grounded in direct sensory experience and may therefore be more universally accessible, even to non-native learners. If so, learners might initially acquire tactile onomatopoeias and then extend their understanding to non-tactile domains through a semantic expansion process. To test this, the study employed SEM to compare two acquisition patterns: one assuming continuous semantic expansion and the other assuming independent, parallel acquisition. The findings supported the parallel acquisition model, indicating that Chinese JFL learners tended to acquire tactile and non-tactile meanings of sound-symbolic words separately, rather than through a unified semantic continuum.

In a related study, Feng and Tamaoka (2018b) examined the roles of phonological similarity and lexical proficiency in the acquisition of two types of sound-symbolic words: phonomimes and phenomimes. Chinese JFL learners were categorized into three proficiency groups (lower, intermediate, and higher) based on their overall lexical knowledge. Using decision tree analysis, the study identified key factors influencing acquisition. The results revealed that phonological similarity significantly facilitated the acquisition of phonomimes, which are often imitative of physical sounds. In contrast, lexical proficiency played a more substantial role in the acquisition of phenomimes, which typically express more abstract or subjective states. These findings suggest that different cognitive and linguistic mechanisms underlie the learning of these two categories of sound-symbolic words.

Consequently, for JFL learners, particularly those whose first language lacks similar lexical categories, acquiring these forms would pose a significant challenge. In fact, various studies (e.g., Cao 2016; Feng and Tamaoka 2018a, 2018b; Jin 1989; Nakaishi et al. 2011; Peng 2007; Xu et al. 2010; Zhang 1989) have consistently reported that Japanese sound-symbolic words are particularly difficult for Chinese JFL learners. According to Tamamura (1989), these expressions are considered the third most difficult aspect of Japanese language acquisition for JFL learners, following the complexity of written Japanese and the use of *keigo* (honorific and polite expressions). Despite their frequency in daily Japanese communication, sound-symbolic words are largely underrepresented in teaching/learning materials. Surveys of beginner (Ochi 2005), intermediate (Moriyama 2006), and advanced (Mikami 2003) Japanese textbooks show that sound-symbolic vocabulary receives minimal systematic attention.

In Japanese language education, sound-symbolic words have not been adequately integrated into formal curricula or pedagogical frameworks (Sakaguchi 1995). As a result, Chinese JFL learners are often deprived of structured opportunities to acquire these forms and may fail to recognize their unique semantic, phonological, and pragmatic properties that distinguish them from general vocabulary. Jin (1989) further observes that in instructional contexts, these words are often treated in the same manner as ordinary vocabulary, taught primarily through pronunciation, morphological structure, and grammar, while neglecting the critical non-verbal dimensions such as imageability, emotion, and sensory associations.

In terms of learning outcomes, some studies (e.g., Rong 2011; Wang 2011) suggest that Chinese JFL learners can, to some extent, associate sound-symbolic forms with appropriate sensory meanings, demonstrating patterns somewhat similar to those of native speakers. However, a larger body of research (e.g., Cao 2016; Feng and Tamaoka 2018a, 2018b; Haryu and Zhao 2007; Jin 1989; Nakaishi et al. 2011; Peng 2007; Xu et al. 2010; Zhang 1989) highlights substantial learning difficulties. These studies argue that sound-symbolic expressions are typically acquired by Chinese JFL learners as isolated lexical items, lacking the integrated and intuitive grasp that native speakers develop through immersive exposure and use. Taken together, the present study seeks to address this gap by systematically examining how Chinese JFL learners acquire phonomimes and phenomimes.

3 | The Current Study

The present study investigates the acquisition of Japanese sound-symbolic words, which often have multiple meanings and usages. Some sound-symbolic words function as both phonomimes and phenomimes (Miyaji 1978), and these dual-use words are believed to be acquired sequentially, starting with phonomimes and followed by phenomimes (Akita 2009; Herlofsky 1998; Ishiguro 1993; Okubo 1967). Using 12 sound-symbolic words that function as both phonomimes and phenomimes (24 items total), the present study examines whether Chinese JFL learners acquire these words in a sequential manner, progressing from phonomimes to phenomimes.

4 | Method

4.1 | Participants

A power analysis conducted using the A-priori Sample Size Calculator for Structural Equation Models (Soper 2025; https://www.danielsoper.com/statcalc/calculator.aspx?id = 89) indicated that a minimum of 87 participants would be required to detect an effect size of 0.10 at a significance level of 0.01 with a desired statistical power of 0.80. A total of 141 participants were recruited for this study, exceeding the recommended minimum sample size (Westland 2010). All participants were native Chinese speakers majoring in Japanese at a university in China. The sample consisted of 127 females and 14 males. This gender distribution reflects the actual demographics of Japanese language programs at Chinese universities, where female students are significantly more numerous than male students, and is not the result of selective sampling. Participants included 60 second-year, 49 third-year, and 32 fourth-year students, ranging in age from 17 to 23 years (M = 20 years and 8 months, SD = 1 year). Informed consent was obtained from all participants, and the study was approved by the university's ethics committee in accordance with ethical research guidelines.

4.2 | Selection of Sound-Symbolic Words Used as Both a Phonomime and a Phenomime

The ABAB pattern is a characteristic morphological structure frequently found in Japanese sound-symbolic words (Amanuma 1974). In this study, we selected 12 ABAB-type words that can function as both phonomimes and phenomimes. These were used to construct 24 test items. Table 1 provides the hiragana forms, phonemic transcriptions, assumed equivalent expressions in Chinese, and other relevant details for each selected word.

As shown in Table 1, some selected Japanese sound-symbolic words possibly resemble familiar sound-symbolic expressions in Mandarin (e.g., *kotsu kotsu* and 咚咚 *kōng kōng*; *gan gan* and 咣咣 *guāng guāng*; *ban ban* and 梆梆 *bāng bāng*). However, these are only assumed equivalents and are neither systematically taught nor consistently recognized by Chinese JFL learners. Although prior studies (e.g., Feng and Tamaoka 2018a) have acknowledged the potential influence of phonological similarity, they also found that Japanese lexical knowledge, rather than surface-level form overlap, plays a more central role in sound-symbolic word acquisition. If learners were heavily relying on phonological resemblance to infer meanings, we would expect to see a pattern of transfer from phonomimes to phenomimes, supporting a sequential acquisition model (Model 1, details explained in the following section).

To determine the difficulty of each item, we referenced the 日本語教育語彙表 Ver. 1.0 (Japanese Language Education Vocabulary List, Ver. 1.0; Nihongo Gakushu Shien Gurupu, 2015, http://jhlee.sakura.ne.jp/JEV/), which includes approximately 18,000 vocabulary items commonly used in Japanese language instruction. This list also provides difficulty ratings assigned by five experienced Japanese language instructors. Based on these ratings, the 12 selected sound-symbolic words were classified into two categories: six intermediate-level words and six advanced-level words. These classifications are indicated in Table 1.

After selecting these 12 sound-symbolic words, we checked whether they were introduced in Japanese language textbooks used by the first- to fourth-year Chinese JFL university students who participated in this study. The eight textbooks include: みんなの日本語 1–2 (*Minna-no Nihongo 1–2*) [*Japanese for Everyone 1-2*], 综合日语 1–4 (*Zōnghé Rìyǔ 1–4*) [*Comprehensive Japanese 1-4*], 日语综合教程 5–8 (*Rìyǔ Zōnghé Jiàochéng 5–8*) [*Comprehensive Japanese Language Course 5-8*], 中級から学ぶ日本語 (*Tyûkyû-kara Manabu Nihongo*) [*Japanese from Intermediate Level*], 上級で学ぶ日本語 (*Zyôkyû-de Manabu Nihongo*) [*Japanese in the Advanced Level*], 日汉笔译教程 (*Rìhàn Bǐyì Jiàochéng*) [*The Japanese-Chinese Writing Translation Course*], 日本文学史 (*Rìběn Wénxuéshǐ*) [*History of Japanese Literature*], 日本近现代文学作品鉴赏 (*Rìběn Jìnxiàndài Wénxuézuòpǐn Jiànshǎng*) [*Appreciation of Modern and Contemporary Japanese Literature*]. These textbooks did not contain the selected 12 sound-symbolic words.

In addition, none of the selected sound-symbolic words appear in the 日本語能力試験出題基準 改定版 (*Japanese Language Proficiency Test Examination Criteria, Revised Edition*; Japan Foundation and Japan Educational Exchange and Services, 2007). This suggests that Chinese JFL learners are unlikely to have encountered these words through formal instruction, whether in university Japanese language courses or in preparation for the JLPT. Consequently, it is reasonable to assume that learners acquired these expressions primarily through informal exposure to Japanese media, such as manga, anime, dramas, and movies, via mass media and the internet outside of the classroom setting.

To explore how Chinese JFL learners acquire Japanese sound-symbolic words, we administered a post-test questionnaire to 141 participants, asking them to identify the sources through which they had encountered the target vocabulary (multiple responses allowed). The results revealed that 100 learners (70.92%) reported encountering these words through "manga and anime," while 97 learners (68.79%) cited "TV dramas and movies" as sources. In total, 125 students (88.65%) reported being exposed to the words via one or both of these forms of entertainment media. In contrast, only 60 participants (42.55%) indicated that they had learned these words through formal classroom instruction. Given that the tested sound-symbolic words are not included in standard Japanese language textbooks or in the Japanese Language Proficiency Test (JLPT), these findings suggest that Chinese JFL learners primarily acquire such vocabulary through informal, media-based exposure outside the classroom. This pattern underscores the important role of extracurricular learning environments, especially popular media such as manga and anime, in supporting the acquisition of sound-symbolic expressions.

To account for intra-group variability associated with Japanese language proficiency, we conducted analyses based on participants' academic year. As all participants had completed at least one year of formal Japanese instruction at the time of testing, they were grouped into second-year (n = 60), third-year (n = 49), and fourth-year (n = 32) cohorts. A one-way ANOVA was then performed to examine whether performance

TABLE 1 | List of 12 sound-symbolic words used for both a phonomime and a phenomime.

JEL Level of difficulties[a]	Sound-symbolic word	ABAB-type sound	Chinese[b] in hanzi	Chinese in pinyin	Inclusion of the textbook[c]	JLPT vocabulary list inclusion[d]	Frequency[e] (Newspaper)	Frequency[f] (BCCWJ)	Frequency[g] (CSJ)
Intermediate level	かちかち	/kati kati/	嘀嗒嘀嗒	dīdā dīdā	None	None	23	151	11
	こつこつ	/kotu kotu/	啌啌	kōng kōng	None	None	423	403	26
	ぱちぱち	/pati pati/	啪啪	pā pā	None	None	11	256	6
	びりびり	/biri biri/	刺啦刺啦	cīlā cīlā	None	None	20	117	0
	ばらばら	/bara bara/	啪嗒啪嗒	pādā pādā	None	None	1,064	1284	113
	つるつる	/turu turu/	哧溜哧溜	chīliū chīliū	None	None	161	336	10
Advanced level	がんがん	/gan gan/	咣咣	guāng guāng	None	None	89	452	41
	さらさら	/sara sara/	沙沙	shā shā	None	None	379	607	18
	ばんばん	/ban ban/	梆梆	bāng bāng	None	None	30	258	25
	ばたばた	/bata bata/	扑棱扑棱	pūlēng pūlēng	None	None	112	578	16
	ぺたぺた	/peta peta/	啪啪	pā pā	None	None	9	42	6
	がたがた	/gata gata/	轰隆轰隆	hōnglōng hōnglōng	None	None	35	382	30

Note: These sound-symbolic words are used for both phonomimes and phenomimes, so that actual questions were made up with 24 questions.

[a] Items taken from *Nihongo Kyōiku Goi-hyō, Ver 1.0* [JEL level] (http://jhlee.sakura.ne.jp/JEV/).

[b] These words are possible Chinese equivalent translations.

[c] Inclusion of the textbook used in the university where participants study.

[d] Inclusion of the former vocabulary list of the Japanese Language Proficiency Test.

[e] Frequency of Mainichi Newspapers from 1998 to 2015 (487,198,373 morphemes), https://tamaoka.org/websearch/index.html.

[f] Frequnecy taken from the Balanced Corpus of Contemporary Written Japanese (BCCW; https://clrd.ninjal.ac.jp/bccwj/en/).

[g] Frequency taken from the Corpus of Spontaneous Japanese (CSJ; https://clrd.ninjal.ac.jp/csj/en/).

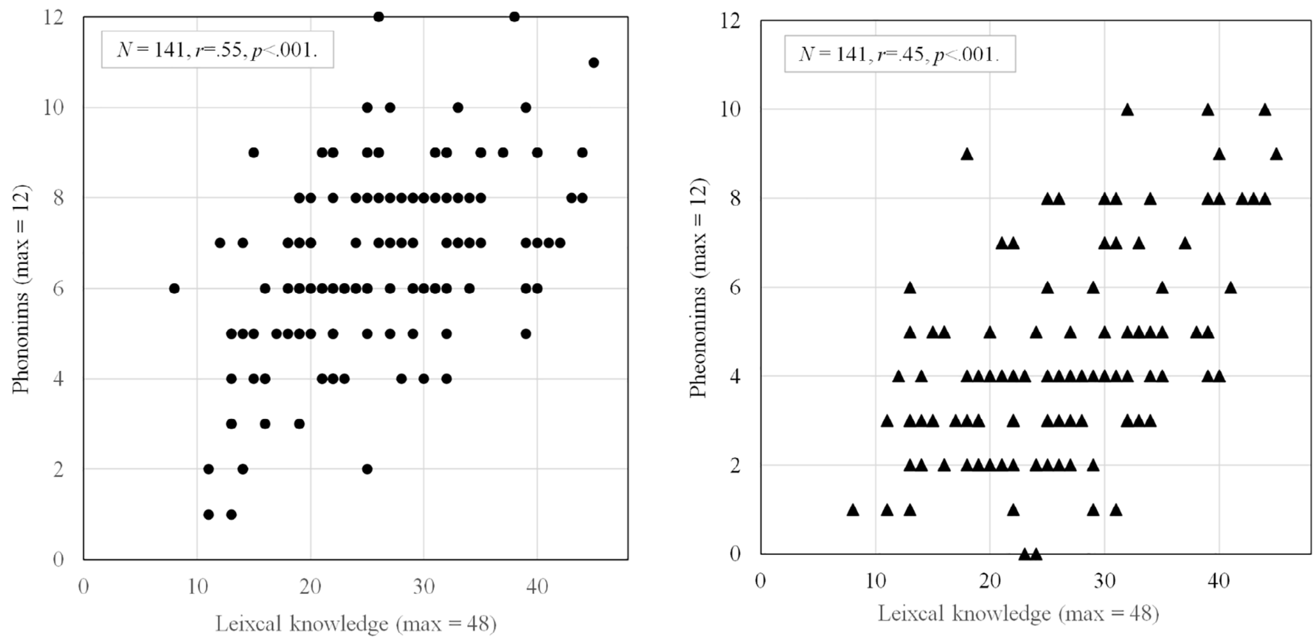

FIGURE 2 | Scatterplots showing the relationship between lexical knowledge and scores for phonomimes and phenomimes.

on the sound-symbolic word tasks differed across these groups. The results revealed significant differences in accuracy scores for both phonomimes and phenomimes. For phonomimes, a significant main effect of academic year was found [$F(2, 138) = 8.08$, $p < 0.001$, $\eta^2 = 0.11$], as was the case for phenomimes [$F(2, 138) = 8.37$, $p < 0.001$, $\eta^2 = 0.11$]. These findings suggest that learners' proficiency, as indexed by academic standing, had a meaningful impact on their comprehension of sound-symbolic words.

Recognition times are typically faster for frequently encountered words than for less familiar ones, a phenomenon known as the *word frequency effect* (for reviews, see Brysbaert et al. 2011, Brysbaert et al. 2018; Yonelinas 2002; for experimental studies, see Barry et al. 2001; Brysbaert et al. 2017; Coltheart et al. 2001; Cop et al. 2015; Hino and Lupker 1998; Monaghan et al. 2017; Monsell et al. 1989; Preston 1935; Taft 1979). This effect is one of the most robust predictors of processing efficiency in word recognition (Brysbaert et al. 2018). To examine the relationships among the frequencies in three corpora and learner performance, frequency data for the 12 sound-symbolic words used in this study were obtained from three major corpora: (1) the Mainichi Shimbun corpus, (2) the Balanced Corpus of Contemporary Written Japanese (BCCWJ), and (3) the Corpus of Spontaneous Japanese (CSJ). The frequency values across these corpora were highly intercorrelated: Mainichi and CSJ ($r(10) = 0.865$, $p < 0.001$), CSJ and BCCWJ ($r(10) = 0.894$, $p < 0.001$), and Mainichi and BCCWJ ($r(10) = 0.837$, $p < 0.001$), suggesting strong consistency in frequency distributions across written and spoken language registers. In addition, independent-samples t-tests showed no significant differences in word frequency between intermediate- and advanced-level sound-symbolic words in any of the three corpora: Mainichi, $t(10) = 0.981$, *ns*; CSJ, $t(10) = 0.192$, *ns*; BCCWJ, $t(10) = 0.272$, *ns*. These results confirm the reliability of using these corpora as frequency references in the present study.

4.3 | Measuring Knowledge of Basic Japanese Vocabulary

Understanding vocabulary in a language encompasses individual word meanings, spelling, connotations, and how words form phrases and sentences. Lexical knowledge is a critical component of language proficiency and communication. In this study, lexical knowledge was measured through four subtests based on grammatical categories developed by Miyaoka et al. (2011). This test consists of 12 questions for each of the three parts of speech—nouns, adjectives, and verbs—as well as 12 questions for functional words that play grammatical roles, totaling 48 vocabulary test questions. The questions are in a four-option single-choice format. All these words were selected from Levels 1 and 2 of the vocabulary list of the JLPT, cross-matched over the three word categories of nouns, adjectives, and verbs. Miyaoka et al. (2011) conducted a reliability study of this 48-question vocabulary test on 281 Chinese JFL learners at a Chinese university and found that Cronbach's alpha reliability coefficient was $\alpha = 0.74$. This test was administered to the 141 Chinese JFL university students in this study. The results showed that Cronbach's alpha reliability coefficient was very high, $\alpha = 0.88$. The vocabulary test used in this study was found to have internal consistency and a high reliability.

Additionally, we examined the relationship between overall Japanese lexical knowledge and sound-symbolic word comprehension. Lexical knowledge was assessed via a vocabulary test comprising four subtests (nouns, verbs, adjectives, and function words; maximum score = 48). The total lexical knowledge score showed significant positive correlations with both phonomime scores ($N = 141$, $r(139) = 0.55$, $p < 0.001$) and phenomime scores ($N = 141$, $r(139) = 0.54$, $p < 0.001$), indicating a developmental trend in which lexical proficiency supports the acquisition of sound-symbolic vocabulary. These relationships are illustrated in the scatterplots presented in Figure 2.

TABLE 2 | Test questions for 12 sound-symbolic words in phonomime and phenomime contexts.

Correct	Type	Question sentences	Three incorrect choices
かちかち /kati kati/	Phonomime	時計が ( ) 鳴る。	さらさら/はらはら/ばらばら
	Phenomime	水が ( ) に凍る。	どろどろ/ふわふわ/こりこり
こつこつ /kotu kotu/	Phonomime	窓ガラスを ( ) 叩く。	しくしく/つるつる/だらだら
	Phenomime	お金を ( ) 貯める。	ざらざら/だぶだぶ/からから
ぱちぱち /pati pati/	Phonomime	子供たちは ( ) 拍手する。	はらはら/からから/ぎりぎり
	Phenomime	目を ( ) させる。	びりびり/とんとん/はらはら
びりびり /biri biri/	Phonomime	用紙を ( ) 破る。	ぎりぎり/しくしく/ぐいぐい
	Phenomime	電気が ( ) 流れる。	ちょろちょろ/びしょびしょ/じろじろ
ばらばら /bara bara/	Phonomime	豆が ( ) 床に落ちた。	しくしく/さらさら/ぎりぎり
	Phenomime	牛が ( ) に解体された。	かちかち/どろどろ/ぴかぴか
つるつる /turu turu/	Phonomime	( ) 蕎麦を食う。	はらはら/ぐるぐる/ざくざく
	Phenomime	濡れた床は ( ) して滑りやすい。	ふわふわ/ざらざら/ごわごわ
がんがん /gan gan/	Phonomime	缶を ( ) 打ち鳴らす。	かさかさ/こそこそ/はらはら
	Phenomime	頭が ( ) 痛む。	ころころ/さらさら/かんかん
さらさら /sara sara/	Phonomime	風が木の葉を ( ) 揺らす。	しくしく/ぴかぴか/じりじり
	Phenomime	髪が ( ) している。	いらいら/ざらざら/ぎりぎり
ばんばん /ban ban/	Phonomime	机を ( ) 叩く。	びんびん/ざらざら/ぎりぎり
	Phenomime	焼肉を ( ) 注文する。	ぴかぴか/とんとん/ざらざら
ばたばた /bata bata/	Phonomime	鳥が羽を ( ) させる。	はらはら/しくしく/ぐるぐる
	Phenomime	仕事が忙しくて ( ) している。	ぴかぴか/がやがや/ざわざわ
ぺたぺた /peta peta/	Phonomime	顔を ( ) 叩く。	さらさら/ざらざら/こそこそ
	Phenomime	絵の具を ( ) 塗る。	くたくた/くらくら/ひたひた
がたがた /gata gata/	Phonomime	地震で家が ( ) 揺れる。	ぼんぼん/さらさら/ちょろちょろ
	Phenomime	怖くて ( ) 震える。	こんこん/おいおい/きょろきょろ

4.4 | Test Question Development for Sound-Symbolic Words

For each of the 12 sound-symbolic words selected for this study, test questions were created to examine their usage as phonomimes and phenomimes. This resulted in a total of 24 questions—two for each word, one for its phonomime usage and one for its phenomime usage. Similar to the vocabulary test by Miyaoka et al. (2011), the questions for sound-symbolic words were designed in a four-option single-choice format, where participants selected the most appropriate word from four choices. The question texts, correct answers, and distractor options for both phonomimes and phenomimes are shown in Table 2.

Taking the sound-symbolic word かちかち (*kati kati*), which is used as a phonomime meaning "tick-tock" in English, as an example, a sentence with a blank was presented: 時計が ( ) 鳴る。*Tokei-ga* ( ) *naru*, translating to "The clock is ticking." Chinese JFL learners were asked to select the correct answer from four choices, all of which are real sound-symbolic words used in Japanese: かちかち (*kati kati*), さらさら (*sara sara*), はらはら (*hara hara*), and ばらばら (*bara bara*). The correct answer that fits into this sentence is the phonomime かちかち. This word is also used as a phenomime. Another question was presented with the sentence: 水が ( ) に凍る。*Mizu-ga* ( ) *ni kôru*, meaning "The water freezes solid." This time, Chinese JFL learners had to choose from a different set of four sound-symbolic words: かちかち (*kati kati*), どろどろ (*doro doro*), ふわふわ (*huwa huwa*), and こりこり (*kori kori*). The correct choice that fits into this sentence is the phenomime usage of かちかち. In this way, two questions were created for each of the 12 sound-symbolic words, making up a total of 24 questions.

The 24 questions in the sound-symbolic word test were developed using authentic example sentences extracted from two authoritative Japanese dictionaries: *Digital Daijisen* (published by Shogakukan) and *Daijirin Third Edition* (published by Sanseido). To ensure rigorous categorical accuracy, we implemented a two-stage validation process. First, all stimulus items were carefully selected to represent prototypical examples of Japanese sound symbolism. Second, five native Japanese linguists independently classified each lexical item as either phonomime or phenomime, with inter-rater consistency reaching 100%. Furthermore, to verify whether the 24 sound-symbolic words appropriately fit into each sentence, we administered the test to native Japanese speakers. A total of 39 native Japanese university students (31 males and 8 females), with an average age of 18 years and 9 months (*SD* = 1 year and 5 months), completed all 24 questions. The mean percentage of correct responses was nearly 100% ($M = 99.17\%$). Therefore, these 24 test questions can be accurately judged by native Japanese speakers.

In addition, the difficulty levels of the words used in the question sentences were also checked using the web-accessible search engine 日本語読解学習支援システム:リーディングチュウ太 (*Japanese Language Reading Tutorial System: Reading Tutor*, Kawamura et al., 2021) (https://chuta.cegloc.tsukuba.ac.jp/). This program automatically assigns each word a level according to the JLPT, ranging from N1 (most difficult) to N5 (easiest). The only noun not included in the JLPT levels was 焼肉 (*yakiniku*), "grilled meat."

All of the adjectives and verbs used in the lexical knowledge test were selected from items included in the Japanese Language Proficiency Test (JLPT). Given that the participants were Chinese JFL learners majoring in Japanese at the university level, it is reasonable to assume that they were familiar with these words. The test combined 48 items from the lexical knowledge test developed by Miyaoka et al. (2011) and 24 items targeting sound-symbolic words, resulting in a single 72-item test. To minimize potential inter-item priming effects, particularly in cases where the same sound-symbolic word appeared in different roles as a phonomime and as a phenomime, pseudo-stratified randomization was applied to the full set of items. This procedure ensured that similar items were spaced apart and that repetition within the sequence was minimized. Although full counterbalancing was not implemented, this method effectively reduced carryover effects and sequence-related bias.

5 | Data Analysis and Results

5.1 | Comparison of Phonomime and Phenomime Acquisition

A total of 141 Chinese JFL learners were tested on 12 pairs of sound-symbolic words. To examine differences in acquisition between phonomime and phenomime usages, we conducted Chi-square tests of goodness-of-fit for each pair. These analyses compared the number of correct responses for phonomimes versus phenomimes, allowing us to assess whether learners exhibited a statistically significant preference or difficulty with either type. For example, in the case of ばんばん /ban ban/, 98 participants answered correctly for the phonomime usage, whereas only 17 did so for the phenomime usage—a substantial difference of 81 participants (98 vs. 17) that reached statistical significance [$\chi^2(1) = 96.34$, $p < 0.001$]. In contrast, for つるつる /turu turu/, 73 participants correctly identified the phonomime usage, while 66 correctly identified the phenomime usage. This smaller difference of 7 participants (73 vs. 66) was not statistically significant [$\chi^2(1) = 0.70$, *ns*], indicating no strong bias toward either usage in this case. Table 3 presents the comparative accuracy rates for all 12 phonomime–phenomime pairs, along with the corresponding Chi-square statistics.

Before conducting SEM analysis, we assessed whether the difficulty level of the sound-symbolic words (intermediate vs. advanced) significantly affected participant performance. A 2 (difficulty level: intermediate vs. advanced) × 2 (sound-symbolic word type: phonomime vs. phenomime) repeated-measures ANOVA was performed using percentage scores as the dependent variable ($N = 141$). The analysis revealed no significant main effect of difficulty level: intermediate words ($M = 46.34\%$, $SD = 24.37\%$) did not differ significantly from advanced words ($M = 43.45\%$, $SD = 24.92\%$) [$F(1, 280) = 1.84$, $p = 0.176$, $\eta^2 = 0.007$]. However, there was a significant main effect of symbolic word type, with phonomimes ($M = 54.91\%$, $SD = 24.37\%$) outperforming phenomimes ($M = 34.88\%$, $SD = 22.57\%$) [$F(1, 280) = 150.99$, $p < 0.001$, $\eta^2 = 0.350$]. The interaction between difficulty level and word type was not significant [$F(1, 280) = 2.21$, $p = 0.138$, $\eta^2 = 0.008$]. These results indicate that the difficulty level of the test items did not meaningfully influence learner performance. Consequently, difficulty level was not included as a factor in the subsequent SEM analysis. For the SEM, we used raw scores based on all 12 items for each symbolic word type (phonomimes and phenomimes), combining across both difficulty levels.

5.2 | Outline of SEM Analysis

There was a significant difference in test scores between phonomimes and phenomimes, indicating that Chinese JFL learners were more likely to acquire phonomimes than phenomimes. However, comparing test scores alone does not demonstrate a causal relationship. To verify the causal relation in the acquisition of phonomimes and phenomimes among Chinese JFL learners, the present study employed SEM (Byrne 2016; Hu and Bentler 1999; Kline 2015; Hair et al. 2010; Schumacker and Lomax 2016). The SEM method allows for the validation of hypothetical causal models based on collected data. In this analysis, the four parts of speech were treated as observed variables, while lexical knowledge constituted the latent variable. Since the aim of this study is to examine the causal relationship between phonomimes and phenomimes, these two variables were treated as observed variables in the SEM model.

5.3 | Three Causal Models for the Acquisitions of Phonomimes and Phenomimes

In this study, six observed variables were examined: four sub-categories of lexical knowledge—nouns, adjectives, verbs, and function words—and two sub-categories of sound-symbolic words—phonomimes and phenomimes. The latent variable of lexical knowledge was represented by four observed variables: nouns, adjectives, verbs, and function words. Using this foundational lexical knowledge, three hypothetical causal models were developed to examine the relationship between the acquisition of phonomimes and phenomimes.

Model 1 is grounded in previous research on native Japanese-speaking infants (Akita 2009; Herlofsky 1998; Ishiguro 1993; Okubo 1967), which suggests that sound-symbolic words are acquired in a developmental sequence from phonomimes to phenomimes via semantic extension. Similar patterns have been observed in studies of Chinese JFL learners (Haryu and Zhao 2007; Wang 2011), where phonomimes are reportedly acquired earlier than phenomimes. As illustrated in Figure 3, Model 1 assumes a sequential progression in which learners first acquire

TABLE 3 | Accuracy comparisons between phonomime and phenomime usages.

Sound-symbolic words	Sound	Number of participants		Mean of accuracies (%)		Chi-square test of independence
		Phonomime	phenomime	Phonomime	phenomime	
ばんばん	/ban ban/	98	17	69.50%	12.06%	$\chi2(1)=96.34$, $p<0.001$
がんがん	/gan gan/	122	51	86.52%	36.17%	$\chi2(1)=75.39$, $p<0.001$
ぱちぱち	/pati pati/	99	38	70.21%	26.95%	$\chi2(1)=52.52$, $p<0.001$
こつこつ	/kotu kotu/	86	46	60.99%	32.62%	$\chi2(1)=22.79$, $p<0.001$
かちかち	/kati kati/	76	38	53.90%	26.95%	$\chi2(1)=21.26$, $p<0.001$
ばらばら	/bara bara/	76	54	53.90%	38.30%	$\chi2(1)=6.91$, $p<0.01$
びりびり	/biri biri/	32	51	22.70%	36.17%	$\chi2(1)=6.16$, $p<0.05$
がたがた	/gata gata/	65	47	46.10%	33.33%	$\chi2(1)=4.80$, $p<0.05$
ぺたぺた	/peta peta/	83	65	58.87%	46.10%	$\chi2(1)=4.61$, $p<0.05$
さらさら	/sara sara/	76	65	53.90%	46.10%	$\chi2(1)=1.72$, *ns*
ばたばた	/bata bata/	43	52	30.50%	36.88%	$\chi2(1)=1.29$, *ns*
つるつる	/turu turu/	73	66	51.77%	46.81%	$\chi2(1)=0.70$, *ns*

Note: Number of participants means those who correctly answered questions.

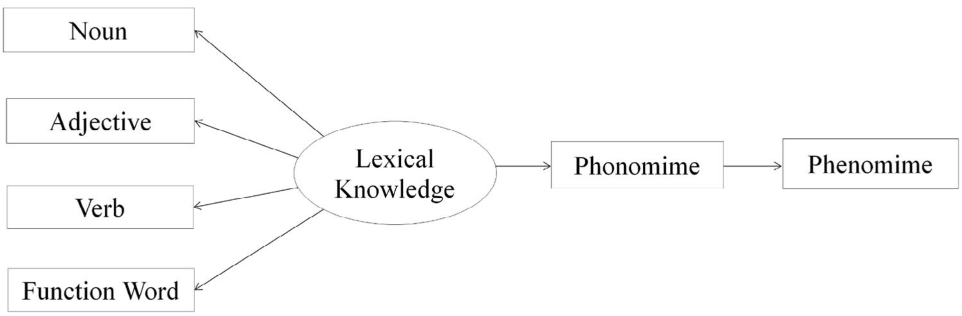

FIGURE 3 | Model 1: Sequential acquisition from phonomimes to phenomimes from lexical knowledge via phonomimes to phenomimes.

general Japanese lexical knowledge, which facilitates the learning of phonomimes. These phonomimes, in turn, serve as a foundation for acquiring phenomimes. This sequential model reflects the assumption that learners build abstract semantic understanding based on more concrete, perceptually grounded forms. Furthermore, as shown in Table 1, several of the Japanese sound-symbolic words in this study have possible phonological analogs in Mandarin. If learners were relying heavily on such cross-linguistic similarity to infer meaning, we would expect this to support the kind of transfer and analogical reasoning reflected in Model 1.

Iida et al. (2012) conducted a study testing Chinese JFL learners' ability to modify verbs with sound-symbolic words. Since the scores on the verb-modification test were strongly related to Japanese reading comprehension, they suggested that sound-symbolic words are primarily learned as vocabulary items rather than being acquired spontaneously. Based on this, and in contrast to previous studies on native Japanese infants (Akita 2009; Herlofsky 1998; Ishiguro 1993; Okubo 1967) and Chinese JFL learners (Haryu and Zhao 2007; Wang 2011), Model 2 assumes no causal relationship between the acquisition of phonomimes and phenomimes. As shown in Figure 4, Model 2 depicts a parallel causal model in which the observed variables of phonomimes and phenomimes are acquired independently, both based on the latent variable of lexical knowledge.

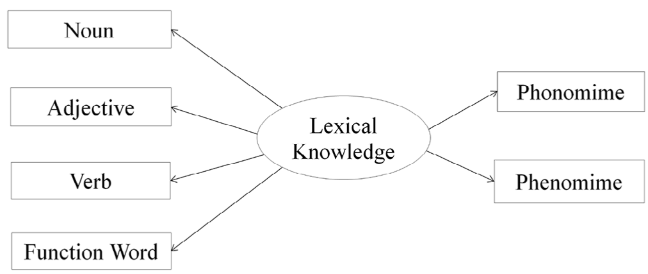

FIGURE 4 | Model 2: No causal relation between phonomimes and phenomimes.

As shown in Figure 5, Model 3 is a combination of Models 1 and 2. This model assumes that both phonomimes and phenomimes are independently acquired from basic lexical knowledge, while at the same time, phonomimes facilitate the acquisition of phenomimes. This is consistent with the idea of semantic expansion "from phonomimes to phenomimes" as observed in native Japanese infants (Akita 2009; Herlofsky 1998; Ishiguro 1993; Okubo 1967).

TABLE 4 | Correlations of six observed variables.

Latent variable	#	Observed variable	Variable # 1	2	3	4	5	6
Lexical knowledge	1	Noun						
	2	Adjective	0.64***	—				
	3	Verb	0.70***	0.66***	—			
	4	Function word	0.69***	0.56***	0.64***	—		
Symbolic words	5	Phonomime	0.50***	0.44***	0.52***	0.42***	—	
	6	Phenomime	0.57***	0.47**	0.45***	0.39***	0.39***	—
Mean			7.30	5.65	6.66	6.14	6.59	4.18
Standard deviation			2.28	2.20	2.62	2.71	2.16	2.15

Note: $N = 141$; $df = 139$.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

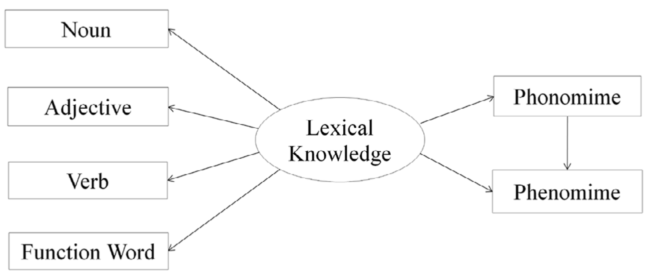

FIGURE 5 | Model 3: Combined causal model.

TABLE 5 | Comparison of goodness-of-fit indices for the three models.

Model	AIC	CAIC	BCC	BIC
Model 1	66.53	113.92	67.80	101.92
Model 2	35.54	82.92	36.80	70.92
Model 3	37.08	88.42	38.45	75.42
Saturated model	42.00	124.92	44.21	103.92
Independence model	421.82	445.51	422.45	439.51

5.4 | Correlations of Six Observed Variables

The Pearson's product-moment correlation coefficients, means, and standard deviations for the test scores of the six observed variables—nouns, adjectives, verbs, function words, phonomimes, and phenomimes—are presented in Table 4. The correlation coefficients for these six variables range from 0.39 to 0.69, all of which are statistically significant. The correlation between phonomimes and phenomimes was measured at $r(139) = 0.39$, $p < 0.001$, indicating a moderate correlation. The four variables representing lexical knowledge (nouns, adjectives, verbs, and function words) showed high correlations with each other. The strongest correlation was found between nouns and verbs ($r(139) = 0.70$, $p < 0.001$).

5.5 | Model Comparison and Fit Evaluation

To evaluate which structural equation model best fits the data, we compared three models using multiple fit indices. Four information criteria of Akaike Information Criterion (AIC) (Akaike 1987), Consistent Akaike Information Criterion (CAIC) (Bozdogan 1987), Browne-Cudeck Criterion (BCC) (Browne and Cudeck 1989), and Bayesian Information Criterion (BIC) (Schwarz 1978) were used for overall model comparison (see Table 5). For all indices, lower values indicate better fit. Model 2, which assumes parallel acquisition of phonomimes and phenomimes, consistently showed the best fit. Specifically, Model 2 had the lowest values across all four indices: AIC = 35.54, CAIC = 82.92, BCC = 36.80, and BIC = 70.92. These values were all lower than both the saturated model and the other two models (Model 1 and Model 3), indicating that Model 2 offered the most parsimonious and robust explanation of the data.

To further evaluate the absolute and incremental fit of Model 2, five additional indices were examined: chi-square (χ^2), Goodness of Fit Index (GFI), Adjusted GFI (AGFI), Comparative Fit Index (CFI) and Root Mean Square Error of Approximation (RMSEA). Model 2 demonstrated an excellent fit across all measures: $\chi^2(9) = 11.54$, *ns* (non-significant, indicating good fit), GFI = 0.97 and AGFI = 0.94 (both > 0.90, see Marsh and Grayson 1995; Schermelleh-Engel et al. 2003), CFI = 0.99 (excellent fit, see Schermelleh-Engel et al. 2003) and RMSEA = 0.045 (below the 0.05 threshold for good fit, see Browne and Cudeck 1993). Taken together, both model comparison and fit evaluation indices provide strong support for Model 2. These findings suggest that, contrary to a developmental progression from phonomimes to phenomimes, the acquisition of these two types of sound-symbolic words occurs independently among Chinese JFL learners.

5.6 | Causal Relationships Among Variables in Model 2

Next, we will examine the results of the analysis of causal relationships among the variables in Model 2 using SEM. The causal relationships among the variables in Model 2 are depicted in Figure 6. First, the factor analysis model of lexical knowledge

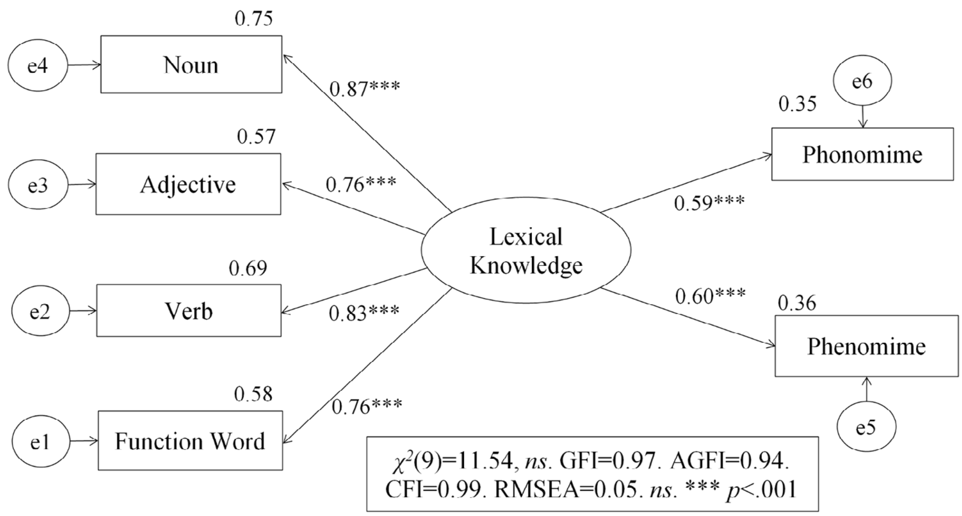

FIGURE 6 | Causal relations in the acquisition of sound-symbolic words using SEM. Note: $N = 141$. *** $p < 0.001$. The numbers on the arrows represent standardized partial regression coefficients (β).

revealed that the four observed variables of nouns, adjectives, verbs, and function words each contributed significantly, with values exceeding 0.76. Notably, nouns made the largest contribution to the latent variable of lexical knowledge, with a factor loading of 0.87. Verbs also contributed significantly, with a factor loading of 0.83. Model 2, which assumes that phonomimes and phenomimes are acquired independently, demonstrated strong causal relationships. The analysis showed a significant causal relationship from lexical knowledge to phonomimes ($\beta = 0.59$, $p < 0.001$) and from lexical knowledge to phenomimes ($\beta = 0.60$, $p < 0.001$).

6 | Discussion

This study examined the acquisition of Japanese sound-symbolic words by Chinese JFL learners, with a particular focus on the developmental sequence commonly observed among native Japanese-speaking children, namely, the progression from phonomimes to phenomimes. Specifically, we investigated whether Chinese JFL learners follow a similar trail, beginning with the acquisition of more concrete, sound-related phonomimes and subsequently acquiring more abstract, context-dependent phenomimes. The findings provide valuable insight into cross-linguistic differences in the acquisition of sound-symbolic vocabulary, which is closely tied to the sensory, linguistic, and cultural frameworks of native speakers.

As noted earlier in the introduction, certain sound–meaning associations appear to reflect universal perceptual tendencies. The "Bouba–Kiki" effect (Köhler 1929; Ramachandran and Hubbard 2001) shows that people across different cultures, including non-literate groups like the Himba of Namibia (Bremner et al. 2013), tend to associate round shapes with soft-sounding words like *bouba* and angular shapes with sharp-sounding words like *kiki*. Similarly, Sapir (1929) found consistent cross-linguistic associations between sounds like *mal* and *mil* and perceived object size. These universal associations provide a foundation for understanding how sound-symbolic words are acquired. In Japanese, children tend to acquire phonomimes first, followed by phenomimes and psychomimes (Akita 2009; Asano et al. 2015; Saji et al. 2011). This developmental sequence appears to align with the cognitive accessibility of different sound-symbolic types, starting with direct sensory experiences and moving toward more abstract representations. In this sense, the progression observed in Japanese-speaking children may reflect a broader, universally grounded pattern of acquisition, which may serve as a useful baseline when examining how second or foreign language learners acquire sound-symbolic forms in Japanese.

While the developmental pattern observed in Japanese children, acquiring phonomimes before phenomimes, may reflect a universally grounded sequence aligned with perceptual salience, the results of our study suggest that such a pattern does not hold for Chinese JFL learners. SEM analysis of Model 2 revealed that lexical knowledge had independent and equally strong effects on both phonomime and phenomime acquisition. This finding contradicts the notion of a universal, staged progression. Instead of acquiring sound-symbolic words through an innate, perceptually driven sequence, Chinese JFL learners appear to acquire both types of words through lexical development, with no evidence of a natural progression from phonomimes to phenomimes. In other words, their acquisition is shaped not by universal sensory pathways, but by the accumulation of lexical knowledge likely influenced by instructional exposure and language input. This dissociation from the native developmental trajectory challenges the assumption that iconicity alone ensures universal learnability of sound-symbolic forms in a second language context.

Additionally, as shown in Table 1, some Japanese sound-symbolic words may resemble familiar expressions in Mandarin. While Feng and Tamaoka (2018a) have acknowledged the potential role of phonological similarity, they also emphasize that Japanese lexical knowledge plays a more critical role than surface-level sound-form overlap in the acquisition of these words. If learners were primarily relying on phonological resemblance to infer meaning, we would expect a sequential acquisition pattern, first acquiring phonomimes and then extending that knowledge to phenomimes, consistent with Model 1. However, this model was not supported by our data. Instead, the structural equation modeling results favored a parallel acquisition model (Model 2), in which phonomimes and phenomimes are learned independently. This pattern suggests that learners approach each sound-symbolic word as a distinct lexical item, rather than making inferences based on cross-linguistic phonological similarity. Therefore, phonological inference is unlikely to be the primary mechanism guiding learner performance in this task.

The shift from a developmental to a parallel acquisition model highlights key cross-linguistic differences in how sound-symbolic words are learned. For native Japanese speakers, acquisition is grounded in early and repeated exposure to culturally embedded expressions through daily interactions and natural language use. In contrast, L2 learners, such as Chinese JFL students, typically lack a comparable system in their native language and have limited access to immersive cultural input. As a result, they tend to approach sound-symbolic words through analytical, vocabulary-oriented learning strategies. Consequently, phonomimes and phenomimes are not acquired in a sequential, developmental order but rather as separate lexical items. These findings underscore the importance of targeted pedagogical approaches that address the distinct cognitive and cultural challenges faced by L2 learners in acquiring Japanese sound-symbolic expressions.

7 | Limitations

This study offers valuable insights into the acquisition of Japanese sound-symbolic words by Chinese JFL learners, but several limitations should be noted. First, the study focused on a small set of 12 dual-use sound-symbolic words. Future research should include a broader range, especially psychomimes, which may present greater abstraction challenges. Second, the cross-sectional design and focus on learners at a specific proficiency level limit generalizability. A longitudinal approach would help clarify developmental changes over time, and examining learners across different levels and contexts would enhance applicability. Third, the use of a multiple-choice test captures only receptive knowledge. Future studies could incorporate experimental tasks (e.g., Wang 2011) to assess deeper processing and production abilities. These limitations point to the need for expanded lexical coverage, more varied learner profiles, and methodological diversity in future research.

8 | Conclusion

This study examined how Chinese JFL learners acquire Japanese sound-symbolic words, focusing on the distinction between phonomimes and phenomimes. While learners performed better on phonomimes than phenomimes, SEM revealed that the two types are acquired independently rather than in a sequential, developmental order. These findings contrast with the acquisition pattern observed in native Japanese speakers and highlight the influence of linguistic and cultural background on JFL learning. The results suggest that JFL learners approach sound-symbolic words as discrete lexical items, underscoring the need for targeted instructional strategies that account for their unique cognitive and cultural challenges.

Acknowledgments

We appreciate the kind support from the Chinese participants of the present study at Hangzhou Normal University, China.

Ethics Statement

The present study involving human participants was reviewed and approved by the Research Ethics Committee of Hangzhou Normal University, China.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in OSF at https://osf.io/jehn6/.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1111/ijal.12838.

References

Akaike, H. 1987. “Factor Analysis and AIC.” *Psychometrika* 52, no. 3: 317–332. https://doi.org/10.1007/BF02294359.

Akita, K. 2009. “A Grammar of Sound-Symbolic Words in Japanese: Theoretical Approaches to Iconic and Lexical Properties of Mimetics.” Doctoral diss., Kobe University.

Amanuma, Y. 1974. *Dictionary of phonomimes and phenomimes*. Tokyodo.

Asano, M., M. Imai, S. Kita, K. Kitajo, H. Okada, and G. Thierry. 2015. “Sound Symbolism Scaffolds Language Development in Preverbal Infants.” *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior* 63: 196–205. https://doi.org/10.1016/j.cortex.2014.08.025.

Barry, C., K. W. Hirsh, R. A. Johnston, and C. L. Williams. 2001. “Age of Acquisition, Word Frequency, and the Locus of Repetition Priming of Picture Naming.” *Journal of Memory and Language* 44, no. 3: 350–375. https://doi.org/10.1006/jmla.2000.2743.

Bozdogan, H. 1987. “Model Selection and Akaike’s Information Criterion (AIC): The General Theory and Its Analytical Extensions.” *Psychometrika* 52: 345–370. https://doi.org/10.1007/BF02294361.

Bremner, A. J., S. Caparos, J. Davidoff, J. de Fockert, K. J. Linnell, and C. Spence. 2013. ““Bouba” and “Kiki” in Namibia? A Remote Culture Makes Similar Shape-Sound Matches, but Different Shape–Taste Matches to Westerners.” *Cognition* 126, no. 2: 165–172. https://doi.org/10.1016/j.cognition.2012.09.007.

Browne, M. W., and R. Cudeck. 1989. “Single Sample Cross-Validation Indices for Covariance Structures.” *Multivariate Behavioral Research* 24, no. 4: 445–455. https://doi.org/10.1207/s15327906mbr2404_4.

Browne, M. W., and R. Cudeck. 1993. “Alternative Ways of Assessing Model Fit.” In *Testing Structural Equation Models*, edited by K. A. Bollen and J. S. Long, 136–162. Sage.

Brysbaert, M., M. Buchmeier, M. Conrad, A. M. Jacobs, J. Bölte, and A. Böhl. 2011. “The Word Frequency Effect: A Review of Recent Developments and Implications for the Choice of Frequency Estimates in German.” *Experimental Psychology* 58, no. 5: 412–424. https://doi.org/10.1027/1618-3169/a000123.

Brysbaert, M., E. Lagrou, and M. Stevens. 2017. “Visual Word Recognition in a Second Language: A Test of the Lexical Entrenchment Hypothesis With Lexical Decision Times.” *Bilingualism: Language and Cognition* 20, no. 3: 530–548. https://doi.org/10.1017/S1366728916000353.

Brysbaert, M., P. Mandera, and E. Keuleers. 2018. "The Word Frequency Effect in Word Processing: An Updated Review." *Current Directions in Psychological Science* 27, no. 1: 45–50. https://doi.org/10.1177/0963721417727521.

Byrne, B. M. 2016. *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*. 3rd ed. Routledge.

Cao, J. 2016. "Research on Mimetic Words in Japanese Language Education: Focusing on the Construction of Learning Content and Instructional Processes." Doctoral diss., Josai International University.

Coltheart, M., K. Rastle, C. Perry, R. Langdon, and J. Ziegler. 2001. "DRC: A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud." *Psychological Review* 108, no. 1: 204–256. https://doi.org/10.1037/0033-295X.108.1.204.

Cop, U., E. Keuleers, D. Drieghe, and W. Duyck. 2015. "Frequency Effects in Monolingual and Bilingual Natural Reading." *Psychonomic Bulletin & Review* 22: 1216–1234. https://doi.org/10.3758/s13423-015-0819-2.

Davis, R. 1961. "The Fitness of Names to Drawings: A Cross-Cultural Study in Tanganyika." *British Journal of Psychology* 52: 259–268. https://doi.org/10.1111/j.2044-8295.1961.tb00788.x.

Feng, Y., and K. Tamaoka. 2018a. "An Investigation on Continuity of Onomatopoeia Acquisition via Extending the Meaning of "Touch to Non-Tactile" by Native Chinese Speakers Learning Japanese." [In Japanese] *Japanese As a Second Language (JSL) Acquisition* 21: 7–24.

Feng, Y., and K. Tamaoka. 2018b. "Factors for Acquiring Onomatopoeic and Mimetic Words by Native Chinese Speakers Learning Japanese." [In Japanese] *Tyûgokugo Wasya-no Tame-no Nihongo Kyôiku Kenkyû* 9: 35–51.

Hair, J. F., W. C. Black, B. J. Babin, and R. E. Anderson. 2010. *Multivariate Data Analysis*. 7th ed. Prentice Hall.

Haryu, E. 2010. "Children's Understanding of the Symbolic Values of Japanese Onomatopoeia: Contribution of Knowledge About Orthography." [In Japanese] *Kyôiku Sinrigaku Kenkyû* 58, no. 3: 275–284. https://doi.org/10.5926/jjep.58.275.

Haryu, E., and L. Zhao. 2007. "Understanding the Symbolic Values of Japanese Onomatopoeia: Comparison of Japanese and Chinese Speakers." [In Japanese] *Sinrigaku Kenkyû* 78, no. 4: 424–432. https://doi.org/10.4992/jjpsy.78.424.

Herlofsky, W. J. 1998. "The Acquisition of Japanese Iconic Expressions: Giongo versus Gitaigo." *Journal of Japanese Linguistics and Education* 5, no. 1: 1–8.

Hino, Y., and S. J. Lupker. 1998. "The Effects of Word Frequency for Japanese Kana and Kanji Words in Naming and Lexical Decision: Can the Dual-Route Model Save the Lexical-Selection Account?" *Journal of Experimental Psychology Human Perception and Performance* 24, no. 5: 1431–1453.

Hu, L. T., and P. M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling* 6: 1–55. https://doi.org/10.1080/10705519909540118.

Iida, K., K. Tamaoka, and S. Hatsu. 2012. "Understanding of Sound-Symbolic Words by Native Chinese Speakers Learning Japanese." [In Japanese] *Journal of Japanese and Chinese Linguistics and Japanese Language Teaching* 5: 46–54.

Imai, M., and S. Kita. 2014. "The Sound Symbolism Bootstrapping Hypothesis for Language Acquisition and Language Evolution." *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, no. 1651: 20130298. https://doi.org/10.1098/rstb.2013.0298.

Imai, M., S. Kita, M. Nagumo, and H. Okada. 2008. "Sound Symbolism Facilitates Early Verb Learning." *Cognition* 109, no. 1: 54–65. https://doi.org/10.1016/j.cognition.2008.07.015.

Imai, M., M. Miyazaki, H. H. Yeung, et al. 2015. "Sound Symbolism Facilitates Word Learning in 14-Month-Olds." *PLoS ONE* 10, no. 2: e0116494. https://doi.org/10.1371/journal.pone.0116494.

Ishiguro, H. 1993. "The Origin of Mimetic Words." [In Japanese] *Gengo* 22, no. 6: 26–33.

Japan Foundation and Japan Educational Exchange and Services. 2007. *The Criteria of Japanese Proficiency Tests–Revised Edition*. Bonjinsha.

Jin, M. 1989. "On the Education of Japanese Phonomimes and Phenomimes in China." [In Japanese] *Journal of Japanese Language Education* 68: 83–98.

Kawamura, Y., T. Kitamura, and R. Hobara. 2021. "Japanese Language Reading Tutorial System: Reading Tutor." Accessed 12, August 2025. https://chuta.cegloc.tsukuba.ac.jp/.

Kline, R. B. 2015. *Principles and Practice of Structural Equation Modeling*. 4th ed. Guilford Press.

Köhler, W. 1929. *Gestalt Psychology*. Liveright.

Maeda, T., and K. Maeda. 1983. *Study on Vocabulary Development in Infants*. Musashino Shoin.

Marsh, H. W., and D. Grayson. 1995. "Latent Variable Models of Multitrait-Multimethod Data." In *Structural Equation Modeling: Concepts, Issues, and Applications*, edited by R. H. Hoyle, 177–198. Sage.

Mikami, K. 2003. "Mimetic Words in Advanced Teaching Materials: Analysis of Syntactic Features and Instructional Perspectives." [In Japanese] *Waseda Journal of Japanese Applied Linguistics* 2: 193–209.

Miyaji, H. 1978. "Morphological Considerations of Phonomimes and Phenomimes." [In Japanese] *Studies in the Japanese Language* 11: 33–39.

Miyaoka, Y., K. Tamaoka, and H. Sakai. 2011. "A Japanese Lexical Knowledge Test and Its Reliability: A Test Evaluation Using Data From Native Chinese Speakers Learning Japanese." [In Japanese] *Hiroshima Keizai Daigaku Kenkyû Ronsyû* 34, no. 1: 1–18.

Monaghan, P., Y. N. Chang, S. Welbourne, and M. Brysbaert. 2017. "Exploring the Relations Between Word Frequency, Language Exposure, and Bilingualism in a Computational Model of Reading." *Journal of Memory and Language* 93: 1–21. https://doi.org/10.1016/j.jml.2016.08.003.

Monsell, S., M. C. Doyle, and P. N. Haggard. 1989. "Effects of Frequency on Visual Word Recognition Tasks: Where Are They?" *Journal of Experimental Psychology: General* 118, no. 1: 43–71. https://doi.org/10.1037//0096-3445.118.1.43.

Moriyama, K. 2006. "Onomatopeoia in Intermediate Level Japanese Language Textbooks." [In Japanese] *Bulletin of the International Student Center, Nagasaki University* 14: 53–62.

Nakaishi, Y., N. Saji, M. Imai, and H. Sakai. 2011. "How do Advanced Chinese Learners of Japanese Use Onomatopoeic Vocabulary? A Production Experiment Using Animations." [In Japanese] *Tyûgokugo Wasya-no Tame-no Nihongo Kyôiku Kenkyû* 2: 42–58.

Nihongo Gakushu Shien Gurupu. 2015. "Nihongo kyôiku goihyô, Ver. 1.0." [In Japanese]. Accessed 12, August 2025. http://jhlee.sakura.ne.jp/JEV/.

Noji, J. 1973–1977. *The Reality of Language Life in Early Childhood (I~IV)*. Bunka Hyoron Shuppan.

Ochi, K. 2005. "Mimetic Words in Japanese Language Textbooks." [In Japanese] *Kuzuno* 9: 17–40.

Okubo, A. 1967. *Development of Infant Language*. Tokyodo shuppan.

Osaka, N. 1999. *Studying the Perceptual and Sentimental Words: Whereabouts of the Mind Explored in Mimetics*. Shinyosha.

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. 1957. *The Measurement of Meaning*. University of Illinois Press.

Peng, F. 2007. "Characteristics of Japanese Mimetic Words as Seen by Non-Native Speakers." [In Japanese] *Nihongogaku* 26, no. 7: 48–56.

Preston, K. A. 1935. "The Speed of Word Perception and Its Relation to Reading Ability." *Journal of General Psychology* 13, no. 1: 199–203. https://doi.org/10.1080/00221309.1935.9917878.

Ramachandran, V. S., and E. M. Hubbard. 2001. "Synaesthesia: A Window Into Perception, Thought and Language." *Journal Consciousness Studies* 8, no. 12: 3–34.

Rong, S. 2011. "Understanding Japanese Mimetic Words by Chinese Native Speakers." Master's diss., Kobe University.

Saji, N., A. Izawa, and M. Imai. 2011. "How do Children Understand the Meanings of Sound Symbolic Words Which are Diverse in the Degree Of Iconicity? A Case Study of Tripartite Classification in Japanese Mimetics." In *Papers from the National Conference of the Japanese Cognitive Linguistics Association*, 153–157. Japanese Cognitive Linguistics Association.

Sakaguchi, M. 1995. "The Representation of Onomatopoeias in Textbooks." [In Japanese] *Nara Kyôiku Daigaku Kokubun: Kenkyû-to Kyôiku* 18: 26–49.

Sapir, E. 1929. "A Study in Phonetic Symbolism." *Journal of Experimental Psychology* 12, no. 3: 225–239. https://doi.org/10.1037/h0070931.

Schermelleh-Engel, K., H. Moosbrugger, and H. Müller. 2003. "Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures." *Methods of Psychological Research* 8, no. 2: 23–74.

Schumacker, R. E., and R. G. Lomax. 2016. *A Beginner's Guide to Structural Equation Modeling*. 4th ed. Routledge.

Schwarz, G. 1978. "Estimating the Dimension of a Model." *Annals of Statistics* 6, no. 2: 461–464. https://doi.org/10.1214/aos/1176344136.

Soper, D. S. 2025. "A-Priori Sample Size Calculator for Structural Equation Models [Software]." Accessed 12, August 2025. https://www.danielsoper.com/statcalc.

Taft, M. 1979. "Recognition of Affixed Words and the Word Frequency Effect." [In Japanese] *Memory and Cognition* 7, no. 4: 263–272. https://doi.org/10.3758/BF03197599.

Tamamura, F. 1989. "Characteristics of Japanese Sound-Symbolic Words and Their Education." *Journal of Japanese Language Education* 68: 1–12.

Tamori, I. 2002. *Japanese Mimetic Words: Enjoying Phonomimes And Phenomimes*. Iwanami Shoten.

Wang, Y. 2011. "Sensory Evaluation of Japanese Mimetic Words by Different Learners of Japanese." [In Japanese] *Nihongo Kyôiku Kenkyû* 57: 65–82.

Westland, J. C. 2010. "Lower Bounds on Sample Size in Structural Equation Modeling." *Electronic Commerce Research and Applications* 9, no. 6: 476–487. https://doi.org/10.1016/j.elerap.2010.07.003.

Wong, L., J. Kwon, Z. Zheng, S. J. Styles, M. Sakamoto, and R. Kitada. 2022. "Japanese Sound-Symbolic Words for Representing the Hardness of an Object Are Judged Similarly by Japanese and English Speakers." *Frontiers in Psychology* 13: 830306. https://doi.org/10.3389/fpsyg.2022.830306.

Xu, Y., Y. Qiao, C. Wu, and J. Shi. 2010. *A Study of Japanese Phonomimes and Phenomimes*. Xuéyuàn Chūbǎnshè.

Yonelinas, A. P. 2002. "The Nature of Recollection and Familiarity: A Review of 30 Years of Research." *Journal of Memory and Language* 46, no. 3: 441–517. https://doi.org/10.1006/jmla.2002.2864.

Zhang, L. 1989. "Phonomimes and Phenomimes in Japanese as Seen by Chinese Learners of Japanese." [In Japanese] *Journal of Japanese Language Education* 68: 128–130.