Identifying a Phonetic Factors of Onomatopoeias Correlated to Sound Symbolic Commons between Japanese and Non-Japanese Speakers

Maki Miyamoto (mymt.makiron@jaist.ac.jp)
Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa, Japan

Shohei Hidaka (shhidaka@jaist.ac.jp)
Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa, Japan

Abstract

Although the relationship between the sound of each word and its meaning is generally arbitrary, onomatopoeias are said to have the unarbitrary link, which called sound-symbolism, between them. In this study, we investigated whether sound symbolic words are widely common in a class of onomatopoeias in some natural language. We conducted an experiment, which asked Japanese and non-Japanese speakers to match each given Japanese-onomatopoeia-like sound with a shape to which the sound referred. The result of analysis showed a similar structure for both speaker groups, in which round shapes were associated to a particular set of sounds and pointed shapes associated to the other set of sounds. Moreover, the round/ pointed shapes are correlated to pseudo onomatopoeias with sonorants/ fricative phonetic features. This finding supports the sound symbolic hypothesis asserting that the major component of Japanese onomatopoeias forms a bouba-kiki like sound-shape correspondence even for non-Japanese speakers.

Keywords: Arbitrariness of Language, Bouba-Kiki Effect, Sound Symbolism, Onomatopoeia

Introduction

A situation as we can see on communication to another language speakers, generally it is difficult for a speaker to communicate information just with speech sounds to a listener who is not a same language speaker as the speaker.

This cross-linguistic communicative barriers may be due to the arbitrariness of language (Saussure, 1916). This refers no necessity of fixed corresponding relationship between the speech sound and what it refers. On the other words, correspondence between the speech sound and what it refers is defined into one in certain language in most cases, however, it can correspond to another meaning or no correspondence against the sound in another language. Thus, this arbitrariness of language makes verbal communication with speech sounds difficult between the speaker and the listener who do not share the same language.

However, a specific class of empirical findings contradicting against this theoretical observation has been reported by Köhler (1929) and Ramachandran & Hubbard (2001), called "bouba-kiki" effect. In their experiment, the participants were asked to guess and match a novel speech sound to a unfamiliar shape to which the sound was expected to refer. At a result, the participants matched the specific speech sound to the specific shape at rates above chance, though participants had no previous knowledge on the relationship between the speech sounds and shapes. This showed that correspondence between speech sounds and meaning to which they referred may not be completely arbitrary. The potential factor related to the bouba-kiki effect is partly identified. For example, McCormick, Kim, List, & Nygaard (2015) suggested that phonetic properties of pseudo words affects the associated shapes in the bouba-kiki experiment.

The bouba-kiki effect is supposed one of examples of *sound symbolism*. The sound symbolism hypothesis is that there is some class of words in natural language which holds a systematic correspondence between speech sounds and referential meanings, and such class of words may support the first language learners at early developmental stage (Imai et al., 2015).

Generally, typical words which familiarly used have arbitrary correspondence between its speech sound and the meaning which the sound refers. However, onomatopoeia, which imitates animals crying or nature sounds, or represents objects' looking or someone's feeling as language sounds, is thought that correspondence between the sound and the meaning which the sound refers are supposed not arbitrary. Thus, onomatopoeia is supposed to be the word which has sound-symbolic correspondence between the sound and the meaning (Tamori & Schourup, 1999; Hamano, 2014). In their research, it is showed that each sound which construct onomatopoeia makes people associate a specific sound-symbolic impression. However, their studies considered only a limited number of onomatopoeias, and thus it has remained unclear whether the whole class of onomatopoeia is systematically sound-symbolic.

In this research, we defined sound symbolism as a property of words which are available for both the speaker and non-speakers of a particular language and systematically maps the sound to the referential meaning. We then hypothesize that Japanese onomatopoeia as a whole, not only a few special words, forms such sound symbolic structure. If so, we further aim to identify the phonetic factors correlated to the sound symbolic structure.

Experiment

On communication with speech sounds, a speaker imagines what he/she wants to communicate and a listener knows what the speaker wants to communicate from speaker's speech sounds. However, it is normally hard to observe objectively whether what the speaker wants to communicate and what the

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

listener knew from speech sounds matching or not before and after communication because they are in a head of both. In this research, we visualized the process of communicating information by setting choices in advance and making the speaker and the listener choose one from them. At this point, we defined success of communicating information as matching choices or choosing similar choices by both.

In our experiment, we set the speaker as a Japanese speaker, the listener as a non-Japanese speaker, and the speaker's speech sound as a sound of Japanese onomatopoeia-like words. This is because Japanese have many onomatopoeias in the vocabulary and onomatopoeias are familiar for most Japanese. We made both Japanese speakers and non-Japanese speakers listen to audio of Japanese onomatopoeia-like sound and choose one of shapes which were guessed the most matching shape for audio. We prepared more choices against one speech sound to investigate how many ranges of similar shape could be covered by one speech sound though two choices were shown in previous research by Ramachandran & Hubbard (2001).

Method

Participants

We recruited 12 Japanese speakers and 27 non-Japanese speakers as the participants to the experiment. All the Japanese speakers were native speakers of Japanese. All the non-Japanese speakers were foreign students living in Japan, who come from Asian countries such as Thai, Myanmar, China, Indonesia, Bangladesh, Kenya, Vietnam, India, Laos, and Malaysia. A half of them hardly had an experience of learning Japanese (less than a year).

Materials

Onomatopoeia-like Words In this study, we employed only a simple type of Japanese sound combinations as onomatopoeia-like pseudo words. According to Tamori (1998), Japanese onomatopoeias can be categorized by phonetic and morphological properties of the word and the most typical form is a repeat of two syllables which are constructed by a consonant and a vowel in this order (CVCV-CVCV) (e.g. KIRAKIRA, DOSUDOSU). Following this categorization, we employed the repeating-style pseudo words for the experiment. we modified the experimental protocol to choose a set of onomatopoeia proposed in previous research (Shimizu et al., 2014), and used to choose the set of onomatopoeia. First, we created 20449 onomatopoeia-like words with the repeat style by all the possible combinations of Japanese moras. It was constructed with all phonemes of Japanese including particular phonemes (prolonged sound /R/, choked sound /Q/, syllabic nasal /N/). Next, we let three native speakers of Japanese judge whether the word was acceptable as a Japanese word or not. Then, we finally obtained the 249 onomatopoeia-like words, which two or more participants accepted as Japanese word.

Sound Stimuli We created two sets of sound stimuli for each of the 249 words, by human voice and synthetic voice. The human-voice stimuli were made by a female native speaker of Japanese, reading it aloud. The synthetic voice was artificially composed like a female voice to control the potential effects of tones and pitches of reading. There are two set of sound stimuli, one consists of 125 human-voice stimuli and 124 synthetic voice ones, and the other consist of 124 human-voice ones and 125 synthetic ones. In each set, each pseudo word is assigned either human-voice or synthetic voice randomly with the probability approximately 1/2 to counterbalance.

List of Shapes We used the list of shapes obtained by Yamaguchi & Shiina (2005) for the experiment. The list included 36 shapes drawn by a group of participants to their experiment. We made the two lists of shapes, the list of 36 shapes for Japanese speakers (Figure 1), and that of 20 shapes for non-Japanese speakers (Figure 2). The reason why the number of shapes were reduced for non-Japanese speakers is to reduce mental stress of non-Japanese speakers, who were more stressful than Japanese speakers to keep listening to the sounds of non-native language. The shapes were located in the visual array so that similar shapes are close to each other, in order to increase participants' accessibility to the shapes.

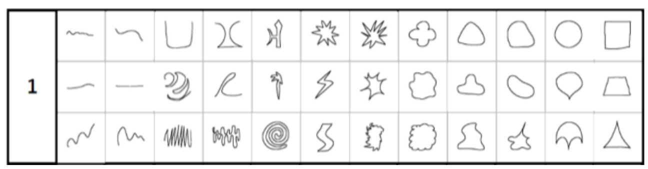

Figure 1: The list of shapes for Japanese speakers.

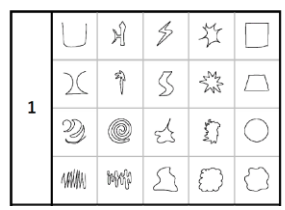

Figure 2: The list of shapes for non-Japanese speakers.

Procedure

Each participant was required to listen to the 249 sounds of onomatopoeia-like words one by one, then to guess and choose one shape out of the given shape list, to which the sound referred. Each sound was played by participant's clicking of the button displayed on a computer screen and they listened to it via a headphone. They were allowed to listen the same sound more than one time. The experimenter instructed to participants (either in Japanese to Japanese speakers or English to non-Japanese ones) "Put on a headphone. Then click an audio player on the screen and listen to a sound. Choose a shape, which is closest to the image you associate with the sound, from a printed list of shapes, and circle one".

Results

Entropy as a Measure of Choice Consistency

We analyzed whether the choice of shapes for each sound by Japanese speakers and non-Japanese speakers had bias or not. In this research, we adopted Shannon Entropy as a measure of consistency of their choices within the participant group. The lower normalized entropy, ranges from the minimum 0 to the maximum 1, for the pseudo words suggests more consistent choices within the group – more participants choses the same shape.

Japanese Speakers Table 1 shows the normalized entropy and correspondence between sounds of each word and shapes by Japanese speakers in ascending order of normalized entropy up to the 21st. The sound with the lowest normalized entropy was 0.000 was 'guruguru', which means all the 12 participants chose the same shape. The other words which normalized entropy was lower were found the common onomatopieas, which Japanese speakers often use.

Table 1: Normalized entropy and the most chosen shape for each of the top 21 pseudo words (Japanese speakers). *: The shape which less than 2 participants chose.

Rank	Word	Normalized Entropy	Most chosen shape (Rate)	
1	guruguru	0.000	(1.000)	-
2	moamoa	0.256	(0.667)	(0.333)
3	kunekune	0.290	(0.750)	(0.167)
3	mofumofu	0.290	(0.750)	(0.167)
5	kachikachi	0.332	(0.667)	(0.250)
6	zaazaa	0.337	(0.750)	*
7	korokoro	0.349	(0.667)	(0.167)
8	togetoge	0.386	(0.583)	(0.250)
9	busubusu	0.396	(0.667)	(0.167)
10	kakukaku	0.454	(0.500)	(0.333)
11	nebaneba	0.460	(0.417)	*
12	gorogoro	0.500	(0.500)	(0.167)
13	goagoa	0.529	(0.500)	(0.250)
14	gaogao	0.543	(0.583)	*
14	kushakusha	0.543	(0.583)	*
14	nazonazo	0.543	(0.583)	*
14	jitojito	0.543	(0.583)	*
14	shimashima	0.543	(0.583)	*
19	zukizuki	0.544	(0.417)	(0.333)
20	gotsugotsu	0.546	(0.500)	(0.167)
20	piripiri	0.546	(0.500)	(0.167)

Non-Japanese Speakers Table 2 shows the normalized entropy and correspondence between sounds of each word and shapes by non-Japanese speakers in ascending order of normalized entropy up to the 20th. The sound which normalized entropy was the lowest was 'ouou'. This result showed a different tendency from Japanese speakers.

Table 2: Normalized entropy and the most chosen shape for each of the top 20 pseudo words (non-Japanese speakers). *: The shape which less than 3 participants chose.

Rank	Word	Normalized Entropy	Most chosen shape (Rate)	
1	ouou	0.618	(0.481)	(0.111)
2	minimini	0.672	(0.296)	(0.148)
2	jirojiro	0.672	(0.370)	(0.111)
4	ishiishi	0.680	(0.259)	(0.222)
5	suesue	0.732	(0.296)	(0.148)
6	woiwoi	0.746	(0.333)	(0.185)
7	chirichiri	0.748	(0.222)	(0.185)

8	oraora	0.752	(0.296)	(0.148)
8	nekoneko	0.752	(0.296)	(0.111)
10	oreore	0.754	(0.333)	(0.111)
11	kechikechi	0.762	(0.222)	(0.111)
11	korokoro	0.762	(0.333)	(0.148)
13	kusekuse	0.767	(0.185)	(0.148)
14	gotogoto	0.768	(0.259)	(0.111)
15	akiaki	0.769	(0.222)	(0.148)
16	mofumofu	0.782	(0.333)	*
17	bakobako	0.787	(0.259)	(0.185)
18	katsukatsu	0.788	(0.185)	(0.148)
19	kya-kya-	0.796	(0.185)	(0.148)
20	ikiiki	0.797	(0.222)	(0.148)

Structural similarity

Correspondence Analysis The entropy analysis showed some different tendency between the Japanese-speakers and non-Japanese speakers. Next, we consider *structural correspondence* between sounds and shapes for the two groups, which may underlie the bouba-kiki effect potentially. Such structural correspondence is considered some second order relationship between sound-sound similarity and shape-shape one. To quantify such correlation, we performed correspondence analysis for each of the participant groups. Correspondence analysis assigns a real value to each of sound and shape, which is originally nominal-scale variable, to maximizes the correlation between real-valued sounds and real-valued shapes for a given cross table of the choice of shapes for the sounds. Correspondence analysis for Japanese speakers, the largest correlation was $\boldsymbol{r}$ =0.674, and for non-Japanese speakers, the largest correlation was weakly positive ($\boldsymbol{r}$ =0.355).

We sorted the order of words and shapes of the cross table in the order of their calculated scores, and visualized it. Figure 3 and Figure 4 shows the score-sorted cross table visualized as a heatmap. The sign of scores of non-Japanese speakers was changed for analysis (the sign of scores can be exchangeable arbitrarily without loss of generality). In these heatmaps, the vertical axis shows the sorted shapes, and the horizontal axis shows sounds. In both Japanese (Figure 3) and non-Japanese speakers (Figure 4), the choice frequency concentration (red cells in the heatmap) distributed from bottom-left to top-right shows the correspondence trend between the sounds and shapes, although non-Japanese speaker's pattern seemed more distributed and thus blurred.

In both groups, sharper (rounder) shapes were at the left (right) hand side, and "shaper" ("rounder") sounds were at the top (bottom) in each axis. Thus, this correspondence analysis suggests that the shape-to-sound choices made by both Japanese and non-Japanese speakers commonly have a similar general trend consistent with the bouba-kiki effect.

Multiple Regression Analysis on Phonetic Features In order to further analyze the phonetic structure of the pseudo onomatopoeia in depth, we performed multiple regression analysis on the sound scores of the largest correlation of the corresponding analysis.

We identified 33 phonetic units, in which some combination of them forms each of the 249 pseudo words. First, we conducted multiple regression analysis of the scores of sounds of Japanese speakers as the dependent variable predicted by the existence of the 33 phonetic units as the independent variables ($R^2 = 0.412$, $p < 0.05$). Then, we conducted stepwise regression analysis to select a subset of independent variables which keep the accuracy of a regression model only with significant variables (R^2 =0.390, p <0.05). Table 3 shows the result of this stepwise regression analysis of Japanese speakers. Similarly, we analyzed the data of non-Japanese speakers with multiple regression analysis (R^2 = 0.5622, p <0.05) and selected a subset of independent variables by stepwise regression analysis ($R^2 =$ 0.551, $p < 0.05$). Table 4 shows the result of Stepwise regression analysis of non-Japanese speakers. In Table 3 and Table 4, gray-shaded phonetic units were the one which were selected in common for both Japanese speakers and non-Japanese speakers. Through this regression analysis, the sign of estimated coefficients was seemed to correlate to the property of shapes: the positive regression coefficients corresponded to more rounded shapes; the negative ones corresponded to shaper/more pointed shapes. In addition, we analyzed correlation between the regression coefficients of Japanese speakers and non-Japanese speakers. The result of correlation analysis was significantly positive ($r = 0.825$, $p < 0.05$). This result confirms the similarity between the two groups of participants in their structural similarity of sounds and shapes. As the result of the stepwise multiple regression analysis of the scores of the sounds, we found that the vowels [o], [u], and [i] and the fricative consonants [d], [ɡ], [k], [z], [ts], [tɕ], [ʑ^{j}], [ɡ^{j}], and [k^{j}] are common significant predictors.

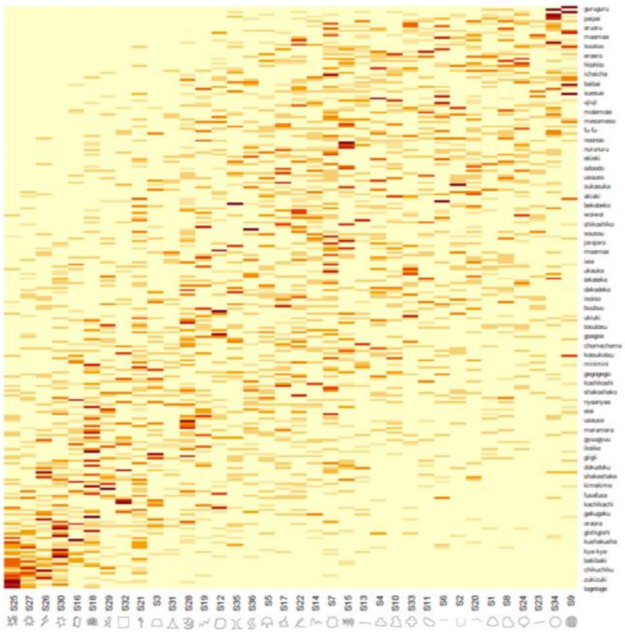

Figure 3: Result of correspondence analysis for Japanese speakers' corresponding sounds to shapes.

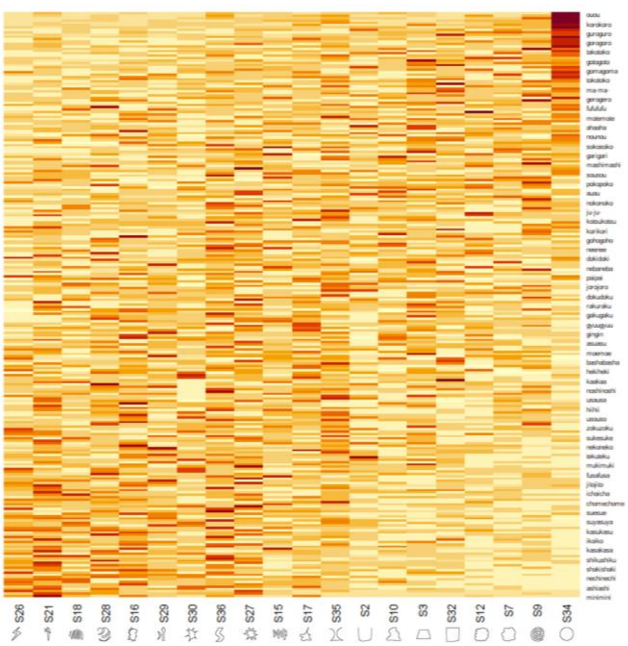

Figure 4: Result of correspondence analysis for non-Japanese speakers' corresponding sounds to shapes.

More specifically, it suggests that both Japanese speakers and non-Japanese speakers associated the vowels [o] and [u] to the round shapes, and did [i] and the fricatives to the pointed shapes.

There was an exceptional difference between the two groups, however: the phonetic unit [n] was associated to the round shapes by Japanese speakers, but it was associated to the pointed shapes by non-Japanese ones.

Rank Correlation As a result of correspondence analysis, properties of shapes, which lined on the edge of the axis of shapes, were similar in both Japanese speakers and non-Japanese speakers. Then, we analyzed correlation of the order of shapes and sounds, which were lined by using scores of correspondence analysis, by Spearman's rank correlation.

Correlation of sounds was weakly positive ($r = 0.246$, $p < 0.05$), and correlation of shapes was positive ($r = 0.680$, $p < 0.05$). These results showed that correspondence of sounds and shapes by both Japanese speakers and non-Japanese speakers were quite similar.

Discussion

In this study, we hypothesized that onomatopoeia in a natural language could be a class of words which had sound-symbolical property not only for the speakers of it but also for non-native speakers of it. Then we aimed to identify the common phonetic factors of onomatopoeias correlated to the sound-symbolic property.

In the experiment, we asked Japanese speakers and non-Japanese speakers to guess and choose the closest shape to which the speech sound of Japanese onomatopoeia-like words referred. The result of entropy analysis revealed some difference between Japanese speakers and non-Japanese speakers in their choice distributions.

To investigate the structural similarity between sounds and shapes for each group, we conducted correspondence analysis of the cross table of shape-to-sound choices (Figure 3, 4). As the result of the analysis, it is suggested that both Japanese and non-Japanese speakers have similar sound-shape correspondences in which a set of similar sounds is associated to rounded shapes, and the other set of similar sounds associated to pointed shapes. These general trends common in the two groups suggested that these sound-shape correspondences reflected their underlying sound symbolism such as the bouba-kiki effect. This suggests that Japanese onomatopoeia-like words have the sound symbolic system.

The stepwise regression analysis on the scores of the pseudo onomatopoeias confirmed the general agreement between Japanese and non-Japanese speakers, and further revealed the common phonetic factors correlated to the shape scores.

Conclusion

In sum, this study revealed a cross-linguistic common correspondence of sound and shapes, and identified its phonetic correlates to it. Although our experiment covered a wide range of pseudo onomatopoeias and various shapes, not limited to something particular for bouba-kiki effect, our result nonetheless showed that one of major structural similarity between shapes and sounds captures an underlying bouba-kiki like patterns. Thus, this finding supports the sound symbolism hypothesis (Imai et al., 2015), asserting that a class of onomatopoeias may serve a pre-linguistic support for potential new language learners.

References

Hamano, S. (2014). *Nihongo no onomatope: O'nsyōtyō to kōzō* [Onomatopoeia in Japanese: Sound symbolism and its structure] Tokyo: Kuroshio Shuppan.

Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H. & Kita, S. (2015). Sound symbolism facilitates word learning in 14-month-olds, *PLoS One 10*(2), e0116494

Köhler, W. (1929). *Gestalt Psychology*, New York, USA: Liveright Publishing Corporation.

McCormick, K., Kim, J., List, S. & Nygaard, L. C. (2015). Sound to Meaning Mappings in the Bouba-Kiki Effect., *Cognitive Science Vol. 2015*, 1565-1570

Ramachandran, V. S. & Hubbard, E. M. (2001). Synaesthesia—a window into perception, thought and language., *Journal of Consciousness Studies 8*(12), 3-34.

Saussure, F. De. (1916). *Course in General Linguistics*.

Shimizu, Y., Doizaki, R. & Sakamoto, M (2014). A System to Estimate an Impression Conveyed by Onomatopoeia, *Transactions of the Japanese Society for Artificial Intelligence, 29*(1), 41-52

Tamori, I. & Schourup, L. (1999). *Onomatope: Keitai to imi* [Onomatopoeia: Form and meaning] Tokyo: Kuroshio Syuppan

Tamori, I. (1998). Nihongo Onomatope: Tayō na Oto to Yōtai no Hyōgen [Japanese Onomatopoeia: Various Sounds and Expressions of Manners], *The Journal of the Acoustical Society of Japan 54*(3), 215-222

Yamaguchi, Y. & Shiina, K. (2005). Communication of affective state through the hand-drawn line-figures, *Library, Information and Media Studies 3*(2), 15-23

Table 3: Result of Stepwise regression analysis for score of sounds of Japanese speakers.

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.107	0.150	0.712	0.477	
n	0.591	0.174	3.407	0.001	***
u	0.365	0.107	3.416	0.001	***
o	0.352	0.104	3.394	0.001	***
e	0.281	0.134	2.101	0.037	*
i	-0.301	0.116	-2.598	0.010	**
k	-0.337	0.107	-3.140	0.002	**
ts	-0.549	0.287	-1.911	0.057	.
b	-0.618	0.219	-2.816	0.005	**
tɕ	-0.756	0.377	-2.004	0.046	*
d	-0.761	0.268	-2.838	0.005	**
zʲ	-0.762	0.375	-2.034	0.043	*
ɕʲ	-0.814	0.349	-2.330	0.021	*
z	-0.940	0.280	-3.358	0.001	***
ɡ	-1.000	0.166	-6.011	$<10^{-8}$	***
ɡʲ	-1.769	0.415	-4.260	$<10^{-4}$	***
kʲ	-1.980	0.823	-2.406	0.017	*

Table 4: Result of Stepwise regression analysis for score of sounds of non-Japanese speakers.

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.485	0.136	3.580	$<10^{-3}$	***
o	0.757	0.088	8.567	$<10^{-14}$	***
r	0.319	0.149	2.133	0.034	*
u	0.162	0.092	1.749	0.082	.
m	-0.254	0.139	-1.827	0.069	.
n	-0.369	0.157	-2.347	0.020	*
ɸ	-0.467	0.293	-1.593	0.113	
d	-0.483	0.241	-2.005	0.046	*
ɡ	-0.485	0.153	-3.171	0.002	**
k	-0.486	0.101	-4.796	$<10^{-5}$	***
ɕ	-0.504	0.208	-2.420	0.016	*
i	-0.521	0.102	-5.090	$<10^{-6}$	***
p	-0.555	0.265	-2.091	0.038	*
j	-0.619	0.367	-1.685	0.093	.
t	-0.636	0.197	-3.226	0.001	**
ɕʲ	-0.725	0.307	-2.365	0.019	*
z	-0.911	0.250	-3.645	$<10^{-3}$	***
ts	-0.939	0.262	-3.592	$<10^{-3}$	***
s	-0.979	0.153	-6.391	$<10^{-9}$	***
tɕ	-1.004	0.332	-3.022	0.003	**
zʲ	-1.113	0.336	-3.318	0.001	**
ɡʲ	-1.128	0.371	-3.043	0.003	**
nʲ	-1.148	0.423	-2.714	0.007	**
tɕʲ	-1.152	0.368	-3.134	0.002	**
kʲ	-2.003	0.719	-2.785	0.006	**