Letter visual similarity of Japanese hiragana and katakana based on reaction times

Hiroki Higuchi¹ · Tessei Kobayashi¹

Accepted: 21 December 2021
© The Author(s) 2022

Abstract

Letter similarity (i.e., perceptual distance) is a critical measure to better understand letter perception and literacy development. Despite its importance, however, measurements of letter similarity for non-alphabetic scripts are limited, and the shortage of letter similarity for non-alphabetic script interferes with the identification of the universality and the uniqueness of letter perception systems across different scripts. In the present study, we provide a comprehensive matrix of letter similarity for Japanese kana letters (hiragana and katakana). We obtained the discrimination reaction times for simultaneously presented letter pairs and calculated the perceptual distance of 4,278 letter pairs by inverting the time. The matrix showed significant correlations with previously obtained letter similarity for hiragana and katakana. An additional experiment showed that letter pairs for the same sounds (ㇰ–ㇱ) produced significantly slower responses compared with those for different sounds (ㇰ–ㇲ). However, the differences in reaction times between the same and different sound conditions were smaller than the sequentially presented conditions, suggesting that the matrix was partially attributable to knowledge-based factors (e.g., letter-sound knowledge). This first comprehensive matrix of letter similarity (i.e., perceptual distance) for Japanese kana letters (hiragana and katakana) will be useful for researchers interested in letter perception and literacy development.

Keywords Letter similarity · Perceptual distance · Hiragana · Katakana · Reading · Japanese

Introduction

Letter perception is a critical step in reading because we cannot read any words without identifying the presented letters. How we represent letters in our mind and develop their representation during childhood are intriguing questions in the field of cognitive and developmental psychology. Letter similarity, which is the visual (but not phonological) similarity of letters, is a powerful tool for understanding these issues. Letter similarity, which has been extensively measured in alphabetic scripts over the past 130 years, has contributed to studies that are theoretically building cognitive models of letter perception and reading as well as empirically investigating the development of reading related skills (see Mueller & Weidemann, 2012 for a review). For example, Treiman et al. (2007) addressed the cross-linguistic

issue of letter-naming acquisition by investigating the relationship between naming errors and letter properties, including letter similarity. The analysis of naming-error patterns showed that letter similarity had a similar effect in English- and Hebrew-speaking toddlers, suggesting that toddlers in both groups make visually-based confusions. In contrast, phonological similarity has a stronger effect on English-speaking toddlers than Hebrew-speaking toddlers because Hebrew has fewer phonologically similar letter pairs. These results imply a similar principle in letter-naming acquisition across different languages. Letter similarity is also useful for quantifying orthographic similarity. Although Coltheart et al. (1977) proposed using such orthographic neighbors as “time,” “lime,” and “mine” as a measure of orthographic similarity, they neglected a letter level of similarity. Despite the importance of letter similarity, less is known about it in non-alphabetic scripts. The present study provides new evidence for letter similarity from other languages, such as Japanese hiragana and katakana.

The Japanese language’s unique writing system consists of logographic kanji and two types of syllabic kana (hiragana and katakana), resulting in three types of writing systems in

✉ Hiroki Higuchi
hiroki.higuchi.sa@hco.ntt.co.jp

¹ Innovative Communication Laboratory, NTT
Communication Science Laboratories, 2-4, Hikaridai,
Seika-cho, Soraku-gun, Kyoto 619-0237, Japan

Japanese orthography (see Wydell & Butterworth, 1999 for details). While kanji characters are derived from Chinese characters, kana letters were created to simplify the surface forms of kanji characters. Kanji characters generally represent nouns (e.g., 会社 company), the root morphemes of inflected verbs (e.g., 行く to go), and adjectives (e.g., 楽しい fun). In contrast, hiragana letters represent the inflections of verbs, adjectives, and adverbs as well as function words. Katakana letters are mainly used for foreign words (e.g., インターネット internet). The Japanese kana system appears to resemble the upper and lower cases in the alphabet because hiragana and katakana letters represent identical sounds (e.g., hiragana あ and katakana ア represent /a/). However, the orthographic rules between kana and alphabetic scripts are completely different. For example, although a letter at the beginning of a sentence is capitalized in alphabetic script, this is not allowed in the Japanese kana writing system. Given these differences in writing systems, it might be fruitful to test whether cognitive models of letter perception and letter-knowledge development, which are established mainly based on studies of alphabetic scripts, are applicable to other language systems. Investigating similarity for Japanese kana (hiragana, katakana) letters is critical.

Despite the importance, however, very few studies have addressed letter similarity in Japanese kana compared with alphabetic script. In previous studies on hiragana, non-readers of Japanese performed a clustering task on 46 pairs of hiragana letters (Dunn-Rankin & Leton, 1973), and a subjective rating was conducted by Japanese-speaking adults on 71 pairs of hiragana letters (Kawakami & Tsuji, 2012). For katakana, the subjective letter similarity for 15 pairs of letters was investigated (Kaiho, 1970). Since the focus of this study was not to obtain a complete matrix of katakana letters, the evaluated pairs of letter similarity were limited. Other studies measured the letter similarity of every possible pair of the 71 katakana letters based on subjective ratings (Kawakami, 2002) and clustering tasks (Yamade & Haga, 2008). Given previous findings on Japanese kana, the subjective measurement of letter similarity for hiragana and katakana was obtained, although not the similarity across such script as hiragana お vs. katakana オ. Obtaining hiragana-katakana similarity across kana scripts is critical for two reasons. First, it can increase understanding of how children learn to name hiragana and katakana letters. In elementary school, Japanese-speaking children first learn hiragana and then katakana (Wydell & Butterworth, 1999). Perhaps a knowledge of hiragana influences the acquisition of katakana. For instance, hiragana-katakana letter pairs with a similar visual form (e.g., か-カ) are more easily acquired than those with a dissimilar form (e.g., あ-ア). Second, comprehensive data of letter similarity, including hiragana-katakana letter pairs, contribute to both theoretical and empirical research that is investigating how literate adults encode

hiragana and katakana letters, especially whether the two types of kana letters are categorically represented.

Although the letter similarity of Japanese kana has been mainly measured by subjective methods, such as rating and clustering, previous studies on alphabetic script have measured it in other ways. The most common procedure is letter-naming tasks (e.g., Fisher et al., 1969; Gilmore et al., 1979), in which letter similarity is measured by counting how many times participants misnamed a presented letter (i.e., confusability). For example, when the letter O was presented, the participants misnamed 14.2% and 0.3% of their responses as Q and Z (van der Heijden et al., 1984). This error rate probably reflects the visual similarity of letters (Grainger et al., 2008). However, Mueller and Weidemann (2012) pointed out that these errors are not a direct measure of letter similarity because a pair of letters was not presented simultaneously (i.e., not compared directly); it also takes a tremendous amount of time to collect enough naming errors because participants correctly identified most of the trials. Therefore, although many studies tried to increase the naming errors by applying psychological techniques, such as noise, speed, or low-contrast images (e.g., Gervais et al., 1984; Gupta et al., 1983; Tinker, 1928), collecting enough naming errors remains time consuming. The second procedure to efficiently measure letter similarity uses letter-rating tasks to evaluate the visual similarity between two letters on a Likert scale (e.g., Boles & Clifford, 1989; Kuennapas & Janson, 1969). However, because this is a direct (but subjective) evaluation of letter pairs, it remains unclear whether the rating values substantially reflect a representation of the letters in the participants' minds. The third procedure measures reaction times (RTs) in same-different judgments for letter pairs. Podgorny and Garner (1979) assumed that similar letter pairs take more time to be judged as the 'same' compared with dissimilar letter pairs, and Courrieu et al. (2004) proposed the inverse of RTs as a perceptual distance of letter pairs (i.e., an indicator of letter similarity). Unlike the number of errors in the letter-naming task, values based on RTs do not include zero elements, showing that a more useful indicator for investigating letter similarity should be free from struggling with zero elements. Thus, RT-based values are not subjective; they are an objective measure, obtained by directly comparing letter pairs. Furthermore, since this method does not require that letters be named, we can presumably purely measure the visual (but not phonological) similarity of letter pairs. In the present study, we applied this method to measure the letter similarity of Japanese kana.

In Experiment 1, we obtained the letter similarity of Japanese kana (hiragana and katakana) by measuring the discrimination RTs in a go/no-go task for presented letter pairs. Following a previous study (Courrieu et al., 2004), we asked the participants to only press a button as fast as possible if the letter pair was different (e.g., #あ#い#) and

not it if the letter pair was identical (e.g., #あ#あ#). By analyzing RTs in different letter pair condition, we obtained perceptual distance as a measure of letter similarity. Note that similarity (as a general concept) can be equally measured by similarity (similar letter pairs have greater values) and dissimilarity (dissimilar letter pairs have greater values). Distance is a similar measure with dissimilarity (dissimilar letter pairs have greater distance) that is required to satisfy some requirement. Given the evidence that it took more time to discriminate between visually similar letter pairs (Podgorny & Garner, 1979), Courrieu et al. (2004) proposed an inversed value of discrimination RTs as an indicator of letter similarity (i.e., perceptual distance) that allows the application of more powerful data analysis methods, such as multidimensional scaling. Like dissimilarity, similar letter pairs have smaller perceptual distances. For example, Courrieu et al. (2004) reported that the perceptual distance for visually similar b-d pairs were 69, although it was 124 for visually dissimilar y-c pairs. By applying this method, we obtained the perceptual distances for 4278 pairs of Japanese kana.¹

Although we are trying to obtain the perceptual distances of kana letters to determine how a letter pair is *visually* similar, there are no guarantees that other factors do not influence them. Since a letter is a visual object that represents sounds, phonological information might be implicitly activated during our task. In addition, as in the case of alphabetic script (a = A), literate Japanese speakers know that letters are arbitrary signs, and a certain pair of hiragana and katakana letters refers to the same sound (え = エ ≠ お). That is, such knowledge-based factors as letter-sound knowledge (conceptual knowledge) might be induced whenever we perceive a letter. Therefore, such a confounding factor is likely to affect the perceptual distance in our task.

One possible way to confirm the contribution of conceptual knowledge is provided by Lupyan et al. (2010) who used a same-different judgment task of alphabetic letter pairs consisting of identical sounds (Bb) and different sounds (Bp). Both lower-case letters (b and p) consist of a straight line and a semicircle, and the visual similarity between Bb and Bp is almost identical. If the participants only respond by a visual comparison of the letter pairs, one would expect that the RTs for the identical (Bb) and different (Bp) condition do not produce significant differences. This effect was observed in the two-presentation condition. In the simultaneous presentation condition, both letters were presented at the same time, as in our Experiment 1. In the sequential presentation condition, however, the first letters were presented before the second letter appeared. Lupyan et al. (2010) predicted

little or no difference in the simultaneous condition because their participants responded before the conceptual knowledge affected the ongoing visual processing. However, in the sequential presentation condition, Lupyan et al. (2010) predicted a greater difference between the same and different sound pairs because of the greater time that the conceptual knowledge interfered with the visual processing. As predicted, Lupyan et al. (2010) found a significant difference between the same and different sound pairs in the sequential condition but not in the simultaneous condition, suggesting a small amount of interference by conceptual knowledge in the simultaneous condition. Based on previous findings, we also simultaneously presented letter pairs in Experiment 1. Therefore, we confirmed that our obtained perceptual distance is mainly attributed by the visual comparison of letter pairs even if a similar phenomenon was observed in Japanese kana letters.

Experiment 2 examined the effect of conceptual knowledge in the same-different judgment based on the previous study (Lupyan et al., 2010). We presented participants with letter pairs of hiragana and katakana, such as え-エ (hiragana /e/ - katakana /e/) in the same-sound condition and え-コ (hiragana /e/ - katakana /ko/) in the different-sound condition. Although エ and コ are visually quite similar, they represent completely different sounds, and therefore different RTs between the same and different sound conditions can be interpreted as the effect of conceptual knowledge. Like in the previous study (Lupyan et al., 2010), we observed this effect when the letter pairs were only presented in the sequentially presented conditions. If there is little or no effect of conceptual knowledge in the simultaneously presented condition, our obtained data can be mainly attributed to the perceptual visual information of letters.

Experiment 1

Method

Participants

A total of 21 participants (13 women and eight men; *M* age = 26.4 years, range: 20–39 years of age) were included in the experiment. We removed three participants from the final sample because they wished to withdraw (*n* = 1), their accuracy was 2 SD below the mean (96.10%, *n* = 1), and they exceeded the predetermined age criteria (over 40 years old, *n* = 1). All participants were literate native Japanese speakers with normal or corrected-to-normal vision. The present experiment was approved by the ethics committee of our institute and written informed consent was obtained from each participant prior to the experiment. After completing

¹ Hereinafter we use ‘letter similarity’ to refer to this general concept.

Fig. 1 Scheme of Experiment 1. Prior to main trial, five hash marks were presented at center of screen. Two letters with hash marks were presented and participants pressed a button if the presented letters were physically different. Trial orders were randomized

the six to ten days of the experiment, each participant received ¥40,000, which is roughly \$365.

Stimuli and Design

The stimuli were Japanese hiragana (46 letters), katakana (46 letters), and a blank space, resulting in 93 letters. We included a blank space because blank spaces indicate clauses in many types of children books, although Japanese texts are not typically separated by spaces. Each participant judged every possible pair of different letters ($93 \times 92/2 = 4278$ pairs). Additionally, they were presented with the same number of identical pairs (4278 pairs), culminating in 46 presentation times for each identical letter pair. Hence, they completed 8556 trials for each set. To stably acquire each participant's RTs, they completed three sets, resulting in 25,668 trials throughout the experiment. We made two stimuli lists, half in ascending order and the other half in descending order, such that half of the participants were exposed to #あ#い# and the others were exposed to its reversed order (い#あ#).

Procedure

The experiment was conducted in accordance with the procedure of a previous study (Courrieu et al., 2004). Each trial started with the presentation of five hash marks (#####) for 750 ms, followed by presentation of the test letters (e.g., #あ#い#) (Fig. 1). The participants were instructed to press a button as fast as possible when the presented pair of letters

was different (go trial) and not to press it when the presented pair was identical (no-go trial). Trial orders were randomized for each participant. We set timeouts at 1.5 s. Go trials, where the participants did not respond before the timeout, were treated as errors. Our RTs analysis only included the responses in the go trials. 25,668 trials were divided into 138 sessions, resulting in 186 trials in each session. The experiment was carried out over 6 to 10 days due to the large number of sessions. Prior to this main experiment each day, we carried out simple detection RT and practice sessions. In the simple detection RT sessions, five jittering hash marks were presented from 600 to 1000 ms, followed by hash marks with steady, fixed symbols (##### or #####). The participants were required to immediately press buttons after the presentation of the fixed symbols, which consisted of 25 trials. Then they completed a practice task of 42 trials for alphabetic letter pairs (e.g., #a#b#). After checking that the participants correctly performed the practice trials, we moved to the main experiment. Stimuli were presented on a 24.5-in. LCD monitor at 60 Hz (BenQ ZOWIE XL2546, 1920 × 1080 pixels), controlled by PsychoPy Version 1.90.1 (Peirce, 2007, 2009; Peirce et al., 2019) on a Windows desktop computer. The participant responses were collected by a Logitech F310 Gamepad. Letters were presented in MS Gothic (height = 30).² The stimuli, shown at the center of the screen, covered a visual angle of about $0.7^\circ \times 4.2^\circ$.

Results and Discussion

Participant-Level Analysis

We first checked the task performance to identify participants who misunderstood the instructions or showed poor concentration. Although most performed almost perfectly ($M = 99.11\%$, $SD = 1.50$), one participant whose accuracy was 2 SD below the mean was removed from the following analysis.

We daily calculated the mean simple detection time (t_0) for each participant because the experiment was carried out across several days. The RTs for the fastest ten trials were used as simple detection RTs (t_0). This time was subtracted from each RT of each item-pair (i.e., $\text{DiffRT} = \text{itemRT} - t_0$). Since the experiments consisted of three sets, three values (DiffRT) were obtained for each pair if they correctly responded for all pairs. We averaged these values and divided the obtained values by their average to normalize the mean discrimination time to 1. If we could not average these values (i.e., they failed to respond for pair_{*i*} over two sets), the values of this pair were treated as missing values.

² 'MS Gothic' is a standard Japanese font in Microsoft Windows and is commonly used in software, e.g., Internet Explorer.

Table 1 Examples of matrix indicating letter similarity (i.e., perceptual distance)

	あ	い	う	え	お	か	き	く	け	こ	さ	し
あ												
い	103											
う	98	106										
え	112	100	96									
お	54	103	107	115								
か	97	111	105	114	103							
き	90	109	100	107	93	104						
く	106	99	104	105	109	105	103					
け	100	94	106	98	97	100	108	113				
こ	104	94	106	100	116	103	100	98	110			
さ	97	106	101	107	105	94	58	108	106	98		
し	106	94	95	107	109	101	114	83	94	101	104	

Missing values in the participant-level analysis were quite rare ($M = 1.47$, $Mdn = 1$).

Presentation Order of Discrimination Time Matrix

We counterbalanced the presentation order of the letter pairs across the participants such that half were exposed to #あ#い#, and the others were exposed in reverse order (#い#あ#). To check the effect of the presentation order, we calculated the mean value of these conditions (e.g., あ-い and い-あ) by averaging the discrimination time of each participant and checked the correlation between two presentation orders. Since we found a significant correlation between two presentation orders ($r = .606$, $p < .001$), we took the mean values of all the participants invariantly over the presentation order of the letter pairs.

Fig. 2 Distribution of perceptual distance (indicator of letter similarity). Perceptual distance of hiragana-katakana pairs (between) tends to be larger than that hiragana-hiragana (hiragana) and katakana-katakana (katakana). Letter pairs that participants recognized as completely identical (ゑ [hiragana]-ゑ [katakana]) were excluded from this analysis. Boxes represent median of first and third quartiles, and bars show min and max without outliers

Converting to Distance

The normalized differences between the item reaction times and the simple detection times over the presentation order were converted to distances by taking the inverse of these values (i.e., $d = 1/t$). The min (d) was 0.392, and the max (d) was 1.294. Since these values did not satisfy the relation $\max(d) \leq 2 \min(d)$, which is satisfactory for the triangle inequality (See Courrieu et al., 2004 for details), we checked the triangle inequality for all possible pairs. All possible pairs satisfied the triangle inequality, suggesting that the obtained values are metric. Additionally, not a single participant responded correctly to ゑ-ゑ (hiragana/he/ and katakana/he/) pairs, suggesting that all of them recognized these letters as completely identical.³ We assigned zero to this pair.

Summary of Perceptual Distance for Kana

The complete matrix for letter similarity (i.e., perceptual distance) multiplied by 100 can be obtained from online [Supplementary Materials](#). Since the obtained data were converted to distance, similar letter pairs have small values. Examples of this matrix are shown in Table 1. For example, the distance for visually similar あ-お pairs (54) was less than あ-い pairs (103). In addition, Fig. 2 shows the overall distribution for the letter pairs of hiragana-hiragana, katakana-katakana, and hiragana-katakana. The perceptual distance for hiragana-katakana was larger than hiragana-hiragana ($t(1608.6) = 14.09$, $p < .001$, $d = 0.59$) and katakana-katakana ($t(1670.5) = 20.83$, $p < .001$, $d = 0.86$). On the other hand,

³ Although these letters appear identical, there is a minor difference in the font we used. Since all of our participants recognized these letters were identical, we removed these pairs from the triangle inequality check.

the differences between hiragana-hiragana and katakana-katakana were significant, although with a weak effect size ($t(2062.5) = 4.84$, $p < .001$, $d = 0.21$). This indicates that the perceptual distance between different scripts (i.e., hiragana-katakana) tends to be more dissimilar than within scripts (i.e., hiragana-hiragana and katakana-katakana).

Comparison with Previous Studies

To confirm whether our matrix is attributed to *visual* perceptual information rather than conceptual knowledge (e.g., letter-sound knowledge), we compared it with previous studies that subjectively rated the visual similarity of hiragana (Kawakami & Tsuji, 2012) and katakana (Kawakami, 2002). Our matrix showed significant correlations with hiragana ($r = -.64$, $p < .001$) and katakana ($r = -.65$, $p < .001$). We also found a significant, but weak correlation ($r = -.38$, $p < .001$)⁴ between our matrix and the visual similarity obtained by a clustering task performed by non-Japanese-speaking adults for 46 hiragana letters (Dunn-Rankin & Leton, 1973). Although a possible reason for the weak correlation is that the previous matrix had too many zero elements, additional investigation is needed before reaching a final conclusion.

Experiment 2

Our findings from Experiment 1 showed that the obtained perceptual distance is highly related to previous data reflecting visual similarity. In Experiment 2, we empirically investigated the contribution of perceptual information and conceptual knowledge in the comparison of letter pairs.

Method

Participants

Sixteen adults (eight women, eight men; M age = 21.0 years; age range: 18–25 years) who did not participate in Experiment 1 were recruited and paid ¥5000 (about \$45). An additional participant was excluded from main (RTs) analysis whose task accuracy was below 90%.

Stimuli and Design

The stimuli are shown in Fig. 3. All pairs consisted of three letters: one hiragana letter (ㇿ) and two katakana letters (ㇰ and ㇱ). The critical pairs were ㇿ-ㇰ in the same-sound condition and ㇿ-ㇱ in the different-sound condition. Since ㇰ and ㇱ are quite similar at the pixel by pixel level, the visual similarity of the ㇿ-ㇰ pair was almost equal to that of ㇿ-ㇱ. In contrast, the ㇿ-ㇰ pair represents the same /e/ sound, although the ㇿ-ㇱ pair does not. Any significant difference in the RTs between these pairs is attributable to the conceptual knowledge of the letters (e.g., letter-sound knowledge). In addition to the critical pairs, ㇿ-ㇿ and ㇰ-ㇰ pairs were selected as stimuli for the no-go trials. Although there are three possible pairs for the no-go trials (ㇿ-ㇿ, ㇰ-ㇰ, and ㇱ-ㇱ), two were selected to match the numbers for the go and no-go trials.

Procedure

The procedure was identical as in Experiment 1 except for the insertion of a sequentially presented condition. The participants pressed a button when the letter pairs were physically different and did not when they were identical. In the sequential presentation condition, the first letter appeared for 150, 300, 450, or 600 ms before the second letter (i.e., there was nonzero stimulus onset asynchrony, or SOA). Since we predicted that longer SOAs would produce additional time to exert conceptual knowledge that leads to greater differences in RTs between the conditions in which the SOAs increased, different SOAs were included in the sequentially presented condition and determined based on a previous study (Lupyan et al., 2010). The first letter (e.g., ㇿ) was randomly presented on the left (e.g., #ㇿ#) or the right (e.g., #ㇿ#) side of the stimuli. An experimental session consisted of 576 trials. The numbers of same-different, within-between, and simultaneous-sequential were equal. Prior to the experimental session, the participants completed 47 practice trials in which alphabetic letters (ex., #e#F#) were used as stimuli. After checking the correct performance of the practice session, the participants moved to the experimental session. The RTs were measured after both letters appeared.

Results and Discussion

Accuracy was sufficiently high ($M = 98.35\%$, $SD = 1.35$). A two-way repeated measures ANOVA for accuracy did not reveal any significant main effects of condition (physically same and different; $F(3,48) = 2.07$, $p = .117$, $\eta^2_G = 0.04$) or SOA (0 [i.e., simultaneous presentation], 150, 300, 450, and 600 ms; $F(4,64) = 0.93$, $p = .450$, $\eta^2_G = 0.00$). Therefore, we only used the responses for different trials (i.e., go trials) for the following RT analysis. Trials for incorrect responses

⁴ A similar analysis was carried out in a previous study that obtained the visual similarity of alphabetic letters by RT data (Courrieu et al., 2004). The obtained perceptual distance, which was compared with the confusion matrix of non-literate toddlers, included many zero elements. Correlation obtained by the previous study was not as strong ($r = -.35$, $p < .001$) as in our study, indicating that the correlation in our study is not necessarily low.

Fig. 3 Stimuli and design for Experiment 2. Two letters were presented and participants pressed a button when presented letters were physically different. Different condition consists of trials that shared identical sounds (same sound) and did not share identical sounds (different sound). Letters were presented simultaneously and sequentially; in sequential condition, second letters appeared after a variable delay (stimulus onset asynchrony, or SOA)

Fig. 4 Mean RTs for same and different sounds conditions. Although difference in RTs between identical and different sounds conditions were significant in all SOAs, size of difference tended to be smaller in simultaneously compared with sequential presentation. Error bars represent ± 1 SE

or outliers (more than 2 SD away from the participant's mean) were also eliminated from the following RT analysis (5.14%).

The results of the RT analysis are shown in Fig. 4. A two-way repeated measures ANOVA determined the

effect of the condition (same-sound and different-sound) and the SOA (0, 150, 300, 450, and 600 ms). There was a significant main effect of condition ($F(1,15) = 26.32$, $p < .001$, $\eta_G^2 = 0.23$), such that the participants responded more slowly to pairs with the same-sound condition (エ-工, $M = 491$ ms, $SD = 49$ ms) than to those with the different-sound condition (エ-コ, $M = 452$ ms, $SD = 49$ ms). Slower responses for the same-sound condition were consistent with the previous findings using alphabetic letter pairs (Lupyan et al., 2010). Moreover, there was a significant main effect of SOA ($F(4,60) = 93.58$, $p < .001$, $\eta_G^2 = 0.46$). The RTs in the simultaneously presented pairs ($M = 523$ ms, $SD = 39$ ms) were slower than those in the sequentially presented pairs (SOA 150 ms: $M = 468$ ms, $SD = 44$ ms; SOA 300 ms: $M = 470$ ms, $SD = 42$ ms; SOA 450 ms: $M = 447$ ms, $SD = 36$ ms; SOA 600 ms: $M = 440$ ms, $SD = 44$ ms). Critically, we found a significant interaction between condition and SOA ($F(4,60) = 3.69$, $p = .009$, $\eta_G^2 = 0.02$). Although the RTs between the same-sound and different-sound conditions at each SOA were significantly different (all $ps < .01$), the difference between the conditions varies depending on the SOAs. While the simultaneously presented condition produced a 23-ms difference with a medium effect size ($t(15) = 3.20$, $p = .006$, $d_p = 0.798$), the sequentially presented condition produced significant differences

with larger effect sizes (SOA 150 ms: $M = 32$ ms, $d_D = 0.843$; SOA 300 ms: $M = 50$ ms, $d_D = 1.317$; SOA 450 ms: $M = 43$ ms, $d_D = 1.144$; SOA 600 ms: $M = 45$ ms, $d_D = 1.148$).

These findings demonstrate that both simultaneous and sequential presentations produced different RTs between the same-sound and different-sound conditions, suggesting that participants made a same-different judgment using conceptual knowledge. Although we cannot deny that our matrix is at least partially influenced by conceptual knowledge (e.g., letter-sound knowledge), note that this effect was smaller in the simultaneously presented condition than the sequentially presented condition.

Furthermore, the magnitude of the differences between same and different sound conditions seemed to vary among the sequentially presented conditions with a peak at SOA = 300 ms. In particular, although the differences of RTs between SOA 150 ms and SOA 300 ms were 18 ms, the differences of RTs for SOA 300 ms – SOA 450 ms and SOA 300 ms – SOA 600 ms were 7 and 5 ms. However, a previous study on alphabetic script showed no difference in RTs between different SOAs, although the variance decreased as SOA increased, implying that this effect becomes more reliable as SOA increases (Lupyan et al., 2010). Consistent with the previous findings, the present study also identified a tendency for the effect to become robust as SOA increases, but it saturated more quickly. Although the previous study did not show significantly different RTs in the simultaneous presentations, the present study showed a weak but significant difference. These differences might have emerged because of the easier access to conceptual knowledge in the present experiments that used a go/no-go task (in which a button was only pressed when different letters were presented) compared to Lupyan et al. (2010) task (in which a button was pressed when the presented letters were physically identical and different). Thus, the participants in the previous study may have relied more heavily on visual information because of a higher task load. However, since these studies used different languages, further investigation is needed to examine whether the differences in the RTs were induced by cross-linguistic differences.

To sum up, our present findings suggest that the discrimination time for simultaneously presented letter pairs might be partially attributed to conceptual knowledge. However, its contribution to the discrimination times was smaller than the sequential presentation condition.

General Discussion

The present study successfully provided a new measure to elucidate letter similarity (i.e., perceptual distance) for 4278 pairs of Japanese hiragana and katakana. Although many studies have addressed letter similarity in alphabetic scripts (see Mueller & Weidemann, 2012 for a review), it has been largely limited to non-alphabetic scripts. This lack of studies disrupts enterprises that seek to identify universality and the specificity of letter perception across different scripts. Therefore, we focused on Japanese kana letters to obtain the perceptual distance based on RTs in the same-different judgment (Courrieu et al., 2004). Although previous studies on Japanese kana subjectively rated letter similarity (Kawakami, 2002; Kawakami & Tsuji, 2012), our approach used RTs, which enable a participant's task engagement to be monitored in terms of accuracy and RTs and extract the implicitly judged similarities of letter pairs.

Another advantage of the present study is that our matrix includes the first cross-categorical similarity between hiragana and katakana letters; it was not obtained in previous studies (Kawakami, 2002; Kawakami & Tsuji, 2012). The letter similarity between these two types of kana will probably contribute to an acquisition study for examining how prior knowledge of hiragana influences the acquisition of katakana letter naming, especially whether visually similar-looking letters (カ/ka/ and カ/ka/) are easily acquirable. Letter similarity might also contribute to the consideration of how we mentally represent kana letters. Hiragana-katakana pairs tend to be more dissimilar than hiragana-hiragana and katakana-katakana pairs, implying that these kana scripts are categorically represented. A similar representation pattern was found in alphabetic letters within the same language. Boles and Clifford (1989) measured the alphabetic letter similarity for lower and upper cases and modeled a representation of alphabetic letters using multidimensional scaling (MDS). Dimension 1 in the MDS is clearly related to upper- and lower-case letters, indicating that they are categorically represented as in Japanese hiragana and katakana letters. These results suggest that categorical letter representation is universal across languages. One possible reason for such universality is that these categories (upper case vs. lower case and hiragana vs. katakana) were evolutionary selected to be visually distinguishable. However, since our focus is to collect letter similarity data and confirm the validity of the perceptual distance in Japanese kana, we did not perform any additional analysis. Further research is needed to model a representation of Japanese kana letters using our perceptual distance.

Our similarity measure of Japanese kana has several properties. First, ours is a distance matrix because all possible pairs satisfied triangular inequality. Second, our matrix does not

simply reflect the perceptual information of letters. Instead, it is reasonable to assume that perceptual distance is due to the information that is partially affected by conceptual knowledge because there is a significant, but weak correlation between our matrix and the previous data for a task with 46 hiragana letters performed by participants who cannot speak Japanese (Dunn-Rankin & Leton, 1973). Moreover, an additional experiment found a significant effect of conceptual knowledge even in simultaneously presented pairs, which is the same condition in tasks that obtain letter similarity. However, this effect was smaller than the sequentially presented pairs, suggesting that conceptual knowledge's effect is relatively (but not perfectly) controlled.

This work, which is the first comprehensive matrix for the letter similarity of Japanese hiragana and katakana, will become a valuable tool for psycholinguistic studies using Japanese kana stimuli and determining orthographic similarity. It will also be useful for disclosing the underlying mechanisms of letter perception. With our matrix, it is intriguing to examine whether the mechanism of letter perception, shown in alphabetic script (e.g., Mueller & Weidemann, 2012), is applicable to non-alphabetic scripts. Finally, our matrix can contribute to a precise understanding of how toddlers acquire the ability to read hiragana and katakana (Kakihana, 2015; Treiman et al., 2007), which is a critical benchmark for literacy development.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12144-021-02664-w>.

Acknowledgements We are grateful to Dr. Takashi Hattori (NTT Communication Science Laboratories) for his helpful suggestions regarding the data analysis.

Availability of data and materials The complete matrix for letter similarity (i.e., perceptual distance) can be obtained in the online supplementary materials.

Code availability The code during and/or analyzed during the current study available from the corresponding author on reasonable request.

Authors' contributions HH and TK were involved in study design and data interpretation. HH was involved in the data analysis. All authors critically revised the report, commented on drafts of the manuscript, and approved the final report.

Declarations

Conflicts of interest/competing interests The authors declare that they have no competing interests.

Ethics approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the ethics committee of NTT Communication Science Laboratories.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Boles, D. B., & Clifford, J. E. (1989). An upper- and lowercase alphabetic similarity matrix, with derived generation similarity values. *Behavior Research Methods, Instruments, & Computers*, 21(6), 579–586. <https://doi.org/10.3758/BF03210580>
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Lawrence Erlbaum.
- Courrieu, P., Farioli, F., & Grainger, J. (2004). Inverse discrimination time as a perceptual distance for alphabetic characters. *Visual Cognition*, 11(7), 901–919. <https://doi.org/10.1080/1350628044000049>
- Dunn-Rankin, P., & Leton, D. A. (1973). Differences between physical template matching and subjective similarity estimates of Japanese letters. *Japanese Psychological Research*, 15(2), 51–58. <https://doi.org/10.4992/psycholres.1954.15.51>
- Fisher, D. F., Monty, R. A., & Glucksberg, S. (1969). Visual confusion matrices: Fact or artifact? *The Journal of Psychology*, 71(1), 111–125. <https://doi.org/10.1080/00223980.1969.10543077>
- Gervais, M. J., Harvey, L. O., & Roberts, J. O. (1984). Identification confusions among letters of the alphabet. *Journal of Experimental Psychology: Human Perception and Performance*, 10(5), 655–666. <https://doi.org/10.1037/0096-1523.10.5.655>
- Gilmore, G. C., Hersh, H., Caramazza, A., & Griffin, J. (1979). Multidimensional letter similarity derived from recognition errors. *Perception & Psychophysics*, 25(5), 425–431. <https://doi.org/10.3758/BF03199852>
- Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: From pixels to pandemonium. *Trends in Cognitive Sciences*, 12(10), 381–387. <https://doi.org/10.1016/j.tics.2008.06.006>
- Gupta, S. M., Geyer, L. H., & Maalouf, J. A. (1983). Effect of font and medium on recognition/confusion. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 144–149. <https://doi.org/10.1145/800045.801599>
- Kaiho, H. (1970). Katakana-moji sougokanno ruijisei handan jigen to miyasusa no kankei [Similarity dimensions and legibility of Kata-Kana letters]. *Shinrigaku Kenkyu [The Japanese Journal of Psychology]*, 40(6), 337–340. <https://doi.org/10.4992/jjpsy.40.337>
- Kakihana, S. (2015). Yoji no kana-moji yomi syutoku ni eikyouosu moji-gawa no syoyouin [The influence of features of Kana letters on children's letter sound knowledge acquisition]. *Hattarusu Shingiraku Kenkyu [The Japanese Journal of Developmental Psychology]*, 26(3), 237–247. <https://doi.org/10.11201/jjdp.26.237>

- Kawakami, M. (2002). Katakana-moji no keitaiteki ruijisei hyoutei chousa [A figurative similarity rating table for 71 Japanese katakana letters]. *Dokuso Kagaku [The Science of Reading]*, 46(4), 132–139.
- Kawakami, M., & Tsuji, H. (2012). Hiragana-moji no keitaiteki ruijisei hyoutei chousa [Figurative similarity ratings for 71 Japanese hiragana]. *Dokuso Kagaku [The Science of Reading]*, 54(3–4), 80–88. https://doi.org/10.19011/sor.54.3-4_80.
- Kuennapas, T., & Janson, A. J. (1969). Multidimensional similarity of letters. *Perceptual and Motor Skills*, 28(1), 3–12. <https://doi.org/10.2466/pms.1969.28.1.3>
- Lupyan, G., Thompson-Schill, S. L., & Swingle, D. (2010). Conceptual penetration of visual processing. *Psychological Science*, 21(5), 682–691. <https://doi.org/10.1177/0956797610366099>
- Mueller, S. T., & Weidemann, C. T. (2012). Alphabetic letter identification: Effects of perceivability, similarity, and bias. *Acta Psychologica*, 139(1), 19–37. <https://doi.org/10.1016/j.actpsy.2011.09.014>
- Peirce, J. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Peirce, J. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2(10), 10. <https://doi.org/10.3389/neuro.11.010.2008>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Podgorny, P., & Garner, W. R. (1979). Reaction time as a measure of inter- and intraobject visual similarity: Letters of the alphabet. *Perception & Psychophysics*, 26(1), 37–52. <https://doi.org/10.3758/BF03199860>
- Tinker, M. A. (1928). The relative legibility of the letters, the digits, and of certain mathematical signs. *The Journal of General Psychology*, 1(3–4), 472–496. <https://doi.org/10.1080/00221309.1928.9918022>
- Treiman, R., Levin, I., & Kessler, B. (2007). Learning of letter names follows similar principles across languages: Evidence from Hebrew. *Journal of Experimental Child Psychology*, 96(2), 87–106. <https://doi.org/10.1016/j.jecp.2006.08.002>
- van der Heijden, A. H. C., Malhas, M. S. M., & van den Roovaart, B. P. (1984). An empirical interletter confusion matrix for continuous-line capitals. *Perception & Psychophysics*, 35(1), 85–88. <https://doi.org/10.3758/BF03205927>
- Wydell, T. N., & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition*, 70(3), 273–305. [https://doi.org/10.1016/S0010-0277\(99\)00016-5](https://doi.org/10.1016/S0010-0277(99)00016-5)
- Yamade, Y., & Haga, S. (2008). Katakana-moji no keitaiteki ruijisei ni taisuru syukanteki hyouka: iyakuhin-mei no ruizisei youin toshite [Subjective evaluation of similarity of appearance of katakana characters in drug names]. *Rikkyo-Daigaku Shinrigaku Kenkyu [Rikkyo Psychological Research]*, 50, 79–85.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.