Journal of Experimental Psychology:
Human Perception and Performance
2017, Vol. 43, No. 1, 126–146

0096-1523/17/$12.00 http://dx.doi.org/10.1037/xhp0000281

Phonological-Orthographic Consistency for Japanese Words and Its Impact on Visual and Auditory Word Recognition

Yasushi Hino and Yuu Kusunose
Waseda University

Shinobu Miyamura
INTAGE Inc., Tokyo, Japan

Stephen J. Lupker
The University of Western Ontario

In most models of word processing, the degrees of consistency in the mappings between orthographic, phonological, and semantic representations are hypothesized to affect reading time. Following Hino, Miyamura, and Lupker's (2011) examination of the orthographic-phonological (O-P) and orthographic-semantic (O-S) consistency for 1,114 Japanese words (339 katakana and 775 kanji words), in the present research, we initially attempted to measure the phonological-orthographic (P-O) consistency for those same words. In contrast to the O-P and O-S consistencies, which were equivalent for kanji and katakana words, the P-O relationships were much more inconsistent for the kanji words than for the katakana words. The impact of kanji words' P-O consistency was then examined in both visual and auditory word recognition tasks. Although there was no effect of P-O consistency in the standard visual lexical-decision task, significant effects were detected in a lexical-decision task with auditory stimuli, in a perceptual identification task using masked visual stimuli, and in a lexical-decision task with degraded visual stimuli. The implications of these results are discussed in terms of the impact of P-O consistency in auditory and visual word recognition.

Keywords: phonological-orthographic consistency, Japanese kanji words, Japanese kana words, orthographic-phonological interaction

An almost universal assumption is that the visual word recognition processes in any given language are affected by the nature of the relationships between orthography, phonology, and semantics for the words of that language (e.g., Fushimi, Ijuin, Patterson, & Tatsumi, 1999; Grainger, Muneaux, Farioli, & Ziegler, 2005; Grainger & Ziegler, 2007; Hino & Lupker, 1996; Hino, Lupker, & Pexman, 2002; Hino, Nakayama, Miyamura, & Kusunose, 2011; Jared, McRae, & Seidenberg, 1990; Pexman, Lupker, & Reggin, 2002; Stone, Vanhoy, & Van Orden, 1997; Ziegler, Montant, & Jacobs, 1997). In the Japanese language, the language used in this investigation, words can be written in one of three different scripts (although a given word is typically written only in one of the scripts), two of which are syllabaries (katakana and hiragana—the "kana" scripts) and one of which is logographic (kanji). Due to the orthographic differences between kana and kanji, a common assumption has been that the nature of orthographic-phonological (O-P) and orthographic-semantic (O-S) relationships must be different for kana versus kanji words (e.g., Feldman & Turvey, 1980; Frost, 2005; Kimura, 1984; Saito, 1981; Wydell, Butterworth, & Patterson, 1995), which should further imply that orthographic, phonological, and semantic processing will be different for words written in the different scripts. Providing an examination of some of these assumptions is the basic goal of the present research.

O-P and O-S Relationships for Kana and Kanji Words

As just noted, Japanese uses three different scripts: kanji, hiragana and katakana. Kanji is a logographic script and each kanji character is considered a morpheme, representing meaning. Thus, in general, one would expect that words sharing a kanji character would tend to have similar meanings (e.g., 男性 [male,/da.N.se.i/] and 男子 [boy,/da.N.si/]). At the same time, however, most kanji characters have multiple pronunciations. According to Tamaoka, Kirsner, Yanase, Miyaoka, and Kawakami (2002), among 1,945 basic Japanese kanji characters, 64.22% (1,249 characters) possess multiple pronunciations: 1,168 possess both an original Japanese pronunciation (the so-called "kun-reading" pronunciation) and a pronunciation that originated in Chinese (the so-called "on-reading" pronunciation), 73 possess more than one on-reading pronunciation but no kun-reading pronunciation and eight possess more than one kun-reading pronunciation but no on-reading pronunciation. As a result, kanji characters are often pronounced in

---

Yasushi Hino and Yuu Kusunose, Faculty of Arts and Sciences, Waseda University; Shinobu Miyamura, Marketing Innovation Division, INTAGE Inc., Tokyo, Japan; Stephen J. Lupker, Department of Psychology, The University of Western Ontario.

This research was supported by JSPS KAKENHI(24530924).

Correspondence concerning this article should be addressed to Yasushi Hino, Faculty of Arts and Sciences, Waseda University, 1-24-1 Toyama, Shinjuku-ku, Tokyo 162-8644, Japan. E-mail: hino@waseda.jp

different ways in different word contexts (e.g., 親父 [father,/*o.ja*.zi/] and 親戚 [relatives,/*si.N*.se.ki/]).[1]

In contrast, the two kana scripts (i.e., hiragana and katakana) are phonetic scripts and, hence, each kana character basically corresponds to a single mora, a rhythmic unit of a constant duration consisting of either a single vowel or the combination of a consonant and a vowel.[2] Thus, any kana character is almost always pronounced the same (e.g., イス [chair,/i.*su*/] and リス [squirrel,/ri.*su*/]). Because kana characters, unlike kanji characters, are not morphemes, it would seem unlikely that words sharing kana characters would have similar meanings (e.g., ポケット [pocket,/po.ke.Q.to/] and ロケット [rocket,/ro.ke.Q.to/]).

Based on these differences in the nature of kanji and kana scripts, it has generally been assumed that whereas O-P relationships are much more consistent for kana words than for kanji words, O-S relationships are more consistent for kanji words than for kana words. Further, these assumptions have led to the theoretical position that words are processed differently depending on their script type, along the lines of the assumptions made by the orthographic depth hypothesis (e.g., Frost, 2005; Frost, Katz, & Bentin, 1987). In particular, phonological coding for kana words is assumed to be accomplished by simply applying print-to-sound correspondence rules (i.e., an "assembly" route like that found in the dual-route cascaded model; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). In contrast, because of the more complicated O-P relationships for kanji words, the presumption is that phonological coding for these words can only be accomplished via the mental lexicon (i.e., a "lexical" route; e.g., Wydell et al., 1995). Further, based on the assumption that O-S relationships are more consistent for kanji words than for kana words, the assumption is that, whereas the process of retrieving lexical/semantic information is driven directly by orthography for kanji words, this process is typically mediated by phonology for kana words (e.g., Kimura, 1984; Saito, 1981).

Initially, there were a number of studies reporting evidence consistent with these theoretical positions (e.g., Feldman & Turvey, 1980; Kimura, 1984; Saito, 1981; Wydell et al., 1995). More recently, however, there have been a number of studies reporting evidence against some of these positions (e.g., Besner & Hildebrandt, 1987; Fushimi et al., 1999; Hino & Lupker, 1998; Hino, Lupker, Sears, & Ogawa, 1998; Yamada, 1992). Further, based on a systematic analysis of the degree of O-P and O-S consistency for 1,114 Japanese words (339 katakana and 775 kanji words), Hino, Miyamura, and Lupker (2011) recently reported that both the O-P and the O-S consistencies were actually quite similar for their kanji and katakana words, in contrast to what is typically assumed about the two word types. Thus, at this point, it is far from clear whether the processes involved in reading kana and kanji words are qualitatively different.

The details of Hino, Miyamura, et al.'s (2011) analyses are as follows. In order to measure O-P and O-S consistencies, Hino, Miyamura, et al. (2011) generated a list of orthographic neighbors of each of their 1,114 words, based on the definition of neighbors provided by Coltheart, Davelaar, Jonasson, and Besner (1977), using National Language Research Institute (1993). The orthographic neighbors were, then, classified as phonological friends or enemies based on whether the shared characters are pronounced the same at the moraic level in the orthographic neighbor and the target word. In addition, Hino, Miyamura, et al. (2011) also measured the degree of similarity in meaning between the orthographic neighbors and the target word using subjective ratings, allowing orthographic neighbors to be classified as semantic friends or enemies. Then, the O-P and O-S consistencies were computed using National Language Research Institute (1970) word frequency norms.

Although many kanji characters possess multiple pronunciations, kanji characters are usually read with on-reading pronunciations when they are used as constituents of compounds and are usually read with kun-reading pronunciations when they are used as single-character words (e.g., Wydell, 1998). In addition, although most katakana characters are mapped to unique morae, the most frequently used katakana character (a macron, "ー" as in "サービス[service]") can be mapped onto five different vowel phonemes. It is, essentially, these two facts that led to the O-P relationships for the kanji compounds being as consistent as those for the katakana words according to Hino, Miyamura, et al.'s (2011) analyses.

As for the O-S consistency, although orthographic neighbors of a kanji word share a kanji character with that word, orthographic neighbors are not necessarily semantic friends because kanji characters are often used in different senses. For example, although "助手 (assistant)" and "拍手 (handclap)" involve the same constituent character, "手," it denotes "a person" and "a hand," respectively and, therefore, these two words are not considered to be semantically similar. Hence, these words were classified as semantic enemies rather than semantic friends. Such was the case for most kanji words analyzed (i.e., only about 12.5% of the neighbors of kanji words were semantic friends). Consequently, Hino, Miyamura, et al.'s (2011) results indicated that the O-S consistency was essentially equivalent for their kanji and katakana words.

The Nature of Feedback (P-O) Relationships for Kana and Kanji Words

O-P and O-S relationships, which both appear to be similar for katakana and kanji words, are not, of course, the only relationships that may impact processing. In particular, as suggested by Stone, Vanhoy, and Van Orden (1997), P-O consistency (what those authors call "feedback consistency" when discussing visual word recognition) may also be important in reading. That is, if one assumes that word recognition processes involve only the bottom-up flow of information, there would be no reason to expect that reading performance is affected by the feedback relationships (i.e., from phonology to orthography, from semantics to orthogra-

---

[1] Japanese is a moraic language with morae being phonological units that correspond, more or less, to syllables. When we describe morae using characters from the Roman alphabet, we will use the format from Tamaoka and Makioka (2004a) with a period (.) denoting a moraic boundary.

[2] When describing the phonological structure of Japanese sounds, we will follow Tamaoka and Makioka's (2004a) definitions of phonemes, morae, and syllables. Note also that, as Tamaoka and Makioka note, there are some exceptional types of morae in Japanese. Although the regular types of morae (with CV or V structure) are identical to syllables, there are three types of exceptional morae which do not correspond to a syllable (/Q/,/N/and/R/). For example, syllables with a geminate (e.g.,/ki.Q/in/ki.Q.te/, a stamp), syllables with a nasal (e.g.,/ri.N/in/ri.N.go/, an apple) and syllables with a long vowel (e.g.,/te.R/in/te.R.pu/, a tape) consist of two morae. Hence, words containing these types of syllables have a syllabic structure that is different from their moraic structure.

phy, etc.) possessed by words. However, if one assumes that word recognition processes involve the top-down flow of information in addition to the bottom-up flow of information as suggested by the interactive-activation model and various connectionist models (e.g., Grainger et al., 2005; Grainger & Ziegler, 2007; McClelland & Rumelhart, 1981; Stone et al., 1997; Ziegler et al., 1997), one would expect that reading performance would be affected by the feedback relationships possessed by words.

For present purposes, the most important consideration is that it appears that kanji words may be quite P-O inconsistent, potentially allowing for a strong manipulation of the P-O consistency factor when using kanji stimuli. For example, Hino, Kusunose, Lupker, and Jared (2013) recently reported that logographic languages (e.g., Chinese and Japanese kanji) generally possess a large number of homophones. That is, in logographic languages, there are a large number of words that are pronounced the same but are written differently. For example, as pointed out by Tan and Perfetti (1998), a given pronunciation can be generated by, on average, 11 different characters in Chinese, meaning that Chinese words typically possess a large number of homophonic mates. Similarly, because Japanese kanji is also a logographic script, homophones often do not have similar orthographic forms (e.g., 目 [eye,/me/] & “芽 [sprout, /me/]”) and there are a large numbers of homophones for kanji words in general. Therefore, P-O consistency does appear to be weak for kanji words which may be an important issue in reading those words.

In contrast, because Japanese kana scripts (both hiragana and katakana) are phonetic and the character-to-mora relationships are transparent, it is essentially impossible to write a katakana word or a hiragana word using different characters. Further, although katakana words can be transcribed into hiragana and hiragana words can be transcribed into katakana, most words are printed in only a single script, so that most katakana words are written only in katakana and most hiragana words are written only in hiragana. As a result, although kana words can occasionally be homophonic with kanji words (e.g., “イカ [squid, /i.ka/],” “医科 [the medical department, /i.ka/],” and “異化 [dissimilation, /i.ka/]”), the number of homophones would be considerably smaller for kana words than for kanji words.

In order to substantiate this idea, we counted the number of homophones for kana and kanji words listed in Amano and Kondo’s (2003b) Japanese word frequency norms. As Hino, Miyamura, et al. (2011) have pointed out, more than 80% of kanji words are two characters in length and more than 80% of kana words are three to five characters in length according to National Language Research Institute (1993). Thus, we counted the number of homophones for kanji words with one or two characters (90,477 words in total) and for kana words with one to five characters (45,045 words in total) listed in Amano and Kondo’s norms. The average numbers of homophones were 5.94 for the 90,477 kanji words and 1.72 for the 45,045 kana words, $F(1, 135{,}520) = 5{,}294.05$, $MSE = 101.09$, $p < .001$.

Because homophones are words having the same pronunciation even though they are spelled differently, homophones, by definition, involve an inconsistent P-O relationship. Thus, it seems quite likely that, on average, P-O relationships are more inconsistent for kanji words than for kana words. Hence, as will be explained below, the present investigation focused on kanji words.

Feedback (P-O) Consistency Effects

As noted, Stone et al. (1997) have argued that P-O (feedback) consistency is important in reading. This conclusion was based on their report of a feedback (P-O) consistency effect in their (English) visual lexical-decision task. This effect has been interpreted in the following way. When we read words, phonological activation arises automatically (e.g., Perfetti, Bell, & Delaney, 1988; Van Orden, 1987) following orthographic processing. The activated phonological codes would then feed activation back to the orthographic level. The strength of the feedback activation would be determined by the P-O consistency, with more consistent relationships producing more focused feedback. Orthographic processing and, hence, lexical decision performance, would be presumed to be facilitated by strong feedback consistency (e.g., Lacruz & Folk, 2004; Perry, 2003; Stone et al., 1997; Ziegler et al., 1997).

Although the studies by Stone et al. (1997) and others have been criticized (e.g., Peereman, Content, & Bonin, 1998; Ziegler, Petrova, & Ferrand, 2008) due to the possibility that their feedback consistency manipulations were confounded with other factors (i.e., experiential familiarity and spelling-sound consistency), results from other studies provide good support for the idea that there is an automatic flow of activation from phonological representations to orthographic representations. In particular, there are a number of studies demonstrating a P-O consistency effect in auditory word recognition experiments, studies that did not suffer from the original confounds (e.g., Peereman et al., 1998; Ziegler & Ferrand, 1998; Ziegler et al., 2008). For example, using the same stimuli in both visual and auditory lexical-decision tasks, Ziegler, Petrova, and Ferrand (2008) examined both O-P and P-O consistency effects. In their auditory task, a significant P-O consistency effect was observed, although no O-P consistency effect emerged. In contrast, neither effect was detected in their visual task. Although the flow of activation from phonological representations to orthographic representations when processing auditory stimuli is a feedforward, rather than a feedback, flow, the existence of these P-O consistency effects in auditory word recognition, does indicate that activation automatically flows in the P-O direction in at least some situations.

As Ziegler et al. (2008) noted, there would be essentially two possible ways to account for their pattern of results, both based on the idea that, in both tasks, lexical decisions are made based on orthographic information.[3] The first is based on the idea that there is bidirectional (i.e., feedforward and feedback) activation between orthography and phonology. When a visual stimulus is presented, that stimulus would activate its orthographic code and its corresponding phonological code with the activated phonological code then feeding activation back to the orthographic level. Since the phonological feedback would activate all the spelling patterns consistent with the word’s phonology, competition would be created at the orthographic level if the P-O relationship is inconsistent. Thus, the orthographic processing necessary to settle on the

---

[3] Ziegler et al. (2008) noted that the lack of a feedforward (O-P) consistency effect in their visual task was consistent with the previous literature (e.g., Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Waters & Seidenberg, 1985), in which an O-P consistency effect was not observed unless the stimulus set involved words with unique spelling-sound relationships (i.e., strange words like “yacht”).

correct orthographic code should take longer for an inconsistent word than for a consistent word. However, a feedback consistency effect would be difficult to detect for visual stimuli because the orthographic codes could be rapidly cleaned up through the use of the visual input because the visual stimulus usually remains visible throughout the time that lexical decisions are being performed. In an auditory lexical-decision task, a P-O inconsistent stimulus would also activate a number of competing orthographic codes (in a feedforward manner). However, because auditory stimuli are temporal in nature, the presented stimulus would not be available to aid in processing. As such, the ambiguity would have to be resolved only through the use of an orthographic analysis with no help from an external stimulus. The process of settling on the correct orthographic code would, therefore, take longer when the P-O relationships are inconsistent, allowing a P-O consistency effect to emerge in an auditory task, as was reported by Ziegler et al. (2008).

Ziegler et al.'s (2008) second potential account was based on the idea that the assumption of feedback activation from phonology to orthography could be abandoned. That is, although the activation spreads from orthography to phonology for visual stimuli and from phonology to orthography for auditory stimuli, once the codes are activated in the other domain (i.e., the phonological code for the visual stimuli and the orthographic code for the auditory stimuli), there is no activation feeding back to the originally activated codes. Hence, no competition would be produced at the orthographic level in a visual task although there would still be a problem created at the orthographic level by the feedforward P-O activation in an auditory task. As such, it would follow that there would be no P-O consistency effect in a visual task but there would be a significant P-O consistency effect in an auditory task.

The Present Research

Although the data are somewhat mixed, there definitely is evidence that P-O consistency is an important factor in word recognition, at least in auditory word recognition. Therefore, if the P-O relationships are often inconsistent for kanji words, it may be easier to observe an impact of P-O consistency for kanji words than for most other types of words. In an effort to evaluate this idea, we initially attempted to measure the P-O consistencies for the 1,114 words used by Hino, Miyamura, et al. (2011).

In order to measure the P-O consistency for a word, it is necessary to capture the degree to which similarly pronounced words are mapped onto similar spelling patterns. Further, in order to be able to compare kanji and katakana words, it is also necessary to compare P-O consistencies for the two word types based on shared phonological properties. To accomplish this goal, we generated phonological neighbors for both kanji and katakana words by replacing a single mora from each target word. We used a mora-based definition of phonological neighbors because it has been suggested by a number of researchers that morae play the most central role when processing Japanese words (e.g., Otake, Hatano, Cutler, & Mehler, 1993; Kureta, Fushimi, & Tatsumi, 2006; Verdonschot et al., 2011). Our definition of whether these phonological neighbors were orthographic "friends" or "enemies" was then based on the nature of the shared characters. In particular, we determined whether the shared morae are printed with the same characters in the word pair. If most of the phonological neighbors generated by replacing a single mora contained the same characters as the original word (i.e., they are orthographic "friends"), that word would be considered to have high P-O consistency. In contrast, if most of phonological neighbors were spelled with different characters from those used in the original word (i.e., they are orthographic "enemies"), that word would be considered to have low P-O consistency. As will be described below, because our calculations of P-O consistency indicated that the P-O consistency was much lower for kanji words than for katakana words, our experiments involved an examination of the impact of P-O consistency in auditory and visual word recognition tasks using kanji words.

The overall goal of our experiments was to ascertain whether there is bidirectional (i.e., feedback) activation across orthography and phonology in visual word recognition, activation which may, then, impact the process of recognizing words. Although Stone et al.'s (1997) results have been challenged, there is one visual word recognition phenomenon that appears to necessitate the assumption of bidirectional activation across orthography and phonology (i.e., P-O feedback) when reading words. In particular, homophone effects have been reported in visual lexical-decision tasks in many different languages (e.g., Chen, Vaid, & Wu, 2009; Edwards, Pexman, & Hudson, 2004; Ferrand & Grainger, 2003; Hino, Kusunose, Lupker, & Jared, 2013; Kerswell, Siakaluk, Pexman, Sears, & Owen, 2007; Pexman & Lupker, 1999; Pexman, Lupker, & Jared, 2001; Pexman et al., 2002; Rubenstein, Lewis, & Rubenstein, 1971; Ziegler, Tan, Perry, & Montant, 2000). In English and French, those effects are inhibitory. Interestingly, using Japanese kanji compounds, Hino et al. (2013) reported that, although a homophony disadvantage arose when homophones possessed only a single homophonic mate (as in other languages), a homophony advantage arose when homophones possessed multiple homophonic mates (as in studies using Chinese). Although it may be possible to account for the homophony advantage with multiple homophonic mates without assuming feedback (P-O) activation, it is difficult to account for any homophony disadvantage without assuming P-O feedback.

Indeed, an obvious question is why a homophony disadvantage is consistently observed in the visual lexical-decision task while Stone et al.'s (1997) feedback (P-O) consistency effect is not. One possibility is that the contrast between homophones and nonhomophones involves a manipulation of P-O consistency at the whole-word level, whereas, in the feedback consistency studies, P-O consistency was manipulated at the subword level. Hence, the difference might be a quantitative one. If so, the effects of phonological feedback may be observed even in visual tasks whenever the consistency manipulation is strong enough, a result that would lead to the conclusion that the assumption of bidirectional activation across orthography and phonology would need to be included in any workable model of the processes involved in reading words. Therefore, our initial goal was to examine the P-O consistencies of kanji and katakana words to verify that the former are quite P-O inconsistent.

The P-O Analysis

In order to measure the P-O consistency for each of the 1,114 words used by Hino, Miyamura, et al. (2011), we first generated a list of each word's phonological neighbors by replacing a single mora from each target word. For example, for the target word, "ゲスト (guest, /ge.su.to/)," the words "キャスト (cast, /kja.su.to/)," "テスト (test, /te.su.to/)," "ベスト (best, /be.su.to/)," and "下水 (sewer, /ge.su.i/)" were all determined to be phonological neighbors because they were all one mora different from the target word. These phonological neighbors were, then, classified as orthographic friends or enemies based on whether the morae shared with the target word are printed with the same characters. Thus, "キャスト," "テスト," and "ベスト" were classified as orthographic friends but "下水" was classified as an orthographic enemy of the target word, "ゲスト." After classifying the phonological neighbors, summed frequencies of the orthographic friends and enemies were determined and the P-O consistency was computed using the following formula:

> P-O Consistency = (Target Frequency + Summed Frequency of Orthographic Friends)/(Target Frequency + Summed Frequency of all the Phonological Neighbors)

As with Hino, Miyamura, et al.'s (2011) O-P consistency index, this index takes a value between 0 and 1 depending on the degree of P-O consistency, being close to 1 for words with more consistent P-O correspondences (indicating that phonologically similar words tend to be printed with the same characters as the target word) and close to 0 for words with highly inconsistent P-O correspondences (indicating that phonologically similar words tend to be printed with different characters from those used in the target word).[4]

Method

Stimuli. The 339 katakana words (three to five characters in length, 3.95 on average), and 775 kanji words (all two characters in length) used by Hino, Miyamura, et al. (2011) were evaluated. All 1,114 words were nouns.

Procedure. For each of the 1,114 words, phonological neighbors were generated using National Language Research Institute (1993) and were classified as orthographic friends or enemies as shown in Table 1. The summed frequencies of the orthographic friends and enemies were, then, determined using the frequency norms of National Language Research Institute (1970). When a neighbor was not listed in the norms, the frequency count was assumed to be zero. Based on the target frequency and the summed frequencies of the orthographic friends and enemies, the P-O consistency was computed for each of the 1,114 words.[5]

Results

The mean P-O consistencies and the mean summed frequencies of the orthographic friends and enemies for the 339 katakana words and the 775 kanji words are shown in Table 2 along with their mean word frequencies, orthographic neighborhood sizes, phonological neighborhood sizes, word lengths, and the numbers of morae. The P-O consistencies were also computed using Amano and Kondo's (2003b) word frequency norms. As described in Appendix A, the results were unchanged when Amano and Kondo's norms were used instead of National Language Research Institute (1970) norms.

Mean numbers of orthographic friends and enemies were 1.37 and 4.80 for the 339 katakana words and 1.54 and 24.64 for the 775 kanji words, respectively. As such, although the number of orthographic friends was comparable for the katakana and kanji words, $F(1, 1112) = .67$, $MSE = 10.29$, the number of orthographic enemies was larger for the kanji words than for the katakana words, $F(1, 1112) = 299.80$, $MSE = 309.66$, $p < .001$, $\eta^2 = .21$. As a result, the P-O consistency was significantly higher for the katakana words (.83) than for the kanji words (.27), $F(1, 1112) = 814.80$, $MSE = .09$, $p < .001$, $\eta^2 = .42$.

Our stimulus set involved 186 katakana words and 53 kanji words with no phonological neighbor listed in the frequency norms (National Language Research Institute, 1970). Although the computed P-O consistency is 1.00 for all of these words, these words clearly possess unique phonological forms and, hence, unique P-O relationships. Giving the consistency values for these words the same weight as those for words that actually have phonological neighbors may produce an overestimate when computing the mean P-O consistency.

What also needs to be noted is that the degree of overestimation is likely different for the kanji versus katakana words because there are more katakana words with no phonological neighbors (186) than there are kanji words with no phonological neighbors (53). In order to address this issue, we recomputed the mean P-O consistencies for the katakana and kanji words after removing words of this sort. The results of that analysis are shown in Table 3. As seen in Table 3, after removing the 186 katakana and 53 kanji words with no phonological neighbor listed in the frequency norms, the mean P-O consistency for the katakana words decreased to .62 (from .83). The mean P-O consistency also decreased, although to a lesser extent, for kanji words (from .27 to .21). Nonetheless, the mean P-O consistency was still significantly higher for the 153 katakana words than for the 722 kanji words, $F(1, 873) = 309.13$, $MSE = .07$, $p < .001$, $\eta^2 = .26$.

Discussion

Hino, Miyamura, et al. (2011) reported that the O-P consistencies for kanji words were higher than previously assumed and essentially equivalent to those for katakana words. That result contrasts sharply with the present analysis of P-O consistencies which show that those consistencies were much higher for the

---

[4] Assuming that the degree of P-O consistency for a word is determined by learning experiences with phonologically-similar words (e.g., Jared et al., 1990; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989), a P-O consistency measure based on token frequency would appear to be a better measure than one based on type frequency. Thus, we used summed token frequencies of friends and enemies to compute P-O consistencies.

[5] Because Hino, Miyamura, et al. (2011) generated orthographic neighbors using National Language Research Institute (1993), consisting of 36,780 word entries, we used the same dictionary to generate phonological neighbors. In addition, as in Hino, Miyamura, et al. (2011), we also used National Language Research Institute (1970) to calculate word frequencies, which lists only words whose word frequency counts are more than four per 940,533.

In addition, when generating the list of phonological neighbors for each target word, we generated that list by replacing a single mora from the target, using the definition of "mora" provided by Tamaoka and Makioka (2004a). As a result, we didn't consider accent type or the type of phonemes involved in each of the phonological neighbors because, at least at present, there is no reason to believe that these factors have any impact on the degree of P-O consistency.

Table 1

An Example of Computing the P-O Consistency for a Katakana Word, "ギャング(gang, /gja.N.gu/)"

Target	Frequency		Frequency
ギャング(gang, /gja.N.gu/)	7		
Orthographic friend		Orthographic enemy	
キング(king, /ki.N.gu/)	14	玩具(toy, /ga.N.gu/)	7
リング(ring, /ri.N.gu/)	0	寝具(bed clothes, /si.N.gu/)	13
		天狗(a long-nosed goblin, /te.N.gu/)	5
		年貢(tribute, /ne.N.gu/)	0
		文具(stationery, /bu.N.gu)	8
Total	21		33

Note. Orthographic neighbors generated using National Language Research Institute (1993) were classified as orthographic friends or enemies based on whether the shared morae between the neighbor and target are printed with the same characters. The frequency counts were taken from National Language Research Institute (1970). The P-O consistency of "ギャング" = 21 / (21 + 33) = .39.

katakana words than for the kanji words. An obvious question, of course, is whether the P-O consistencies were so different for the katakana and kanji words due to some artifact in the way we calculated consistencies.

One objection that could be raised is that our results might be a function of how we defined P-O consistency because that definition has slightly different implications for the calculation of P-O consistency for words in the different scripts. This issue arises because katakana and kanji words differ in the nature of their character-to-mora correspondences. Whereas each katakana character corresponds to a single mora, kanji characters often correspond to multiple morae.

Table 2

Mean P-O Consistency, Target Frequency Plus Summed Frequency of the Phonological Friends (Friends), Summed Frequency of the Phonological Enemies (Enemies), Word Frequency (Freq), Orthographic Neighborhood Size (ON), Phonological Neighborhood Size (PN), Word Length, and Number of Morae for the 339 Katakana Words and the 775 Kanji Words

	Script type	
Variable	339 katakana words	775 kanji words
P-O consistency	.83	.27
Friends	32.13	47.57
Enemies	47.37	268.51
Freq	18.67	27.15
ON	1.77	47.59
PN	6.17	26.18
Word length	3.95	2.00
Number of morae	3.83	3.66

Note. Mean word frequencies were higher for the kanji words than for the katakana words, $F(1, 1112) = 7.67$, $MSE = 2213.46$, $p < .01$, $\eta^2 = .01$. Mean orthographic neighborhood sizes were higher for the kanji words than for the katakana words, $F(1, 1112) = 900.80$, $MSE = 549.85$, $p < .001$, $\eta^2 = .45$. Mean phonological neighborhood sizes were also higher for the kanji words than for the katakana words, $F(1, 1112) = 277.15$, $MSE = 340.75$, $p < .001$, $\eta^2 = .20$. In contrast, mean word lengths were significantly greater for the katakana words than for the kanji words, $F(1, 1112) = 5175.23$, $MSE = .17$, $p < .001$, $\eta^2 = .82$. Mean numbers of morae were also greater for the katakana words than for the kanji words, $F(1, 1112) = 17.37$, $MSE = .36$, $p < .001$, $\eta^2 = .02$.

In order to get a better handle on the issue of character-to-mora correspondences for kanji characters, we initially counted the number of morae for 1,945 basic kanji characters listed in Tamaoka and Makioka (2004b). Because a number of kanji characters possessed multiple pronunciations, we counted the smallest and largest numbers of morae for each character (for characters

Table 3

Mean P-O Consistency, Target Frequency Plus Summed Frequency of the Orthographic Friends (Friends), Summed Frequency of The Orthographic Enemies (Enemies), Word Frequency (Freq), Orthographic Neighborhood Size (ON), Phonological Neighborhood Size (PN), Word Length, and Number of Morae for the 153 Katakana Words and the 722 Kanji Words After Removing the Words With no Phonological Neighbors Listed in National Language Research Institute (1970)

	Script type	
Variable	153 katakana words	722 kanji words
P-O consistency	.62	.21
Friends	47.15	49.17
Enemies	104.97	288.22
Freq	17.33	27.25
ON	3.64	47.89
PN	12.88	27.89
Word length	3.48	2.00
Number of morae	3.38	3.64

Note. Mean number of orthographic friends was significantly larger for the 153 katakana words (2.76) than for the 722 kanji words (1.64), $F(1, 873) = 12.98$, $MSE = 12.24$, $p < .001$, $\eta^2 = .02$. In contrast, the mean number of orthographic enemies was significantly larger for the kanji words (26.25 than for the katakana words (10.12), $F(1, 873) = 92.87$, $MSE = 353.57$, $p < .001$, $\eta^2 = .10$. In addition, mean word frequencies were higher the kanji words than for the katakana words, $F(1, 873) = 5.10$, $MSE = 2433.56$, $p < .05$, $\eta^2 = .01$. Mean orthographic neighborhood sizes were larger for the kanji words than for the katakana words, $F(1, 873) = 385.65$, $MSE = 645.97$, $p < .001$, $\eta^2 = .31$. Mean phonological neighborhood sizes were larger for the kanji words than for the katakana words, $F(1, 873) = 74.11$, $MSE = 383.55$, $p < .001$, $\eta^2 = .08$. Mean numbers of morae were greater for the kanji words than for the katakana words, $F(1, 873) = 28.67$, $MSE = .29$, $p < .001$, $\eta^2 = .03$. Mean word lengths were smaller for the kanji words than for the katakana words, $F(1, 873) = 4316.15$, $MSE = .06$, $p < .001$, $\eta^2 = .83$.

with a single pronunciation, the same numbers were used in the smallest and the largest counts). For the smallest counts, the number of morae ranged from one to three, with an average of 1.62 and 61.54% (1,197 characters) corresponded to more than one mora. For the largest counts, the number of morae ranged from one to five, with an average of 2.03 and 88.53% (1,722 characters) corresponded to multiple morae. As such, it appears that at least 60% of kanji characters possess pronunciations with multiple morae. What is important for this discussion is the fact that, when kanji words consist only of characters with multiple morae, it would be essentially impossible to find orthographic friends in their phonological neighborhood using the definitions we used and, hence, all phonological neighbors must, by necessity, be orthographic enemies.

The explanation for why this is so is as follows: As shown in Table 1, we define a phonological neighbor as an orthographic friend if *all* the morae shared with the target are printed using the same characters as those in the target. Consider what happens with a two-character kanji word target with four morae, where each character corresponds to two morae. All phonological neighbors will share three morae with that target. While it is possible to find phonological neighbors in which two of the three shared morae are printed using the same character as the target, the third shared mora must come from different characters in the two words. For example, "確実 (certain, /ka.ku.zi.tu/)" is a phonological neighbor of the target, "着実 (steady, /tja.ku.zi.tu/)." The word, "確実" is not an orthographic friend of the target, "着実," however, because only two (rather than three) of the four morae are printed using the same character (e.g., 実 [/zi.tu/]). Such would be the case for any pairs of words of this sort which means that it is impossible for this type of kanji word to have an orthographic friend in its phonological neighborhood using our definitions. In contrast, for katakana words with four morae, it is common to find orthographic friends in the phonological neighborhood (e.g., ジャケット [jacket,/zja.ke.Q.to/] and ソケット [socket,/so.ke.Q.to/] for ポケット [pocket,/po.ke.Q.to/]) because, in most cases, each character corresponds to a single mora. The fact that it is impossible for some kanji words to have orthographic friends in their phonological neighborhoods means that the computed P-O consistencies must, almost by necessity, be smaller for kanji words than for katakana words.

If this factor were the only reason for the small P-O consistencies for kanji words, those small P-O consistencies for kanji words would essentially be an artifact of our definition of orthographic friends in the phonological neighborhoods. In order to address this issue, we reclassified the phonological neighbors of kanji words as orthographic friends if they are orthographic neighbors of the target regardless of whether all the shared morae are written using the same characters.

As noted above, the word, "確実 (certain, /ka.ku.zi.tu/)" was originally classified as an orthographic enemy of "着実 (steady, /tja.ku.zi.tu/)" because not all the shared morae (/ku.zi.tu/) were written with the same characters (i.e., only two morae,/zi.tu/, were written with the same character, "実"). With the new classification scheme, however, "確実" was now classified as an orthographic friend because it is an orthographic neighbor of the target, "着実." With this classification scheme, the mean P-O consistency for the 775 kanji words increased to .48 (from .27), however, that value was still significantly smaller than that for the 339 katakana words (.83), $F(1, 1112) = 257.42$, $MSE = .11$, $p < .001$, $\eta^2 = .19$. Similarly, for the 722 kanji words possessing phonological neighbors listed in the frequency norms, although the mean P-O consistency increased to .44 from .21, their P-O consistency was still significantly smaller than that for the 153 katakana words with phonological neighbors listed in the frequency norms (.62), $F(1, 873) = 35.49$, $MSE = .11$, $p < .001$, $\eta^2 = .04$. These results clearly indicate that, although some of the difference between kanji and katakana words in terms of P-O consistency is due to how we defined orthographic friends and enemies, it is not the only reason for the difference. Rather, there must be another reason why the P-O relationships are more inconsistent for kanji words than for katakana words.

What appears to be the main factor here is simply that there are many more homophonic characters in kanji than in katakana. As noted by Hino et al. (2013), there are, on average, 9.04 homophonic characters for each single-character kanji word. Therefore, there would be a strong chance for kanji words to have phonological neighbors that involved different kanji characters (i.e., words that were kanji-written, orthographic enemies). For katakana words, on the other hand, most characters have only one pronunciation. Therefore, virtually all phonological neighbors of katakana words that were also written in katakana would, by definition, be orthographic friends. In fact, the only possible homophones of katakana words would be words written in kanji (or hiragana). While it would be possible to find words written in other scripts that are phonological neighbors of a particular word as well as being, of course, orthographic enemies of that word, their number would be somewhat limited. As a result, one would expect that katakana words would have higher P-O consistencies than kanji words.

In order to examine this idea more fully, we classified the phonological neighbors of the 153 katakana and 722 kanji words in terms of their script types. Phonological neighbors of each word were classified as being (a) katakana words, (b) hiragana words, (c) kanji words, and (d) others, which consisted of words written in a combination of different scripts or words written in other scripts (such as the Roman alphabet). Then, the mean proportions of katakana and kanji words in the phonological neighborhood were calculated for the katakana and kanji words, respectively. For the 153 katakana words, the proportion of katakana words in their phonological neighborhoods was 45.90%, whereas the proportion of kanji words in the phonological neighborhoods of these words was 36.07%. These numbers contrast with those for the 722 kanji words, for which the proportion of kanji words in their phonological neighborhoods was 94.33% and the proportion of katakana words in their phonological neighborhoods was 0.69%. These results indicate that whereas the phonological neighborhoods of katakana words are somewhat heterogeneous, the majority of those words are written in katakana which means that, due to the nature of katakana, most, if not all, would be orthographic friends. In contrast, most of the phonological neighbors of kanji words are written in kanji. Because there are many homophonic kanji characters, the phonological neighborhoods of kanji words would involve many words with homophonic kanji characters (i.e., orthographic enemies). As a result, the P-O relationships would be expected to be more inconsistent for the kanji words than for the katakana words.

Given that P-O relationships are more inconsistent for kanji words than for katakana words, it should be easier to examine the impact of P-O consistency in visual and auditory word recognition by using kanji words. That is, by using kanji words, we should be able to manipulate P-O consistency much more strongly than in previous studies using alphabetic languages (e.g., Lacruz & Folk, 2004; Peereman et al., 1998; Perry, 2003; Stone et al., 1997; Ziegler et al., 1997; Ziegler & Ferrand, 1998; Ziegler et al., 2008). Therefore, if the bidirectional activation assumption is correct, there would be a good chance of observing a P-O consistency effect for kanji words not only in an auditory lexical-decision task (Experiment 1) but also in a visual lexical-decision task (Experiment 2).

If it were possible to observe P-O consistency effects for kanji words not only in the auditory task but also in the visual task, the results would clearly suggest that bidirectional activation across orthography and phonology would have to be assumed in models of the reading process. In contrast, if we failed to observe a feedback consistency effect in the visual task, it would still be unclear at that point, as explained by Ziegler et al. (2008), whether there is bidirectional activation across orthography and phonology when reading words. Note, however, that we would be able to further examine this issue by using visually degraded stimuli. That is, if the lack of a feedback consistency effect in visual tasks is due to the fact that an orthographic code could be rapidly cleaned up by information obtained directly from the visual stimulus, there would be a greater chance of observing a feedback consistency effect for visually degraded stimuli because it would be more difficult to obtain clean-up information directly from degraded stimuli. Thus, in such a circumstance, a feedback consistency effect may emerge.

If there is no bidirectional activation across orthography and phonology, on the other hand, there would be no real reason to expect a feedback consistency effect in a visual task even when the visual stimuli are degraded. In order to evaluate this issue, we also examined the feedback consistency effect for kanji words in a perceptual identification task with masked visual stimuli (Experiment 3) and in a visual lexical-decision task with degraded stimuli created by a luminance reduction (Experiment 4).

Experiments 1 and 2

Method

Participants. Seventy undergraduate and graduate students from Waseda University participated in these experiments. Thirty participated in Experiment 1 (the auditory task) and the rest participated in Experiment 2 (the visual task). They were paid a small amount of money (500 yen) in exchange for their participation. All were native Japanese speakers who had normal or corrected-to-normal vision.

Stimuli. Based on the data from the P-O analysis, 24 kanji words with high P-O consistency and 24 kanji words with low P-O consistency were selected. These were all two-character kanji words with three morae.[6] The mean P-O consistencies were .755 for the more consistent words and .039 for the less consistent words according to National Language Research Institute (1970) frequency norms, $F(1, 46) = 585.28$, $MSE = .01$, $p < .001$. According to Amano and Kondo's (2003b) frequency norms, the mean P-O consistencies were .571 and .070 for the more and less consistent words, respectively, $F(1, 46) = 82.20$, $MSE = .04$, $p < .001$. As described in Appendix A, the results from the analyses of P-O consistencies were essentially the same regardless of which frequency norms are used. In order to describe our experimental stimuli, therefore, we only report the values based on Amano and Kondo's frequency norms because these norms are newer and larger.

In addition, we also computed the P-O consistencies when orthographic neighbors were all classified as orthographic friends in the phonological neighborhoods (using Amano & Kondo's, 2003b, norms). The mean values were significantly higher for the more consistent words (.594) than for the less consistent words (.206), $F(1, 46) = 27.71$, $MSE = .07$, $p < .001$.

According to Amano and Kondo (2003b), mean word frequency counts (per 287,792,787 words) were equivalent for the more consistent words (4,128.13) and the less consistent words (4,117.04), $F(1, 46) = .00$, $MSE = 19{,}560{,}959.90$. In addition, as shown in Table 4, because Amano and Kondo's (2003a) database involved familiarity ratings for visual and auditory stimuli, we equated our two word groups on both types of familiarity ratings, all Fs < 1. Orthographic neighborhood size (computed using National Language Research Institute, 1993), summed character frequency (from Amano & Kondo, 2003b), and O-P consistencies (computed using Amano & Kondo's, 2003b, frequency norms) were also equated across the two word groups, all Fs < 1.

In contrast, mean phonological neighborhood size was larger for the less consistent words (34.96) than for the more consistent words (20.29), $F(1, 46) = 8.71$, $MSE = 296.30$, $p < .01$. As noted by Grainger, Muneaux, Farioli, and Ziegler (2005), when P-O consistency is manipulated for words, it will typically affect the sizes of orthographic and phonological neighborhoods. As Grainger et al. (2005) explain, while most phonological neighbors are orthographic friends (i.e., orthographic neighbors) for words with more consistent P-O relationships, for words with less consistent P-O relationships, most phonological neighbors would be orthographic enemies (i.e., not orthographic neighbors). Thus, when orthographic neighborhood size is equated for the more consistent and less consistent words, the less consistent words would have to have a large number of orthographic enemies in the phonological neighborhood. Hence, the phonological neighborhood size would become larger for the less consistent words than for the more consistent words. As such, by manipulating P-O consistency for words with similar orthographic neighborhood sizes, such a difference in the phonological neighborhood size

---

[6] In the corpus contained in National Language Research Institute (1993), 60.35% of word entries are kanji words, 17.80% are kana words, and 21.66% are words written in a combination of kana and kanji characters. As such, kanji words are the most frequently used words in Japanese vocabularies. Further, as noted, 80% of kanji words are two-character words. Thus, although one could argue that it is unclear whether one can generalize our results to words written in other scripts or shorter or longer words, these results should be generalizable to a very large percentage of Japanese words. One could, of course, also question whether our results, based only on nouns, would generalize to other syntactic classes of words. There is no reason to believe that they would not. Unlike the distinctions between kanji and kana scripts and between two-character kanji words and one- or three-character kanji words, the distinction between nouns and other syntactic classes is not structural. Hence, it, presumably, would not impact the nature of either orthographic or phonological neighborhoods.

Table 4
Stimulus Characteristics of the Two Groups of Kanji Words Used in Experiments 1, 2, 3, and 4

	P-O consistency	
Variable	More	Less
Morae	3.00	3.00
Freq	4,128.13	4,117.04
Visual fam	5.64	5.63
Auditory fam	5.56	5.46
ON	49.71	52.67
PN	20.29	34.96
CF	59,2243.75	56,5175.46
O-P consistency	.781	.776
P-O consistency	.571	.070
P-O consistency (ON)	.594	.206
Rel to left	5.50	5.50
Rel to right	4.80	4.71
Left FS	17.67	24.58
Right FS	32.04	28.08
Left FF	61,312.17	70,234.29
Right FF	141,826.33	95,367.46

Note. Morae, Freq, Visual Fam, Auditory Fam, ON, PN, and CF stand for mean number of morae, word frequency, familiarity rating for visual stimuli, familiarity rating for auditory stimuli, orthographic neighborhood size, phonological neighborhood size, and summed character frequency, respectively. O-P consistency and P-O consistency are mean O-P and P-O consistency values, respectively. P-O consistency (ON) is the mean P-O consistency value when orthographic neighbors were all classified as orthographic friends. Rel to left (Right) stands for mean relatedness rating between the left (right) constituent character and the Kanji compound. Left (right) FS and Left (right) FF stand for mean family size and mean family frequency of the left (right) constituent, respectively.

would be a natural outcome. We attempted to address this potential issue when analyzing the data.

In addition, because our word stimuli were all compound words, we attempted to equate the degree of transparency between the constituent kanji characters and the compound words across the word groups. For this purpose, we collected relatedness ratings between the constituent characters and the compound words. Using 60 two-character kanji compounds including the 48 words used in Experiments 1 and 2, a questionnaire was created, in which each of the 60 kanji words was presented twice: once paired with its left constituent kanji character and the other time paired with its right constituent character. The 120 pairs were, then, randomly ordered and listed in the questionnaire. Each kanji compound–kanji character pair was accompanied by a 7-point scale ranging from 1 (*unrelated*) to 7 (*related*). Twenty-six participants who did not participate in any of the present experiments were asked to rate the relatedness of these pairs by circling the appropriate number on the scale. Mean ratings between the compound and the left constituent character as well as the mean ratings between the compound and the right constituent character were comparable across the two word groups, all Fs < 1.

Further, in order to equate morphological connectivity for the constituents of compound words across the two word groups, family size and family frequency of the left (right) constituent character with the target compound were computed using Amano and Kondo's (2003b) frequency norms (e.g., Kuperman, Schreuder, Bertram, & Baayen, 2009). The family sizes of the left constituents were comparable across the two word groups, $F(1, 46) = 1.21$, $MSE = 472.81$. Similarly, the family sizes of the right constituents were also comparable across the two word groups, $F(1, 46) = .42$, $MSE = 446.10$. The family frequencies of the left constituents, $F(1, 46) = .08$, $MSE = 11{,}455{,}853{,}489.22$, and the family frequencies of the right constituents, $F(1, 46) = 1.16$, $MSE = 22{,}347{,}414{,}532.20$, were also both comparable across the two word groups. The two groups of kanji words are listed in Appendix B.

In addition to the 48 kanji words, 12 two-character kanji words with three morae were included as fillers in the stimulus set. In addition, 60 kanji nonwords were created by pairing two unrelated kanji characters and were included in the stimulus set. Because these kanji nonwords consisted of characters with single pronunciations, these nonwords were all pronounceable and consisted of three morae. Furthermore, 8 two-character kanji words and 8 two-character kanji nonwords that were not among the 120 experimental stimuli were used as practice stimuli.

Whereas all the stimuli were presented visually in the visual lexical-decision task in Experiment 2, these stimuli were presented aurally through headphones in the auditory lexical-decision task in Experiment 1. For Experiment 1, each of the stimuli were pronounced with a female voice and recorded as a WAV file on a PC. The sound lengths of these files were equated for the two word groups. The mean lengths were 499.96 ms for the more consistent words and 500.42 ms for the less consistent words, $F(1, 46) = 1.17$, $MSE = 2.15$. The mean sound length of the 60 nonwords was 500.28 ms.

Procedure. Participants were tested individually in a normally lit room. In Experiment 1, stimuli were presented binaurally through headphones (STAX, SR-307) connected to a PC with a driver unit (STAX, SRM-323S). Participants were seated in front of the video monitor (Iiyama, HM204DA) at a distance of about 50 cm. They were asked to decide whether or not an auditory stimulus presented through the headphones is a real word and respond by pressing either the "Word" or "Nonword" key on a response box interfaced to the PC through an IO card (Contec, PIO-16/16T(PCI)H).

In Experiment 2, stimuli were presented in the center of the video monitor. They were presented in white on a black background at a luminance of 33.28 cd/m^2 as measured by a luminance meter (Konica Minolta, Spectroradiometer conditional stimulus (CS)-1000A) using a 10 mm × 10 mm square at the center of the video monitor in a darkened room. Participants were asked to decide whether or not a kanji character string that appeared at the center of the video monitor was a word and respond by pressing either the "Word" or "Nonword" key on the response box interfaced to the PC. In both experiments, participants were told that their responses should be made as quickly and as accurately as possible. The "Word" response was always made using the participant's dominant hand. Sixteen practice trials were given prior to the 120 experimental trials. The order of stimulus presentation for the experimental trials was randomized for each participant.

In both experiments, each trial was initiated with a 50 ms 400 Hz warning tone, after which a fixation point appeared at the center of the video monitor. In Experiment 1, 1 s after the onset of the fixation point, a stimulus was presented through the headphones. The fixation point remained on the video monitor until the participant's key press. In Experiment 2, a stimulus was visually presented directly above the fixation point on the video monitor 1 s after the onset of the fixation point. The participant's response terminated the presentation of the fixation point and the stimulus. In both experiments, the response latencies from the onset of the stimulus to the participant's key press and whether the response was correct were automatically recorded by the PC. The intertrial interval was 2 s.

Results

Experiment 1 (auditory lexical-decision task). Lexical decision latencies were classified as outliers if they were out of the range of 2.5 standard deviations (*SD*s) from the mean for each participant. With this procedure, 3.61% (52 data points) of the experimental "Word" trials were classified as outliers and, thus, excluded from the statistical analyses. After excluding the outliers, 12.03% (167 data points) of the remaining experimental "Word" trials were errors and, thus, these trials were excluded from the latency analyses. Mean lexical decision latencies and error rates for the experimental "Word" trials are presented in Table 5. Lexical decision latencies for the correct experimental "Word" trials were analyzed using a linear mixed-effects (LME) model analysis (Baayen, Davidson, & Bates, 2008). Similarly, errors for these trials were analyzed using a logit mixed model analysis (Jaeger, 2008). Across all our experiments, we first fitted a model that included random intercepts for subjects and items (i.e., the intercept model). Following Barr, Levy, Scheepers, and Tily (2013), we also attempted to fit a maximal model including random slopes as well as random intercepts.[7]

In Experiment 1, the models were fitted to log-transformed lexical decision latencies and errors with P-O consistency as a fixed factor (Consistency). Instead of using a categorical variable for P-O consistency, we entered the actual P-O consistency values (computed based on Amano & Kondo's, 2003b, frequency norms) into the models. In addition, as we previously noted, because we could not control phonological neighborhood sizes across the word groups, we also entered Phonological Neighborhood Size (PhonoN) into the models as a control variable. According to model comparisons, the fit of the model was not significantly improved by entering the P-O Consistency × Phonological Neighborhood Size interaction into the model. Thus, we report the results from the models without that interaction.

In the analyses using the intercept model, the "lmer" syntax for R was "*Y* ~ *Consistency* + *PhonoN* + *(1* | *Subject)* + *(1* | *Item),*" in which *Y* stands for the dependent variable. In the analyses using the maximal model, on the other hand, a by-subject slope was assumed for P-O consistency in addition to the random intercepts for subjects and items because P-O consistency was a within-subject factor. Thus, the "lmer" syntax of this model was "*Y* ~ *Consistency* + *PhonoN* + *(1* + *Consistency* | *Subject)* + *(1* | *Item).*"

In the latency analysis using the intercept model, the effect of P-O consistency was significant, estimated coef. = −0.048, *SE* = 0.021, $t = -2.258$, $p = .029$, whereas the effect of phonological neighborhood size was not, estimated coef. = 0.000, *SE* = 0.000, $t = 1.218$. Similarly, using the maximal model, the effect of P-O consistency was significant, estimated coef. = −0.048, *SE* = 0.021, $t = -2.239$, $p = .030$, whereas the effect of phonological neighborhood size was not, estimated coef. = −0.000, *SE* = 0.000, $t = 1.208$.

In the error analyses, neither the effect of P-O consistency, estimated coef. = −1.932, *SE* = 1.363, $z = -1.418$, nor the effect of phonological neighborhood size, estimated coef. = 0.021, *SE* = 0.022, $z = .921$, was significant in the intercept model. The results were the same in the maximal model: neither the effect of P-O consistency, estimated coef. = −1.932, *SE* = 1.363, $z = -1.418$, nor the effect of phonological neighborhood size was significant, estimated coef. = 0.021, *SE* = 0.022, $z = 0.920$.

Experiment 2 (visual lexical-decision task with clear stimuli). As in Experiment 1, lexical decision latencies were classified as outliers if they were out of the range of 2.5 *SD*s from the mean for each participant. With this procedure, 3.70% (71 data points) of the experimental "Word" trials were classified as outliers and, thus, were excluded from the statistical analyses. After excluding the outliers, 3.46% (64 data points) of the remaining experimental "Word" trials were errors and, thus, these trials were excluded from the latency analyses. Mean lexical decision latencies and error rates for the experimental "Word" trials are presented in Table 5. As in Experiment 1, the analyses were conducted on log-transformed lexical decision latencies for the correct experimental trials and errors for the experimental trials using the same intercept and maximal models.

In the latency analyses, neither the effect of P-O consistency, estimated coef. = −0.017, *SE* = 0.013, $t = -1.322$, nor the effect of phonological neighborhood size, estimated coef. = 0.000, *SE* = 0.000, $t = 0.213$, was significant in the intercept model. In the maximal model, neither the effect of P-O consistency, estimated coef. = −0.017, *SE* = 0.013, $t = -1.291$, nor the effect of phonological neighborhood size was significant, estimated coef. = 0.000, *SE* = 0.000, $t = 0.218$. That is, no significant effect was detected in either model.

In the error analyses, neither the effect of P-O consistency, estimated coef. = −0.585, *SE* = 0.984, $z = -.594$, nor the effect of phonological neighborhood size, estimated coef. = −0.004, *SE* = 0.017, $z = -.241$, was significant in the intercept model. Neither the effect of P-O consistency, estimated coef. = −0.416, *SE* = 1.104, $z = -0.377$, nor the effect of phonological neighborhood size, estimated coef. = −0.004, *SE* = 0.017, $z = -0.253$, was significant in the maximal model either.

Combined analyses of Experiments 1 and 2. In order to compare the results from the two experiments, combined analyses were conducted on the log-transformed lexical decision latencies and errors of the experimental trials in Experiments 1 and 2. In these analyses, Experiment (a categorical variable denoting Experiments 1 and 2) and the Experiment × P-O consistency interaction were entered into the models in addition to P-O consistency. Because an effect of phonological neighborhood size was not detected in any of the data analyses described above, we did not include this factor in the combined analyses.

Thus, the "lmer" syntax for the intercept model was "*Y* ~ *Consistency* * *Experiment* + *(1* | *Subject)* + *(1* | *Item).*" In the maximal model, on the other hand, a by-item slope for Experiment

---

[7] In our LME analyses, *p*-values are provided using the "lmerTest" R-package which calculates *p*-values using the degrees of freedom based on the Satterthwaite approximation.

Table 5

Mean Lexical Decision Latencies (RT) in Milliseconds and Error Rates (ER) in Percent for the Two Types of Kanji Words in Experiments 1 and 2

			P-O consistency effect	
Condition	RT (ms)	ER (%)	RT (ms)	ER (%)
Experiment 1 (auditory lexical-decision task)				
More P-O consistent words	798 (6.28)	10.46 (1.16)	+58	+3.16
Less P-O consistent words	856 (7.21)	13.62 (1.31)		
Experiment 2 (visual lexical-decision task with clear stimuli)				
More P-O consistent words	556 (4.01)	3.55 (.61)	0	−.18
Less P-O consistent words	556 (4.05)	3.37 (.60)		

Note. Standard error of the mean is in parenthesis. In Experiment 1, mean lexical decision latency and error rate for the 60 nonwords were 909 ms (SEM = 5.31) and 6.20% (SEM = .58), respectively. In Experiment 2, mean lexical decision latency and error rate for the nonwords were 688 ms (SEM = 6.10) and 8.08% (SEM = .57), respectively.

and a by-subject slope for P-O consistency were assumed because Experiment was a within-item factor and P-O consistency was a within-subject factor. Hence, the "lmer" syntax was "*Y ~ Consistency * Experiment + (1 + Consistency | Subject) + (1 + Experiment | Item)"* for the maximal model. In the error analysis, however, the maximal model failed to converge. Following Barr et al. (2013), therefore, the version of the maximal model we used assumed no random slope-intercept correlations. The "lmer" syntax of this model was "*Y ~ Consistency * Experiment + (0 + Consistency | Subject) + (1 | Subject) + (0 + Experiment | Item) + (1 | Item).*"

In the latency analyses, the effect of P-O consistency, estimated coef. $= -0.038$, $SE = 0.011$, $t = -3.373$, $p = .002$, the effect of Experiment, estimated coef. $= 0.094$, $SE = 0.006$, $t = 15.913$, $p < .001$, and the interaction between P-O consistency and Experiment, estimated coef. $= -0.021$, $SE = 0.004$, $t = -5.216$, $p < .001$, were all significant in the intercept model. Similarly, the effect of P-O consistency, estimated coef. $= -0.039$, $SE = 0.013$, $t = -3.001$, $p = .004$, the effect of Experiment, estimated coef. $= 0.096$, $SE = 0.007$, $t = 12.956$, $p < .001$, and the interaction between the two factors, estimated coef. $= -0.020$, $SE = 0.010$, $t = -2.072$, $p = .044$, were all significant in the maximal model. The significant P-O Consistency × Experiment interaction reflects the fact that a large P-O consistency effect was observed in the auditory task (58 ms) but no effect was observed in the visual task (0 ms).

In the error analyses, the two models produced somewhat different results. In the intercept model, the effect of Experiment, estimated coef. $= 0.960$, $SE = 0.122$, $z = 7.900$, $p < .001$, and the interaction between P-O consistency and Experiment, estimated coef. $= -0.534$, $SE = 0.270$, $z = -1.977$, $p = .048$, were significant, although the effect of P-O consistency was not, estimated coef. $= -1.535$, $SE = 0.973$, $z = -1.577$. In the maximal model with no random correlations, the effect of Experiment was significant, estimated coef. $= 0.876$, $SE = 0.254$, $z = 3.451$, $p = .001$, and the effect of P-O consistency was marginal, estimated coef. $= -1.598$, $SE = 0.960$, $z = -1.665$, $p = .096$. The interaction between the two factors was not significant, estimated coef. $= -0.765$, $SE = 0.599$, $z = -1.278$.

Discussion

Using the same words, a significant P-O consistency effect was observed in our auditory lexical-decision task in Experiment 1 but no such effect emerged in our visual lexical-decision task in Experiment 2. These results are quite consistent with those of Ziegler et al. (2008), in that they also reported a significant P-O consistency effect in their auditory lexical-decision task but not in their visual task. Note as well that the results were also consistent with Peereman, Content, and Bonin (1998) who reported a significant P-O consistency effect in their writing task using auditory stimuli but failed to observe a significant effect in their visual lexical-decision task.

In our first two experiments, we manipulated P-O consistency using a logographic script (i.e., Japanese kanji words) and, as in a number of the previous feedback consistency studies using alphabetic languages, we failed to observe a significant effect in a visual lexical-decision task. We expected that our P-O consistency manipulation for kanji words would be stronger than those in the previous studies using alphabetic languages because kanji characters generally possess a number of homophonic characters. Hence, there was good reason to believe that we would observe a feedback consistency effect in our visual task. That was not the case, however. Instead, our results were quite similar to those in the previous studies (e.g., Peereman et al., 1998; Ziegler et al., 2008). Thus, together with the results from those studies, our results suggest that a feedback consistency effect arises when hearing words but typically not when reading words.

As previously noted, given a lexical structure with direct linkages between orthography and phonology, there are at least two ways to account for the null P-O consistency effect in the visual lexical-decision task in Experiment 2: The first assumes bidirectional activation between orthography and phonology whereas the second abandons that assumption (Ziegler et al., 2008). At the same time, however, other available data do appear to necessitate the assumption of feedback activation from phonology to orthography when reading words (i.e., the homophone disadvantage), suggesting that the first of Ziegler et al.'s (2008) two accounts is likely the correct one.

In order to further examine Ziegler et al.'s (2008) proposals for why phonological feedback effects often do not emerge in visual tasks, in Experiments 3 and 4, we examined the feedback consistency effect in situations in which the quality of the visual stimuli was degraded. Using the kanji words from Experiments 1 and 2, we conducted a perceptual identification task with masked visual stimuli in Experiment 3. In this experiment, each kanji word was briefly presented and was preceded and followed by a mask stimulus. Participants were asked to identify the presented word. In Experiment 4, we conducted a visual lexical-decision task with the luminance of the visual stimuli being degraded. If there were no feedback activation from phonology to orthography in visual tasks, there would be no reason to expect a feedback consistency effect to emerge in Experiments 3 and 4. In contrast, if feedback activation does operate in visual tasks but its impact is difficult to observe when the visual stimulus is clearly available during processing (Ziegler et al., 2008), these experiments would provide a good opportunity to observe a P-O consistency effect. That is, due to the fact that the stimuli in both experiments are only seen in a degraded form, it would be quite difficult to simply clean up their orthographic codes based on information from ongoing visual input. As a result, an impact of P-O consistency may emerge in both experiments.

Experiments 3 and 4

Method

Participants. Seventy-eight undergraduate and graduate students from Waseda University participated in these experiments. Thirty-six participated in Experiment 3 (the perceptual identification task) and the rest participated in Experiment 4 (the visual lexical-decision task with degraded stimuli). Each was paid a small amount of money (500 yen) in exchange for participating. All were native Japanese speakers who had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Stimuli. In Experiment 3, the stimuli were the 60 experimental kanji words used in Experiments 1 and 2 involving 24 more P-O consistent words, 24 less P-O consistent words and 12 fillers. An additional 16 two-character kanji words were also selected and used as stimuli in the practice trials. In Experiment 4, the 120 experimental stimuli (60 kanji words and 60 kanji nonwords) and 16 practice stimuli (8 kanji words and 8 kanji nonwords) were the same as those used in Experiments 1 and 2.

Procedure. In Experiment 3, as in Experiments 1 and 2, participants were tested individually in a normally lit room. In this experiment, the luminance of the stimuli presented on the video monitor was the same as that in Experiment 2 (33.28 cd/m^2). Each trial was initiated with a 50 ms 400 Hz warning tone, after which an array of four number signs (####) appeared at the center of the video monitor as a fixation stimulus, and 1 s later, the fixation stimulus was replaced by a kanji word. The kanji word was presented for 47 ms and immediately replaced by a 50% random dot pattern mask. The pattern mask remained on the video monitor until the participant's key press. Participants were asked to write down the kanji word presented on the video monitor on an answer sheet and to press a key on the response box in order to initiate the next trial.

In Experiment 4, participants were tested individually in a darkened room. In this experiment, the luminance of the stimuli presented on the video monitor was 0.04 cd/m^2 as measured in the same manner as in Experiment 2. In all the other respects, the procedure in Experiment 4 was the same as that in Experiment 2.

Results

Experiment 3 (perceptual identification task). In this experiment, the data were scored in two different ways. One technique involved scoring the number of correctly reported kanji characters on each trial. Because all the stimuli were two-character kanji compounds, a 2 was given when the compound was correctly reported, a 1 was given if one of the characters was correctly reported, and a 0 was given otherwise (either incorrect or no answer). The other technique involved scoring the trial as 1 if the stimulus was reported correctly and 0 if not. Mean numbers of correct characters and correct word reports are presented in Table 6. Because the numbers of correct characters are count data, those values were square-root-transformed and analyzed using the LME models with P-O consistency and phonological neighborhood size as fixed factors as in the previous experiments. The (untransformed) correct word reports were analyzed in the same fashion but using the logit mixed models.

In the analyses of the numbers of correct characters, the effect of P-O consistency was significant, estimated coef. = 0.198, SE = 0.073, t = 2.706, p = .010, although the effect of phonological neighborhood size was not, estimated coef. = 0.000, SE = 0.001, t = 0.097, in the intercept model. Similarly, the effect of P-O consistency was significant, estimated coef. = 0.198, SE = 0.075, t = 2.630, p = .012, although the effect of phonological neighborhood size was not, estimated coef. = 0.000, SE = 0.001, t = 0.097, in the maximal model.

In the analyses of correct word reports, the effect of P-O consistency was significant, estimated coef. = 1.147, SE = 0.462, z = 2.483, p = .013, although the effect of phonological neighborhood size was not, estimated coef. = 0.003, SE = 0.008, z = 0.378, in the intercept model. Similarly in the maximal model, the effect of P-O consistency was significant, estimated coef. = 1.265, SE = 0.495, z = 2.556, p = .011, but the effect of phonological neighborhood size was not, estimated coef. = 0.003 SE = 0.008, z = 0.400.

Experiment 4 (visual lexical-decision task with degraded stimuli). As in Experiments 1 and 2, lexical decision latencies were classified as outliers if they were out of the range of 2.5 *SD*s from the mean for each participant. With this procedure, 2.93% (59 data points) of the experimental "word" trials were classified as

Table 6
Mean Numbers of Correct Characters (NCC) and Correct Word Reporting Rates (CR) in Percent for the Two Types of Kanji Words in the Perceptual Identification Task in Experiment 3

			P-O consistency effect	
Condition	NCC	CR (%)	NCC	CR (%)
More P-O consistent words	1.64 (.02)	74.54 (1.48)	+.15	+7.87
Less P-O consistent words	1.49 (.03)	66.67 (1.60)		

outliers and excluded from the statistical analyses. After excluding the outliers, 3.07% (60 data points) of the experimental "word" trials were errors and, thus, these trials were excluded from the latency analysis. Mean lexical decision latencies and error rates for the experimental trials are presented in Table 7. As in Experiments 1 and 2, log-transformed lexical decision latencies were analyzed using the same LME models and errors were also analyzed using the same logit mixed models.

In the latency analyses, the effect of P-O consistency was significant, estimated coef. = −0.037, *SE* = 0.018, $t = -2.036$, $p < .001$, although the effect of phonological neighborhood size was not, estimated coef. = −0.000, *SE* = 0.000, $t = -0.027$, in the intercept model. Similarly, while the effect of P-O consistency was significant, estimated coef. = −0.037, *SE* = 0.018, $t = -2.039$, $p = .048$, the effect of phonological neighborhood size was not, estimated coef. = −0.000, *SE* = 0.000, $t = -0.026$, in the maximal model.

In the error analysis, neither the effect of P-O consistency, estimated coef. = −0.954, *SE* = 0.860, $z = -1.110$, nor the effect of phonological neighborhood size, estimated coef. = −0.012, *SE* = 0.015, $z = -0.800$, was significant in the intercept model. Also in the maximal model, neither the effect of P-O consistency, estimated coef. = −0.720, *SE* = 0.932, $z = -0.773$, nor the effect of phonological neighborhood size, estimated coef. = −0.012, *SE* = 0.015, $z = -0.805$, was significant.

Combined analyses of Experiments 2 and 4. In order to directly contrast visual lexical decision performance for clear and degraded stimuli, combined analyses were further conducted for log-transformed lexical decision latencies and errors for the experimental trials in Experiments 2 and 4. The procedures for these analyses were the same as those used in the combined analyses of Experiments 1 and 2.

In the latency analyses, the effect of P-O consistency, estimated coef. = −0.028, *SE* = 0.013, $t = -2.183$, $p = .034$, the effect of Experiment, estimated coef. = 0.126, *SE* = 0.008, $t = 15.979$, $p <$ .001, and the P-O Consistency × Experiment interaction, estimated coef. = −0.009, *SE* = 0.004, $t = -1.982$, $p = .048$, were all significant in the intercept model. In the maximal model, however, while the effect of P-O consistency, estimated coef. = −0.028, *SE* = 0.013, $t = -2.203$, $p = .033$, and the effect of Experiment, estimated coef. = 0.126, *SE* = 0.008, $t = 14.944$, $p < .001$, were significant, the P-O Consistency × Experiment interaction was not significant, estimated coef. = −0.009, *SE* = 0.007, $t = -1.389$. The significant P-O Consistency × Experiment interaction that was observed in the intercept model reflects the size difference between the P-O consistency effect in the task with degraded stimuli (a 61 ms effect) and the effect in the task with clear stimuli (a 0 ms effect).

In the error analyses, neither the effect of P-O consistency, estimated coef. = −0.537, *SE* = 0.631, $z = -0.852$, nor the effect of Experiment, estimated coef. = −0.037, *SE* = 0.159, $z = -0.235$, nor the interaction between the two factors, estimated coef. = −0.133, *SE* = 0.307, $z = -0.431$, was significant in the intercept model. Similarly, neither the effect of P-O consistency, estimated coef. = −0.542, *SE* = 0.640, $z = -0.846$, nor the effect of Experiment, estimated coef. = −0.019, *SE* = 0.207, $z = -0.094$, nor the interaction between P-O consistency and Experiment, estimated coef. = −0.135, *SE* = 0.431, $z = -0.314$, was significant in the maximal model with no random correlations (this model was used because the maximal model failed to converge).

Table 7
Mean Lexical Decision Latencies (RT) in Milliseconds and Error Rates (ER) in Percent for the Two Types of Kanji Words in the Visual Lexical-Decision Task With Degraded Stimuli in Experiment 4

Condition	RT (ms)	ER (%)	P-O consistency effect RT (ms)	P-O consistency effect ER (%)
More P-O consistent words	977 (9.87)	2.65 (.51)	+61	+.83
Less P-O consistent words	1038 (13.79)	3.48 (.59)		

Note. Standard error of the mean is in parenthesis. Mean lexical decision latency and error rate for the 60 nonwords were 1,319 ms (SEM = 15.36) and 4.00% (SEM = .40), respectively.

Discussion

A significant feedback consistency effect was observed in both the perceptual identification task (Experiment 3) and the visual lexical-decision task with degraded stimuli (Experiment 4). These results clearly indicate that feedback activation from phonology to orthography does affect task performance in visual tasks although potentially only when the visual stimuli are degraded. As noted, the reasoning is that, when visual stimuli are degraded, it becomes more difficult to obtain clarifying orthographic information directly from those stimuli. Hence, assuming that there is automatic phonological activation for visually presented words (i.e., feedforward activation) followed by feedback activation to the orthographic level, it is more difficult for participants to clarify any conflicts created at the orthographic level than in a conventional visual lexical-decision task. As a result, a feedback consistency effect emerges in tasks when the visual stimuli are degraded (due to masking as in Experiment 3 or due to luminance reduction as in Experiment 4).

One should also note that our results are somewhat inconsistent with the results reported by Peereman et al. (1998). Using the same consistent and inconsistent words that produced a significant P-O consistency effect in their writing task, Peereman et al. (1998) failed to observe a consistency effect in their visual lexical-decision task even when the stimuli were masked. That is, whereas their Experiment 1a was a conventional visual lexical-decision task, in their Experiment 2, the lexical-decision task involved the presentation of a backward mask following the briefly presented stimulus (47 ms on average). No P-O consistency effect was detected in either experiment.

Why are their results different from ours? In their tasks, the error rates were higher in the masked task (20.2% on average) than in the conventional task (11.4% on average) but their mean lexical decision latencies were quite similar in the two tasks (688 ms and 677 ms on average in the conventional and masked tasks, respectively). In contrast, in our tasks with clear and degraded stimuli (in Experiments 2 and 4), lexical decision latencies were much longer in the task with degraded stimuli (1,008 ms on average) than in the task with clear stimuli (556 ms on average), as reflected by the significant effect of Experiment in the combined analyses of lexical decision latencies (in both statistical models), although the error rates were similar in the two tasks (3.46% and 3.07% on average in the tasks with clear and degraded stimuli, respectively), as reflected by the lack of an effect of Experiment in the combined analyses of errors (in both statistical models). These results suggest that the impact of our degradation manipulation (i.e., an approximately 450 ms increase in lexical decision latencies) was much stronger than the degradation manipulation (i.e., an approximately 9% increase in error rates) employed by Peereman et al. (1998). That is, our participants likely had considerably more difficulty using the orthographic information directly available from the stimulus to resolve the competition created at the orthographic level in our Experiment 4 than did Peereman et al.'s (1998) participants in their Experiment 2. In fact, it is quite possible that Peereman et al.'s (1998) participants often had a reasonably good orthographic representation of the presented stimuli available to them, allowing for a fairly rapid clean-up of the orthographic codes. Therefore, it is, perhaps, not surprising that their data showed little effect of feedback activation from phonology in their degraded condition.

General Discussion

In order to examine the impact of P-O relationships for words, we began by evaluating the P-O consistencies for the 339 katakana and 775 kanji words examined by Hino, Miyamura, et al. (2011). Hino, Miyamura, et al.'s (2011) results had indicated that the consistencies of both O-P and O-S relationships were comparable for kana and kanji words. In contrast, the present results clearly indicated that the P-O relationships were reasonably consistent for kana words but not for kanji words. Therefore, the present research focused on kanji words.

The Nature of Feedback (P-O) Relationships for Kana and Kanji Words

An obvious question is why the P-O relationships were more inconsistent for kanji words than for kana words. As previously noted, the most reasonable answer is because there are a large number of homophonic kanji characters, which means that there would be more chance for kanji words to have phonological neighbors involving different kanji characters (i.e., orthographic enemies). For katakana words, on the other hand, many of the words in their phonological neighborhoods would likely be katakana words, most of which would be orthographic friends. As a result, P-O relationships should be more consistent for katakana words than for kanji words, as was found.

One issue that we also needed to examine was the possibility that the reason for the highly inconsistent P-O relationships for kanji words was the specific nature of character-to-mora correspondences for kanji words coupled with the way we defined P-O consistency. While each kana character basically corresponds to a single mora (ポケット, [pocket,/po.ke.Q.to/]), kanji characters often correspond to multiple-mora pronunciations (e.g., 確実 [certain,/ka.ku.zi.tu/]). For words of this sort, it would be impossible, according to our definitions, to have orthographic friends in their phonological neighborhood. As previously noted, for a two-character kanji word with four morae, when each character corresponds to two morae, it is impossible to find an orthographic friend in the phonological neighborhood because there is no phonological neighbor in which the shared three morae are printed using the same characters as those in the target (e.g., 確実 [certain,/ka.*ku.zi.tu*/] and 着実 [steady,/tja.*ku.zi.tu*/]). Thus, the computed P-O consistency values would inevitably be smaller for kanji words than for katakana words using our original definitions. Our analysis indicates, however, that this issue only accounted for a small part of the more inconsistent P-O relationships for kanji words than for katakana words. That is, even when we classified all the orthographic neighbors in the phonological neighborhoods as orthographic friends, the mean P-O consistency for the kanji words (.44) was still significantly lower than that for the katakana words (.62). Our conclusion, therefore, is that the fact that there are many homophonic characters for kanji words has led to the P-O relationships being more inconsistent for kanji words than for kana words.

P-O Consistency Effects

Although there are a number of studies demonstrating that auditory word recognition is affected by the nature of a word's P-O consistency (e.g., Peereman et al., 1998; Ziegler & Ferrand, 1998; Ziegler et al., 2008), it was much less clear whether visual word recognition is also affected by a word's P-O consistency because there are quite a few studies reporting a null P-O consistency effect in the visual domain (e.g., Peereman et al., 1998; Ziegler et al., 2008, a "feedback consistency" effect). For example, Ziegler et al. (2008) examined the P-O consistency effect using both visual and auditory lexical-decision tasks. Although a significant P-O consistency effect was detected in their auditory task, no P-O consistency effect emerged in their visual task. In order to explain these results, Ziegler et al. (2008) suggested two accounts: one maintaining the assumption of bidirectional feedback activation across orthography and phonology, and the other abandoning that assumption.

According to Ziegler et al.'s (2008) first account, even if bidirectional feedback activation across orthography and phonology does occur, a P-O consistency effect would be more difficult to detect in a visual lexical-decision task with clear stimuli because orthographic codes could be rapidly cleaned up through the continuing visual input and, hence, any effect due to phonological feedback would be minimal in that situation. In contrast, in the auditory lexical-decision task, competing orthographic information activated by the feedforward activation from phonology to orthography when words are presented aurally would have to be resolved through an orthographic analysis in order to make an accurate decision, an analysis that could not be based on a currently available stimulus. Therefore, a P-O consistency effect would be much more likely to emerge in an auditory task.

In Ziegler et al.'s (2008) second account, the assumption of bidirectional feedback activation between orthography and phonology was simply abandoned. That is, the reason there was no feedback consistency effect in the visual task was because there was no P-O feedback while the P-O consistency effect for auditory stimuli would be merely a feedforward consistency effect.

The former of these accounts would seem to be more likely, a priori, however, because there is independent evidence suggesting bidirectional activation across orthography and phonology when reading words. As previously noted, a homophony disadvantage has been reported in a number of studies when homophones possess only a single homophonic mate (e.g., Edwards et al., 2004; Ferrand & Grainger, 2003; Hino et al., 2013; Kerswell et al., 2007; Pexman & Lupker, 1999; Pexman et al., 2001; Pexman et al., 2002; Rubenstein et al., 1971) and that effect appears to be due to the inconsistent P-O feedback for the homophones. Further, in order to account for the interaction between orthographic and phonological neighborhood sizes in their visual and auditory tasks, Grainger and colleagues (Grainger et al., 2005; Grainger & Ziegler, 2007) have also argued that there must be bidirectional activation across orthography and phonology (see also Hino, Nakayama, et al., 2011, for similar results using Japanese katakana words).

In order to discriminate between these ideas, we conducted four experiments in which P-O consistency was manipulated for kanji words. The first two experiments involved auditory and visual lexical-decision tasks. Because kanji characters can possess multiple homophonic mates, the P-O consistency manipulation for kanji words should be quite strong, potentially stronger than the P-O consistency manipulation in the previous studies using alphabetic languages. If so, and if there is bidirectional activation across orthography and phonology, the supposition was that we may be able to observe a P-O consistency effect not only in the auditory task but also in the visual task. Nonetheless, our results were consistent with those reported by Ziegler et al. (2008): A significant P-O consistency effect was observed only in the auditory task.

In order to further examine the question of whether bidirectional activation arises when reading words, therefore, we conducted a perceptual identification task with masked visual stimuli (in Experiment 3) and a visual lexical-decision task with degraded stimuli using luminance reduction (in Experiment 4). If there is no bidirectional activation across orthography and phonology, there would be no reason to expect a feedback consistency effect even for the degraded visual stimuli. Assuming bidirectional activation across orthography and phonology, on the other hand, a feedback consistency effect would be expected to emerge for the degraded stimuli because it would be difficult to clean up the orthographic code using the visual input. That is, phonological feedback would create competition at the orthographic level for the inconsistent words and, hence, a feedback consistency effect should emerge for the degraded stimuli.

The results in Experiments 3 and 4 were consistent with these predictions as a significant feedback consistency effect was observed in both experiments. Our overall results, therefore, support the assumption of bidirectional activation across orthography and phonology. That is, in the auditory task, multiple orthographic codes would be activated for the less consistent words, creating competition that can only be resolved through extended orthographic processing. As a result, a P-O consistency effect will typically arise in an auditory task, as was observed in Experiment 1. In the visual task with clear stimuli, however, the phonological feedback will typically play little role because the clear visual stimuli would be able to provide enough information to readily resolve any orthographic competition. In visual tasks with degraded stimuli, on the other hand, it is harder to obtain orthographic information directly from the visual stimuli. In such a circumstance, a feedback consistency effect would be expected, consistent with the results in Experiments 3 and 4.

Phonological Neighborhood Size Effects?

Although we successfully observed a P-O consistency effect in Experiments 1, 3, and 4, one thing to note is that our P-O consistency manipulation was confounded with phonological neighborhood size. That is, as shown in Table 4, phonological neighborhood size was larger for the less P-O consistent words (34.96) than for the more P-O consistent words (20.29). As described earlier, Grainger et al. (2005) explained why a confound of this sort is virtually inevitable when one is manipulating P-O consistency for word groups with similar orthographic neighborhood sizes. Nonetheless, due to the fact that larger phonological neighborhoods do produce longer latencies in the auditory lexical-decision task (e.g., Vitevitch & Luce, 1999; Ziegler, Muneaux, & Grainger, 2003), this situation does raise the possibility that at least some part of our P-O consistency effects was due to the difference in phonological neighborhood sizes across word groups. In order to address this potential issue, we entered phonological neighborhood size as a control variable in our LME analyses in our experiments. In these analyses, no significant effect of phonological neighborhood size was detected in any of our experiments, indicating clearly that the consistency effects we observed in Experiments 1, 3, and 4 were not due to phonological neighborhood size.

P-O Consistency Effects for Ambiguous Stimuli

In our experiments, a P-O consistency effect emerged for both auditory and degraded visual stimuli. We have characterized our results by saying that the effect emerges whenever the visual evidence available during processing is not sufficient to overcome the competition created by feedback at the orthographic level. In contrast, one could characterize our results by noting that the P-O consistency effect in our experiments only arose when the stimuli were presented in a less than fully clear fashion (due to the use of either auditory stimuli or degraded visual stimuli). One could further suggest that, as a result of not having a clear stimulus available, our participants may have brought some processing (in particular, guessing) strategy to bear on at least a subset of trials and it is that strategy that produced the benefit for the more P-O consistent words.

If such a strategy did operate in the auditory lexical-decision task of Experiment 1, task performance likely would have been modulated strongly by the number of phonological neighbors. That is, a target would have been more difficult to isolate/identify/guess successfully if it has more phonologically similar words and, hence, lexical decision responses would have been slower and less accurate for words with larger neighborhood sizes. In contrast with this expectation, there was no effect of phonological neighborhood size in Experiment 1. Further, in Experiments 3 and 4, it is unclear how any guessing strategy would have had an impact on the P-O consistency manipulation. In those experiments, the impoverished stimulus that the system had to deal with (and guess from) was visual. Hence, orthographic neighborhoods would be most relevant and those were both large and equal across the two word groups (49.71 and 52.67 for the more and less P-O consistent words, respectively). Because the mean number of orthographic neighbors was equal for the words in the two groups, there would be no reason to expect an advantage for more P-O consistent words. One could argue, of course, that at least some phonological activation may also have arisen based on the extracted partial spelling information in these tasks and, hence, a target would have been more difficult to identify if it had more phonologically similar words (as proposed just above when discussing potential strategy use in Experiment 1). Note, however, that the results of the LME analyses in Experiments 3 and 4 clearly indicated that the task performance was not modulated by phonological neighborhood size in those experiments either. Everything considered, therefore, it seems unlikely that a guessing strategy based on partial information could explain the P-O consistency effects observed in our experiments.

A second alternative account of our results would be based on the idea that a different type of processing is implemented as the stimuli become harder to process. There are, for example, some studies reporting orthographic/lexical/semantic effects in speech processing tasks that only emerge when stimuli are degraded (e.g., Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005; Obleser, Wise, Dresner, & Scott, 2007; Pattamadilok, De Morais, & Kolinsky, 2011). In particular, using a shadowing task, Pattamadilok, De Morais, and Kolinsky (2011) reported a significant P-O consistency effect when auditory stimuli were presented with noise, although no effect of consistency emerged when the stimuli were presented clearly (i.e., without noise). Based on these results, Pattamadilok et al. (2011) suggested that the cognitive system adjusts itself by employing additional processing resources when the task becomes more difficult (in order to maintain a good level of performance). When clear stimuli are presented, it may be possible to maintain a good level of performance by simply employing normal processes. That is, it may be possible to accurately perceive the acoustic signals and to produce shadowing responses without there being extensive interactions within the processing system. When the stimuli are degraded, however, making the task more difficult, the cognitive system may attempt to maintain a good level of performance by recruiting additional processing involving orthographic-phonological interactions. As a result, orthographic knowledge would affect shadowing performance in the more difficult noise condition, leading to a P-O consistency effect.

In our experiments with visual stimuli, we observed a significant P-O consistency effect for degraded stimuli (in Experiments 3 and 4) but not for clear stimuli (in Experiments 2). Thus, in line with Pattamadilok et al.'s (2011) argument, one could suggest that a significant consistency effect was observed only for degraded visual stimuli in our experiments because additional processing resources were employed for the degraded stimuli in order to maintain a good level of performance.[8]

Pattamadilok et al.'s (2011) account essentially differs from ours in that it regards the P-O interactions as arising automatically to aid processing only when the stimuli are difficult to process (see also Price & Devlin, 2011, for a similar account). A similar claim has also been advanced recently by Barnhart and Goldinger (2010) based on a slightly different manipulation. Using Stone et al.'s (1997) stimuli, Barnhart and Goldinger (2010) examined feedback consistency effects for handwritten stimuli as well as stimuli with a normal type font in their lexical decision and naming experiments. In all their experiments, there was a P-O feedback consistency effect with the effects in the lexical-decision task being similar in size but with the effect being larger for the handwritten stimuli than for the stimuli with a normal font in the naming task. Because the false alarm rate (incorrect "word" decisions to nonword stimuli) was quite high for handwritten stimuli in their lexical decision experiments, Barnhart and Goldinger (2010) suggested that the effect size comparisons were somewhat compromised in their lexical decision experiments, focusing their discussion instead on their naming task. Based on the increased feedback consistency effect in their naming experiments with handwritten stimuli, Barnhart and Goldinger (2010) concluded that the use of handwritten stimuli increased the feedback consistency effect due to the fact that the orthographic codes obtainable from the presented stimuli are less clear when one is reading handwriting. Hence, phonological feedback played a larger role in cleaning up the orthographic code, producing a larger P-O consistency effect.

Essentially, the issue is that there are two ways to think about the timing/creation of phonological feedback. One would be to assume that the orthographic-phonological interaction occurs automatically during an early stage of processing as suggested by Ziegler et al. (2008), and as proposed here, with the effect of phonological feedback being more likely to have an impact (i.e., to produce a P-O consistency effect) when the stimuli are degraded and/or ambiguous. Hence, whether there is a feedback consistency effect or not is a matter of the degree to which the participant can take advantage of external information in identifying the stimulus. Alternatively, one could also assume, as done by Pattamadilok et al. (2011), that when the stimuli are more difficult to process, orthographic-phonological interactions are recruited in order to maintain a good level of task performance.

One thing to note about the present results is that lexical decision latencies were significantly slower in the auditory task (in Experiment 1) than in the visual task (in Experiment 2) and a significant consistency effect was observed only in the auditory task. Similarly, comparing Experiments 2 and 4, there was also a tendency for lexical decision responses to be slower in the visual task with degraded stimuli (in Experiment 4) than in the visual task with clear stimuli (in Experiment 2) and a significant consistency effect was observed only in the task with degraded stimuli. As such, consistent with Pattamadilok et al.'s (2011) argument, it was the case that significant P-O consistency effects emerged only when the stimuli were more difficult (and, hence, took more time) to process in our lexical-decision tasks. Following Pattamadilok et al.'s (2011) logic, therefore, one could argue that the orthographic-phonological interaction operated in Experiments 1 and 4 in order to maintain a good level of task performance for the difficult stimuli. As a result, a significant P-O consistency effect emerged in Experiments 1 and 4 but not in Experiment 2.

In general, based on our data, it is not really possible to discriminate between these alternatives. We think, however, that the former interpretation is preferable because it is unclear whether homophone effects in visual lexical-decision tasks (e.g., Chen et al., 2009; Edwards et al., 2004; Ferrand & Grainger, 2003; Hino et al., 2013; Kerswell et al., 2007; Pexman & Lupker, 1999; Pexman et al., 2001; Pexman et al., 2002; Rubenstein et al., 1971; Ziegler et al., 2000) could be accounted for by assuming that the orthographic-phonological interactions operate only when the stimuli are difficult to process. Needless to say, more research is needed to discriminate between these alternatives.

---

[8] We thank one of our anonymous reviewers for bringing this type of account to our attention.

Word Recognition Models With Orthographic-Phonological Interaction

Regardless of whether the orthographic-phonological interaction arises in an early or a late stage of processing, our results suggest that any type of model of word recognition would need to incorporate the assumption of orthographic-phonological interactions. Indeed, there are attempts to model these types of interactions already in the literature. For example, Grainger and colleagues (e.g., Grainger et al., 2005; Grainger & Ziegler, 2007) have implemented orthographic-phonological interactions in their bimodal interactive-activation model in an attempt to account for the interaction between orthographic and phonological neighborhood sizes found in their visual and auditory tasks (e.g., Grainger et al., 2005; Ziegler, Muneaux, & Grainger, 2003). In addition, models based on connectionist frameworks which incorporate interactions between orthography and phonology have been with us for a number of years now (e.g., Stone et al., 1997; Ziegler et al., 1997). In these models, when a visual stimulus is presented, orthographic units would be first activated and send feedforward signals to phonological units. The activated phonological units, then, send feedback signals to the orthographic units and the lexical network gradually moves into a stable state (i.e., a basin of attraction), which represents a learned pattern of orthographic-phonological association. As such, these types of models would seem to be able to explain the nature of relationships between orthography and phonology and how they play an important role when reading and hearing words.

What is missing from these models, however, is an explanation of why there is no P-O consistency effect in lexical-decision tasks for visual stimuli and in shadowing tasks for auditory stimuli when the stimuli are clear (i.e., the data in our Experiment 2 as well as those in Pattamadilok et al., 2011; Peereman et al., 1998; and Ziegler et al., 2008). If one adopts the assumption that the orthographic-phonological interactions arise automatically early in processing, the obvious explanation for any null effect would be that orthographic/phonological codes are rapidly cleaned up using the information derived from the stimuli when the stimuli are sufficiently clear in either the visual or auditory domain. In contrast, if one were to assume that the interaction arises when encountering a difficult-to-process stimulus, one may need to consider some mechanism to disable the orthographic-phonological interaction when the input stimuli are clear and to enable it when the input stimuli are degraded and ambiguous, allowing phonological feedback to become available. Either way, word recognition models would need to incorporate some additional assumption/mechanism in order to be able to explain the typical lack of a P-O consistency effect in most conventional lexical-decision tasks with visual stimuli as well as in Pattamadilok et al.'s (2011) shadowing tasks with auditory stimuli.

Potential Processing Differences for Kana and Kanji Words

Given that the P-O relationships are generally more inconsistent for kanji words than for kana words, our results suggest that processing kanji words may become selectively more difficult under conditions where phonological feedback can play a larger role. For example, assuming that literate listeners tend to retrieve orthographic codes in order to resolve the ambiguity of auditory stimuli in auditory word recognition, the feedforward activation from phonology to orthography may make this task more difficult for kanji words than for kana words.

In the visual domain, when visual stimuli are clearly presented, this P-O consistency difference would typically not create a processing difference for kana and kanji words. For people with reduced visual acuity, however, it may become more difficult to obtain enough information to clean up the orthographic codes through the visual input. In such a situation, phonological feedback may make orthographic processing more difficult for kanji words than for kana words. Similarly, under conditions where obtaining visual information is more difficult for normal readers such as when one tries to read words in a darkened room or when one tries to read words printed in a very small font, kanji word reading may be more strongly impacted due to kanji words' inconsistent P-O relationships. Further research is needed to evaluate these ideas as well as to more clearly understand the nature of processing required when reading Japanese kana versus kanji words.

Conclusions

Using the same katakana and kanji words examined by Hino, Miyamura, et al. (2011), we initially evaluated the degree of P-O consistency for those words. Our results indicated that P-O relationships were more inconsistent for the kanji words than for the katakana words. We then investigated the possibility that reading and hearing processes for kanji words may be strongly affected by the consistency of their P-O relationships. The results of our experiments using kanji words suggested that whereas visual word recognition is typically not affected by phonological feedback when the visual stimuli are clear (as in Experiment 2), the effect of P-O consistency can be witnessed when the visual stimuli are degraded (by masking or luminance reduction as in Experiments 3 and 4) or when auditory stimuli are used (as in Experiment 1). Based on these results, our basic conclusion is that bidirectional activation across orthography and phonology arises in both auditory and visual word recognition.

References

Amano, N., & Kondo, K. (2003a). *NTT Detabesu Shirizu: Nihongo No Goi-tokusei Dai 1-ki CD-ROM-ban* [NTT database series: Lexical Properties of Japanese, Vol. 1, CD-ROM Version]. Tokyo, Japan: Sanseido.

Amano, N., & Kondo, K. (2003b). *NTT Detabesu Shirizu: Nihongo No Goi-tokusei Dai 2-ki CD-ROM-ban* [NTT database series: Lexical Properties of Japanese, Vol. 2, CD-ROM Version]. Tokyo, Japan: Sanseido.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. http://dx.doi.org/10.1016/j.jml.2007.12.005

Barnhart, A. S., & Goldinger, S. D. (2010). Interpreting chicken-scratch: Lexical access for handwritten words. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 906–923. http://dx.doi.org/10.1037/a0019258

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001

Besner, D., & Hildebrandt, N. (1987). Orthographic and phonological codes in the oral reading of Japanese kana. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 13, 335–343. http://dx.doi.org/10.1037/0278-7393.13.2.335

Chen, H.-C., Vaid, J., & Wu, J.-T. (2009). Homophone density and phonological frequency in Chinese word recognition. *Language and Cognitive Processes, 24,* 967–982. http://dx.doi.org/10.1080/01690960902804515

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). New York, NY: Academic Press.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256. http://dx.doi.org/10.1037/0033-295X.108.1.204

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General, 134,* 222–241. http://dx.doi.org/10.1037/0096-3445.134.2.222

Edwards, J. D., Pexman, P. M., & Hudson, C. E. (2004). Exploring the dynamics of the visual word recognition system: Homophone effects in LDT and naming. *Language and Cognitive Processes, 19,* 503–532. http://dx.doi.org/10.1080/01690960344000215

Feldman, L. B., & Turvey, M. T. (1980). Words written in Kana are named faster than the same words written in Kanji. *Language and Speech, 23,* 141–147.

Ferrand, L., & Grainger, J. (2003). Homophone interference effects in visual word recognition. *The Quarterly Journal of Experimental Psychology, 56,* 403–419. http://dx.doi.org/10.1080/02724980244000422

Frost, R. (2005). Orthographic systems and skilled word recognition processes in reading. In M. J. Snowling & C. Hulmes (Eds.), *The science of reading: A handbook* (pp. 272–295). Malden, MA: Blackwell Publishing. http://dx.doi.org/10.1002/9780470757642.ch15

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 104–115. http://dx.doi.org/10.1037/0096-1523.13.1.104

Fushimi, T., Ijuin, M., Patterson, K., & Tatsumi, I. (1999). Consistency, frequency, and lexicality effects in naming Japanese Kanji. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 382–407. http://dx.doi.org/10.1037/0096-1523.25.2.382

Grainger, J., Muneaux, M., Farioli, F., & Ziegler, J. C. (2005). Effects of phonological and orthographic neighbourhood density interact in visual word recognition. *The Quarterly Journal of Experimental Psychology, 58,* 981–998. http://dx.doi.org/10.1080/02724980443000386

Grainger, J., & Ziegler, J. C. (2007). Cross-code consistency effects in visual word recognition. In E. L. Grigorenko & A. Naples (Eds.), *Single-word reading: Biological and behavioral perspectives* (pp. 129–157). Mahwah, NJ: Erlbaum.

Hino, Y., Kusunose, Y., Lupker, S. J., & Jared, D. (2013). The processing advantage and disadvantage for homophones in lexical decision tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 529–551. http://dx.doi.org/10.1037/a0029122

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1331–1356. http://dx.doi.org/10.1037/0096-1523.22.6.1331

Hino, Y., & Lupker, S. J. (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,* 1431–1453. http://dx.doi.org/10.1037/0096-1523.24.5.1431

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 686–713.

Hino, Y., Lupker, S. J., Sears, C. R., & Ogawa, T. (1998). The effects of polysemy for Japanese katakana words. *Reading and Writing: An Interdisciplinary Journal, 10,* 395–424. http://dx.doi.org/10.1023/A:1008060924384

Hino, Y., Miyamura, S., & Lupker, S. J. (2011). The nature of orthographic-phonological and orthographic-semantic relationships for Japanese kana and kanji words. *Behavior Research Methods, 43,* 1110–1151. http://dx.doi.org/10.3758/s13428-011-0101-0

Hino, Y., Nakayama, M., Miyamura, S., & Kusunose, Y. (2011). Goi-handan-kadai niokeru katakana-go no keitai/onin-rinsetsugosuu-kouka [Orthoraphic and phonological neighborhood size effects for Japanese Katakana words in a lexical decision task]. *Japanese Journal of Psychology, 81,* 576–596.

Jaeger, T. F. (2008). Categorization data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434–446. http://dx.doi.org/10.1016/j.jml.2007.11.007

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687–715. http://dx.doi.org/10.1016/0749-596X(90)90044-Z

Kerswell, L., Siakaluk, P. D., Pexman, P. M., Sears, C. R., & Owen, W. J. (2007). Homophone effects in visual word recognition depend on homophone type and task demands. *Canadian Journal of Experimental Psychology, 61,* 322–327. http://dx.doi.org/10.1037/cjep2007032

Kimura, Y. (1984). Concurrent vocal interference: Its effects on Kana and Kanji. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 36,* 117–131. http://dx.doi.org/10.1080/14640748408401506

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 876–895. http://dx.doi.org/10.1037/a0013484

Kureta, Y., Fushimi, T., & Tatsumi, I. F. (2006). The functional unit in phonological encoding: Evidence for moraic representation in native Japanese speakers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1102–1119.

Lacruz, I., & Folk, J. (2004). Feedforward and feedback consistency effects for high- and low-frequency words in lexical decision and naming. *The Quarterly Journal of Experimental Psychology, 57,* 1261–1284. http://dx.doi.org/10.1080/02724980343000756

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

National Language Research Institute. (1970). *Denshi-keisanki niyoru shinbun no goichousa* [Studies on the vocabulary of modern newspapers (Vol. 1), General descriptions and vocabulary frequency tables]. Tokyo, Japan: Shuei Shuppan.

National Language Research Institute. (1993). *Bunrui-goi hyou (Furoppi Ban)* [Thesaurus (floppy disk version)]. Tokyo, Japan: Shuei Shuppan.

Obleser, J., Wise, R. J. S., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *The Journal of Neuroscience, 27,* 2283–2289. http://dx.doi.org/10.1523/JNEUROSCI.4663-06.2007

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language, 32,* 258–278. http://dx.doi.org/10.1006/jmla.1993.1014

Pattamadilok, C., De Morais, J. J., & Kolinsky, R. (2011). Naming in noise: The contribution of orthographic knowledge to speech repetition. *Frontiers in Psychology, 2,* 1–12.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition.

Journal of Memory and Language, 39, 151–174. http://dx.doi.org/10.1006/jmla.1998.2573

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent reading: Evidence from backward masking. *Journal of Memory and Language, 27,* 59–70. http://dx.doi.org/10.1016/0749-596X(88)90048-4

Perry, C. (2003). A phoneme-grapheme feedback consistency effect. *Psychonomic Bulletin & Review, 10,* 392–397. http://dx.doi.org/10.3758/BF03196497

Pexman, P. M., & Lupker, S. J. (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects? *Canadian Journal of Experimental Psychology, 53,* 323–334. http://dx.doi.org/10.1037/h0087320

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 139–156.

Pexman, P. M., Lupker, S. J., & Reggin, L. D. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 572–584. http://dx.doi.org/10.1037/0278-7393.28.3.572

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115. http://dx.doi.org/10.1037/0033-295X.103.1.56

Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences, 15,* 246–253. http://dx.doi.org/10.1016/j.tics.2011.04.001

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645–657. http://dx.doi.org/10.1016/S0022-5371(71)80071-3

Saito, H. (1981). Kanji to Kana no yomi niokeru keitaitekifugouka oyobi onintekifugouka no kentou [Use of graphemic and phonemic encoding in reading Kanji and Kana]. *Japanese Journal of Psychology, 52,* 266–273. http://dx.doi.org/10.4992/jjpsy.52.266

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568. http://dx.doi.org/10.1037/0033-295X.96.4.523

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404. http://dx.doi.org/10.1016/S0022-5371(84)90270-6

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36,* 337–359. http://dx.doi.org/10.1006/jmla.1996.2487

Tamaoka, K., Kirsner, K., Yanase, Y., Miyaoka, Y., & Kawakami, M. (2002). A Web-accessible database of characteristics of the 1,945 basic Japanese kanji. *Behavior Research Methods, Instruments & Computers, 34,* 260–275. http://dx.doi.org/10.3758/BF03195454

Tamaoka, K., & Makioka, S. (2004a). Frequency of occurrence for units of phonemes, morae, and syllables appearing in a lexical corpus of a Japanese newspaper. *Behavior Research Methods, Instruments, & Computers, 36,* 531–547. http://dx.doi.org/10.3758/BF03195600

Tamaoka, K., & Makioka, S. (2004b). New figures for a web-accessible database of the 1,945 basic Japanese kanji, 4th ed. *Behavior Research Methods, Instruments, & Computers, 36,* 548–558.

Tan, L. H., & Perfetti, C. A. (1998). Phonological codes as early sources of constraint in reading Chinese: A review of current discoveries and theoretical accounts. *Reading and Writing: An Interdisciplinary Journal, 10,* 165–220. http://dx.doi.org/10.1023/A:1008086231343

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181–198. http://dx.doi.org/10.3758/BF03197716

Verdonschot, R. G., Kiyama, S., Tamaoka, K., Kinoshita, S., Heij, W. L., & Schiller, N. O. (2011). The functional unit of Japanese word naming: Evidence from masked priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1458–1473. http://dx.doi.org/10.1037/a0024491

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40,* 374–408. http://dx.doi.org/10.1006/jmla.1998.2618

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition, 13,* 557–572. http://dx.doi.org/10.3758/BF03198326

Wydell, T. N. (1998). What matter in kanji word naming: Consistency, regularity, or On/Kun-reading difference? *Reading and Writing, 10,* 359–373. http://dx.doi.org/10.1023/A:1008083513500

Wydell, T. N., Butterworth, B., & Patterson, K. (1995). The inconsistency of consistency effects in reading: The case of Japanese Kanji. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1155–1168. http://dx.doi.org/10.1037/0278-7393.21.5.1155

Yamada, J. (1992). Why are kana words named faster than kanji words? *Brain and Language, 43,* 682–693. http://dx.doi.org/10.1016/0093-934X(92)90090-2

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5,* 683–689. http://dx.doi.org/10.3758/BF03208845

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37,* 533–554. http://dx.doi.org/10.1006/jmla.1997.2525

Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language, 48,* 779–793. http://dx.doi.org/10.1016/S0749-596X(03)00006-8

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 643–661. http://dx.doi.org/10.1037/0278-7393.34.3.643

Ziegler, J. C., Tan, L. H., Perry, C., & Montant, M. (2000). Phonology matters: The phonological frequency effect in written Chinese. *Psychological Science, 11,* 234–238. http://dx.doi.org/10.1111/1467-9280.00247

Appendix A

P-O Analyses Using Amano and Kondo's (2003b) Word Frequency Norms

Because Hino et al. (2011) used National Language Research Institute (1970) word frequency norms when analyzing the O-P and O-S consistencies for their 1,114 words, we used the same word frequency norms to compute the P-O consistencies here, norms which list only words whose word frequency counts are more than four per 940,533. At the same time, however, we also examined Amano and Kondo's (2003b) word frequency norms because these norms are newer and larger, involving 287,792,787.

Because the P-O consistencies using Amano and Kondo's (2003b) norms were highly correlated with those from National Language Research Institute (1970) for the 1,114 words, $r = .89$, $p < .001$, the results were essentially the same. That is, the mean P-O consistencies were .78 and .21 for the 339 katakana and 775 kanji words, respectively, using Amano and Kondo's (2003b) norms (vs. .83 and .27 using National Language Research Institute, 1970 norms), $F(1, 112) = 942.93$, $MSE = .08$, $p < .001$, $\eta^2 = .46$. After removing words with no phonological neighbors, the mean P-O consistencies were .58 and .17 for the 153 katakana and 722 kanji words, respectively, using Amano and Kondo's (2003b) norms (vs. .62 and .21 using the National Language Research Institute, 1970 norms), $F(1, 873) = 333.59$, $MSE = .06$, $p < .001$, $\eta^2 = .28$.

In addition, after classifying all the orthographic neighbors as orthographic friends in the phonological neighborhoods, we also recomputed P-O consistencies using Amano and Kondo's (2003b) norms. The mean P-O consistency based on Amano and Kondo's (2003b) norms was .44 for the 775 kanji words, which was significantly smaller than that for the 339 katakana words (.78), $F(1, 112) = 259.91$, $MSE = .11$, $p < .001$, $\eta^2 = .19$ (the values from National Language Research Institute, 1970 norms were .48 and .83, respectively). Similarly, the mean P-O consistency based on Amano and Kondo's (2003b) norms was .41 for the 722 kanji words with phonological neighbors, which was significantly smaller than that for the 153 katakana words with phonological neighbors (.58), $F(1, 873) = 34.60$, $MSE = .11$, $p < .001$, $\eta^2 = .04$ (the values from National Language Research Institute, 1970 norms were .44 and .62, respectively). As such, the patterns were essentially the same regardless of which word frequency norms were used.

(*Appendices continue*)

Appendix B

The More P-O Consistent and Less P-O Consistent Kanji Words Used in Experiments 1, 2, 3, and 4 Along With Their English Translations

More P-O consistent words		Less P-O consistent words	
Kanji word	English translation	Kanji word	English translation
語学	(study of) languages	炊事	cooking
受託	being given something in trust	旅券	a passport
水着	bathing suit	手芸	handicrafts
孤立	isolation	筆記	note-taking
懸念	fear, anxiety	修理	repair
庶民	the (common) people	財務	financial affairs
未定	undecided	世相	social conditions
荷物	luggage	翻後	old age
出前	meal delivery service	保存	conservation, storage
祖国	mother country	批評	criticism
武力	military force	若手	a young man
路上	on the road	保全	conservation, maintenance
序盤	opening, beginning	近所	the neighborhood
和服	Japanese clothes	下着	underwear
派出	sending out	下期	second half of the year
給油	refueling	任務	one's duty
背広	a jacket, a suit	手製	handmade
便秘	constipation	頭痛	headache
実務	(practical) business	不便	inconvenience
場面	a scene	利息	interest (on a loan)
秩序	order, regularity, system	専務	a managing director
序曲	an overture	根拠	basis, evidence
貯蓄	savings	社説	an editorial (article)
在庫	stock	談話	talk, conversation

Received December 2, 2015
Revision received June 1, 2016
Accepted June 4, 2016 ■