

## The Inconsistency of Consistency Effects in Reading: The Case of Japanese Kanji

Taeko N. Wydell and Brian Butterworth  
University College London

Karalyn Patterson  
Medical Research Council Applied Psychology Unit

Most Japanese Kanji characters have several different pronunciations, at least one ON-reading (of Chinese origin) and a KUN-reading (of Japanese origin); the appropriate pronunciation is determined by intraword context. There are also Kanji characters that have a single ON-reading and no KUN-reading. With 2-character ON-reading Kanji words as stimuli, naming experiments were carried out to investigate print-to-sound consistency effects. The consistent Kanji words were those in which neither constituent character had an alternative ON-reading or a KUN-reading, hence there can be no pronunciation ambiguity for these words. The inconsistent items were ON-reading words composed of characters that have KUN-readings that are appropriate to other words in which the characters occur, hence there should be some ambiguity about the pronunciation of the constituent characters. Six experiments yielded reliable effects of both word and character frequency and familiarity on speed and accuracy of word naming but virtually no evidence for consistency effects. It was concluded that for Kanji, phonology is dominantly computed at the word rather than at the character level.

A number of theoretical issues regarding reading processes may fruitfully be addressed by comparing different written languages or writing systems, for example, alphabetic English and logographic Japanese Kanji. The broader issue considered here is the process by which phonology is computed from orthography, and the more specific question is whether various “sizes of unit” contribute to this process for the two writing systems. A prominent technique for investigating the question of unit size is to assess the impact, on accuracy and speed of word naming, of the consistency of spelling–sound correspondences within neighborhoods of orthographically similar words. The principle underlying this line of investigation is that all whole words have a single conventionally correct pronunciation (except for rare homographs like *wind* and *lead* in English, in which the reader requires extraword context to select between two pronunciations); however, in many writing systems, including both English and Japanese, subword sized units may take several different alternative pronunciations, with the correct one determined by intraword context. If the reader translates spelling to sound at the level of the whole word only, then the existence of several competing pronuncia-

tions for subword chunks should not interfere; however, if the computation of phonology from orthography also occurs for units smaller than the word, then words containing “inconsistent” units—that is, those with different pronunciations in other words—might be at a disadvantage in accuracy or speed (or both) in word-naming tasks.

Many studies have addressed this question in English (e.g., Andrews, 1982, 1989, 1992; Glushko, 1979; Jared, McRae, & Seidenberg, 1990; Parkin, 1982; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Stanhope & Parkin, 1987; Taraban & McClelland, 1987; Waters & Seidenberg, 1985). On the basis of the frequent observation (e.g., Coltheart, 1978; Parkin, 1982; Waters & Seidenberg, 1985) that words with regular or typical spelling–sound correspondences (such as *five*) produce shorter naming latencies and lower error rates than words with exceptional correspondences (e.g., *give*), regularity was originally considered to be the critical variable. However, Glushko (1979) argued that consistency, rather than rule-defined regularity, provided a better account of empirical results. Although *five* may be a regular word “by rule,” its spelling–sound relationship is inconsistent with orthographically similar words such as *give*. To the extent that the process of computing phonology from orthography is sensitive to the characteristics of the neighborhood, then performance on a regular but inconsistent word like *five* should also adversely be affected. Glushko did indeed demonstrate longer response times (RTs) for regular inconsistent words than for regular words from consistent neighborhoods, though this result was not always obtained in subsequent experiments (e.g., Stanhope & Parkin, 1987; Taraban & McClelland, 1987; see also Brown, 1987; Patterson & Coltheart, 1987). In all studies manipulating either regularity or consistency, the effect has been shown to be stronger for low- than for high-frequency target items and is often statistically reliable only for the lower frequency words.

Jared et al. (1990) produced a more sophisticated hypothesis that captures aspects of results not handled by previous

---

Taeko N. Wydell and Brian Butterworth, Department of Psychology, University College London, London, United Kingdom; Karalyn Patterson, Medical Research Council Applied Psychology Unit, Cambridge, United Kingdom.

Part of this research was presented at a joint meeting of the Brain, Behaviour and Cognitive Science Society of Canada and the Experimental Psychology Society of Britain in Toronto, Canada, July 1993. The research was supported by a project grant (G9015838N) from the Medical Research Council.

We are grateful to A. Jonckheere for statistical advice and to Debra Jared for helpful comments on the original version of this article.

Correspondence concerning this article should be addressed to Taeko N. Wydell, Department of Psychology, University College London, Gower Street, London WC1E 6BT United Kingdom. Electronic mail may be sent via Internet to ucjrtw@ucl.ac.uk.

accounts referring solely to either regularity or consistency. According to Jared et al., the magnitude of the consistency effect for a given word is related to the summed frequency of that word's "friends" (words with a similar spelling pattern and a similar pronunciation) and of its "enemies" (words with a similar spelling pattern but with a different pronunciation). For example, an inconsistent word such as *haste* has a number of friends (e.g., *waste*, *paste*, *taste*, and so forth) and just a single enemy, *caste*. Against the strength of friends, the single enemy (which in the case of *caste*, moreover, is a rarely encountered word) cannot exert a marked influence; its negative impact on computing the pronunciation of *haste* will thus be small or even undetectable. An inconsistent word like *gown*, however, with many enemies (e.g., *blown*, *mown*, *sown*, and so forth), as well as friends (e.g., *town*, *clown*, and *frown*), gives rise to a more substantial effect. The commonly observed effect of regularity also finds a natural explanation within this account because many words with irregular spelling-sound correspondences defined by rule (such as *caste* and *pint*) happen to have few or no friends and many enemies. The theory (and supporting data) of Jared et al. also mesh well with (a) other results demonstrating the inadequacy of a simple regular-irregular dichotomy, such as the *degrees of regularity* effect described by Shallice, Warrington, and McCarthy (1983), and (b) computational models of translation from print to sound that demonstrate strong frequency and consistency effects (Plaut, McClelland, Seidenberg, & Patterson, 1994; Seidenberg & McClelland, 1989).

It thus appears that the computation of a word's pronunciation in English has contributions from both whole-word and subword levels; the former is demonstrated by the fact that consistency effects are modulated by the frequency of the target word, with very common words less vulnerable to the existence of unfriendly neighbors; the impact of the consistency of a word's orthographic neighbors provides evidence of generalization at the subword level. In a series of experiments reported in this article, we attempted to establish whether Japanese readers naming words written in Kanji show consistency effects paralleling those for English. To demonstrate what is meant by consistency in Kanji, we begin with a brief description of Japanese orthography.

### Japanese Orthography and Pronunciation Consistency

The Japanese orthography comprises two scripts: syllabic Kana and logographic Kanji (see Sampson, 1985; Sasanuma, 1986; Wydell, Patterson, & Humphreys, 1993, for more detailed descriptions). The feature of Kanji that is particularly germane to an understanding of how Japanese readers might translate orthography to phonology is that a typical Kanji character has two (or sometimes more) pronunciations: a KUN-reading and one or more ON-readings. A KUN-reading pronunciation is part of the original Japanese language and was assigned to Chinese characters from their meanings when Japan adopted the Chinese writing system. At the time of importing Chinese characters, however, not only were Chinese characters introduced, but Chinese words and their pronunciations were also added to the Japanese vocabulary—these are ON-readings. Also, a number of Kanji characters now have more than one ON-reading, as a consequence of the same

Chinese characters being introduced to Japan at several different periods.

A Kanji word may have from one to four characters; words with more than one character typically, but by no means always, take ON-readings. For Kanji characters with more than a single ON-reading (be it a KUN- or an additional ON-reading), the appropriate pronunciation is determined by the intraword context, that is, the other character(s) with which the particular character combines to constitute the word in question. For example, the Kanji character 親 (meaning parent) is pronounced "oya" (the KUN-reading) when it occurs in compound words such as 親方 *oyakata* (meaning foreman) or 父親 *chichioya* (meaning father); however, the same character is assigned the ON-reading "shin" when it occurs in the compound word 親類 *shinrui* (relatives) or 両親 *ryoshin* (parents). Any word containing the character 親 can therefore be described as an inconsistent word because pronunciation of the character varies across the orthographic neighborhood of words containing it.

There are, however, some Kanji characters that do not have KUN-readings and also have just a single ON-reading. Kanji compound words made up of these characters, therefore, have no pronunciation ambiguity because each constituent character has but a single possible pronunciation (e.g., for 氣候 *kikou*, meaning climate, each constituent character 気 *ki* and 候 *kou* has no KUN-reading and no other ON-reading). These words, which can be described as consistent, must have been the ones added to the Japanese vocabulary at the time of borrowing Chinese characters.

It should be noted that there are at least two respects in which issues of "regularity" and "consistency" of print-sound relationships differ between Kanji and English. The first of these is that, despite some ambiguity about how to define regularity in English, most investigators would agree that—for the inconsistent spelling pattern *int*, for example—words like *mint*, *hint*, and *lint* have the regular or typical pronunciation, with *pint* being the exception. There is no exact equivalent of such a regular or typical correspondence at the subword level in an inconsistent two-character Kanji word. It could be argued that the (dominant) ON pronunciation is the more typical because multicharacter words most often have ON pronunciations. On the other hand, because single-character words with KUN-readings almost always take that KUN-reading, it could also be argued that, at the character level, the KUN-reading for these items is the more typical. Because Kanji words do not lend themselves to classification as regular or exceptional (as well as for the reasons given above favoring a description in terms of consistency even for English), we refer to consistency rather than regularity in Kanji print-sound correspondences. The issue does, however, have some bearing on the design of empirical investigations. In the first experiments we reported in this article, we assess consistency effects by comparing two-character ON-reading consistent and inconsistent words primarily because there is a reasonable (though limited) number of characters with just a single ON-reading; there are very few characters with just a KUN-reading, which would make it impossible to select a set of multicharacter consistent KUN-reading words. On the grounds that the ON-reading of an inconsistent character in a multicharacter word might in

some sense be the more typical, however—which could make consistency effects more difficult to detect (just as consistency effects are harder to see in regular inconsistent than in exception inconsistent words in English)—in the last two experiments we focus on KUN-reading words.

A second notable difference between English and Japanese is that each KANJI character is a morphographic element that cannot phonemically be decomposed in the way that an alphabetic word can be. There are no separate components of the character 𪛗 that correspond to the individual phonemes in *oya* or *shin*. Whereas incompatible words from a neighborhood of shared-body words in English (such as *haste* and *caste*) will often have several phonemes in common, incompatible words from a neighborhood of shared-character words in Kanji (such as *oyakata* and *shinrui*) will typically have no phonological elements in common. Furthermore, because each character of an inconsistent Kanji word is a morpheme (with not only whole-morpheme pronunciation but also with meaning), in some sense the English equivalent would be not so much an example like *haste* versus *caste* but rather something more like English compound words containing heterophonic homographs (e.g., *bowtie* vs. *bow-wow* or *lead-in* vs. *lead-free*).<sup>1</sup> We know of no evidence on consistency effects in such English compounds, nor indeed is it clear that there are a sufficient number of compounds containing heterophonic homographs to form the stimuli for an experiment to investigate such putative effects.

Despite these rather different senses of consistency in alphabetic writing systems like English and in the morphographic multireadings system of Japanese Kanji, with regard to the potential impact of inconsistency on word naming, the theoretical interpretation seems approximately parallel in the two cases. The fact that inconsistent English words give rise to longer RTs and higher error rates than consistent words suggests that subword levels play a significant role in the computation of word pronunciation. Analogously, if inconsistent Kanji words were to yield longer RTs and higher error rates than consistent words, this would then suggest that the subword level of constituent character plays a major role in computing the pronunciation of written Kanji words.

In this study, therefore, we designed the experiments to evaluate whether Kanji words with consistent character-to-sound correspondences have a naming advantage over Kanji words with inconsistent correspondences. A significant impact of consistency should reveal itself either in longer RTs or more errors (or both) to inconsistent words. A further prediction is that just as errors to inconsistent words in English typically take the form of assigning a pronunciation appropriate to other members of the orthographic neighborhood (e.g., *pint* pronounced to rhyme with *hint*), a consistency effect on accuracy in Kanji word naming should produce errors in which inconsistent characters are assigned an alternative ON- or KUN-reading.

## Experiment 1

### Method

#### Participants

Twenty Japanese male and female native speakers (aged between 18 and 41 years), who were brought up in Japan until at least age 18 and

had lived in the United Kingdom for less than 18 months, served as participants in this experiment. All had normal or corrected-to-normal vision. Each participant was paid a small fee for participating.

#### Apparatus

A Macintosh II computer with the software Psychlab (Gum & Bub, 1988) was used to run the experiment. Japanese Kanji stimulus words were generated by the Macintosh software SweetJAM 4.5 (A & A Company, Ltd., Tokyo, 1990) installed on the English operating system. RTs were recorded by the computer via a throat microphone connected to a voice key. Responses were also tape recorded for later checking.

#### Stimuli

All stimulus words in this and all subsequent experiments in the series were composed of characters from approximately the 3,000 most commonly used Kanji characters and would therefore be known by any average literate Japanese adult. In Experiment 1, 80 two-character ON-reading Kanji nouns, all with three-syllable (or rather three-morae) pronunciations, were divided into four different conditions of 20 words each as follows.

*Consistent.* Each component character of a two-character word has only a single ON-reading, hence the pronunciations of words in this condition are consistent.

*Intermediate-1.* Either the first or the second character of a two-character word has a KUN-reading as well as an ON-reading, but the KUN-readings do not apply to any two-character words. Therefore, the pronunciations of these words are still fairly consistent.

*Intermediate-2.* Both the first and second characters of a word have KUN-readings, but these KUN-readings do not occur in other two-character words.

*Inconsistent.* Either one or both component characters have KUN-readings, and these KUN-readings are appropriate to other words containing the character(s). Hence, the pronunciations of these words are inconsistent.

In an experiment comparing groups of words (here, differing in consistency), it is obviously important to balance the different sets of stimuli for other factors, such as word frequency or familiarity, that might influence performance. Unfortunately, there are no published word norms for Japanese Kanji; therefore, values for word frequency and word familiarity were taken from a corpus of 2,357 Japanese nouns rated for frequency, familiarity, and imageability on a scale of 1 (*very low*) to 7 (*very high*) that were collected by Wydell (1991; see also Wydell, Quinlan, & Butterworth, 1995). A pilot study preceding Experiment 1 established rated familiarity as a better predictor than rated frequency of naming RT. The same pilot study revealed that, in addition to effects of whole-word familiarity, naming RT was significantly predicted by individual Kanji character frequency, a measure for which there are published norms in *Gendai Shinbun No Kanji* (National Language Institute, 1976, *A Study of Uses of Kanji Characters in Modern Newspapers*). Unfortunately, it is not possible to achieve perfect matching of component character frequencies across conditions because a character's frequency of occurrence and its ambiguity of pronunciation are closely correlated in the Japanese language: Characters with multiple pronunciations appear in many more words than do characters with consistent pronunciations.

Stimulus words were matched across conditions as closely as possible for both rated word frequency and rated word familiarity (mean values for the four conditions of Experiment 1 are given in Table 1). We made an attempt to match both first and second character frequencies across the four conditions, although (for the

<sup>1</sup> We are grateful to an anonymous reviewer for raising this point.

reasons explained above) the matching achieved was far from ideal; this imbalance was handled with an analysis of covariance (ANCOVA). Finally, the initial phonemes of the stimulus words were matched as closely as possible across the four conditions, in which it was not possible to match at the level of the identical initial phoneme, then at least matching was achieved within groups of similar phonemes (e.g., fricatives, nasals, and so forth).

### Procedure

Each trial began with the presentation of a fixation dot for 500 ms. The fixation dot was then replaced by a target word, which remained visible until a response was made or a maximum of 3,000 ms had elapsed. The order of the stimulus words was randomized by the computer program for each participant. We instructed the participants to pronounce the target word as quickly and accurately as possible on its presentation. Naming latencies were recorded by the computer. T. Wydell sat next to each participant throughout the experiment and made notes of errors and accidental triggerings of the voice key (e.g., throat clearing before the response). Twenty practice trials were given before the experimental trials. For the analysis, a cutoff point of two standard deviations above or below the mean for each participant for that condition was used to clean the raw RT data.

### Results

Table 1 shows mean RTs (in milliseconds) and percentage of error rates for each condition.

#### RTs

A one-way analysis of variance (ANOVA) with experimental condition as a variable, performed on the RTs correspond-

Table 1  
Word Characteristics, Mean Reaction Times (RTs; in Milliseconds), and Error Rates (in %) in Kanji Naming With Four Conditions in Experiment 1 (Straightforward Naming) and Experiment 2 (Speeded Naming)

Variable	Condition			
	1	2	3	4
Word characteristics				
Word frequency				
<i>M</i>	4.16	3.97	4.00	3.73
<i>SD</i>	0.85	0.84	0.88	0.69
Word familiarity				
<i>M</i>	4.10	4.15	4.12	4.08
<i>SD</i>	0.71	0.67	0.71	0.84
Experiment 1				
RT				
<i>M</i>	657	672	672	661
<i>SD</i>	70.73	75.05	105.72	91.87
Adjusted RT	660	673	674	656
Error rate	6	3	4	6
Experiment 2				
RT				
<i>M</i>	575	584	581	569
<i>SD</i>	65.34	74.85	77.06	76.42
Adjusted RT	579	585	584	569
Error rate	6	12	8	10

Note. 1 = consistent condition; 2 = Intermediate-1 condition; 3 = Intermediate-2 condition; 4 = inconsistent condition.

Table 2

Number of Different Types of Errors for the Four Conditions in Experiments 1 and 2

Error type	Experiment 1				Experiment 2			
	Condition				Condition			
	1	2	3	4	1	2	3	4
ON-KUN reading	0	0	1	1	0	0	5	3
Start-Stop	7	8	7	15	13	34	17	29
Visual	10	3	2	2	9	8	2	6
Semantic	5	1	0	1	2	0	1	0
Other	1	0	4	4	1	5	6	3
Total	23	12	14	23	25	47	31	41

Note. 1 = consistent condition; 2 = Intermediate-1 condition; 3 = Intermediate-2 condition; 4 = inconsistent condition.

ing to correct word-naming responses, yielded no significant difference because of condition, either over subjects,  $F_1(3, 19) < 1$ , or over items,  $F_2(3, 76) < 1$ .

Analyses of regression and covariance, with whole-word familiarity and individual character frequencies (log transformed) as variables, revealed only a significant effect of familiarity: for the regression,  $t(74) = -3.02, p < .01$ ; for the ANCOVA,  $t(74) = -3.20, p < .01$ . The adjusted mean RTs for the four conditions (as if the covariate of character frequency had properly been matched across the four conditions) are shown in Table 1.

#### Errors

Error scores were square root transformed after the addition of a constant of 0.1 to each error score and then submitted to a one-way ANOVA. There was no significant difference between the four conditions ( $F < 1$ ). A priori mean comparisons between the four conditions also revealed no significant differences between these conditions.

The errors were categorized according to the error types in Table 2. The error types here (a) alternative ON- or KUN-reading (one of the characters was given a KUN-reading or a different inappropriate ON-reading), (b) start-stop errors (the participant uttered the first character correctly then stopped and named the whole word correctly), (c) visual error (a different word visually similar to the target word was produced), (d) semantic error (a different word semantically related to the target was produced), and (e) other (including nonwords). For the consistent condition, there were 10 (out of 23 total errors) visual errors, but one particular item produced all 10 errors. Similarly, in the consistent group, one word 保護 (a relatively uncommon word, meaning protection) yielded all five semantic errors; for the inconsistent condition, a particular item produced 6 out of 15 start-stop errors. There was no obvious meaningful pattern in the error types.

#### Discussion

The results of Experiment 1 are straightforward: Word-naming latency did not vary as a function of the degree of character-to-sound consistency (as analyzed by both ANOVA

and ANCOVA). The only variable significantly influencing naming latency was rated word familiarity, with faster RTs to high- than to low-familiarity Kanji words. Error analyses also failed to reveal significant consistency effects. If pronunciation ambiguity of constituent characters were having a marked impact, then one would expect to see errors for Intermediate-2 and especially inconsistent words, in which participants substitute a KUN-reading for the correct ON-reading. Only one error of this type occurred in each of these two conditions.

## Experiment 2

One possible reason for our failure to observe consistency effects in Experiment 1 is that character-to-sound consistency may have its impact very early in processing; by the time that a word is uttered—when of course any ambiguity must be resolved—these effects might be attenuated (Monsell, Doyle, & Haggard, 1989). In this case, the paradigm of straightforward naming may not ideally be suited to observing consistency effects. In Experiment 2, therefore, a speeded naming technique was used. Strain, Patterson, and Seidenberg (1995), using this technique for naming of English words, obtained a high rate of regularization errors to lower frequency words with atypical spelling-sound correspondences. They argued that this may indicate the prominence of correspondences between subword orthographic and phonological segments in the direct translation from spelling to sound; these will especially reveal their strength when the whole-word influence of meaning-level representations is prevented from having an impact on the naming process, as in the case of speeded naming. If the same argument applied to Kanji, then perhaps speeded naming would produce the kind of ON-KUN confusion errors that we failed to observe in Experiment 1.

## Method

### Participants

Twenty Japanese male and female native speakers (aged between 19 and 25 years), who were brought up in Japan until at least age 18 and had lived in the United Kingdom for less than 18 months, served as participants in this experiment. All had normal or corrected-to-normal vision. Each participant was paid a small fee for participating.

### Apparatus and Stimuli

The apparatus and stimuli were identical to those used in Experiment 1.

### Procedure

Each trial began with the presentation of a fixation cross (for 500 ms), followed by a target word exposed for a duration calibrated individually for each participant in the course of 50 practice trials. Participants were asked to try to start naming each target word before the deadline, as indicated by the disappearance of the stimulus word from the screen. Target-word exposure durations varied across participants between 316 ms and 449 ms (average was 384 ms). To enable settings of the exposure durations, we increased the number of practice trials from 30 to 50. Practice began with a 600-ms duration,

which was decreased by 15-ms steps until the participant reached a critical exposure duration, defined as 80%–85% correct responses.

Response latencies were recorded by the computer from the onset of the target. Despite the fact that participants were instructed to try to “beat” the deadline, all reaction times were included in the analyses, provided that the words were pronounced correctly. This is the procedure used in other experiments with speeded naming of either words (Strain et al., 1995) or pictures (Vitkovitch & Humphreys, 1991). In all other respects, the procedure was identical to that in Experiment 1.

## Results

Table 1 shows correct mean RTs (in milliseconds) and percentage of error rates for the four conditions.

### RTs

Correct RTs were submitted to a one-way ANOVA with experimental condition as a variable, which yielded a significant difference between conditions over subjects,  $F_1(1, 19) = 4.10$ ,  $MSE = 223.48$ ,  $p < .01$ , but not over items,  $F_2(3, 76) < 1$ . Planned comparisons between conditions over subject RTs revealed that RTs in the inconsistent condition were significantly faster than those in both intermediate-1,  $F(1, 57) = 10.40$ ,  $p < .002$ , and Intermediate-2,  $F(1, 57) = 6.60$ ,  $p < .012$ . Other comparisons were nonsignificant.

As with Experiment 1, ANCOVA with first and second character frequencies (log transformed), word frequency, familiarity, and imageability as covariates yielded no significant differences between the four experimental conditions ( $F < 1$ ). The ANCOVA-adjusted means for the four conditions are shown in Table 1. A multiple regression analysis with the same variables showed, as in Experiment 1, a significant effect only for rated word familiarity,  $t(73) = -3.28$ ,  $p < .002$ .

### Errors

A one-way ANOVA on error scores (square root transformed after a constant of 0.1 was added to each error score) just failed to reach significance,  $F(3, 19) = 2.46$ ,  $MSE = 0.215$ ,  $p = .07$ . However, the a priori comparison between the consistent and inconsistent groups revealed that more errors were made to words in the inconsistent condition,  $F(1, 57) = 4.14$ ,  $p < .046$ . Also, Intermediate-1 words produced significantly more errors than consistent words,  $F(1, 57) = 6.21$ ,  $p < .015$ .

As with Experiment 1, errors were categorized according to the error types in Table 2. By comparison with straightforward naming in Experiment 1, speeded naming produced a higher rate of errors, except in the consistent condition. This increase is almost entirely attributable to one category of response, start-stop, which in a sense is not a true error. These are cases in which the participant pronounced the first character (or, much more rarely, the second character) of a two-character Kanji word correctly, paused, and then pronounced the whole word correctly. These responses, which clearly result from the pressure to start speaking quickly, probably correspond to the anticipation errors characteristic of speeded naming in English, in which the participant utters the first phoneme or two

of the target word, pauses, and then produces the whole word (see Strain et al., 1995). Speeded naming also yielded a few more errors, relative to straightforward naming, of the type labeled *visual errors*, in which participants produced a different two-character Kanji word containing the same first (or, much more rarely, second) character as the target word.

Although the increased incidence of visual errors and especially start-stop errors under conditions of speeded naming is not without interest, it is not especially germane to the primary issue of consistency effects because the likelihood of these errors did not vary systematically across the four conditions. In particular, the largest number of start-stop errors occurred in the Intermediate-1 condition. The errors that would be most indicative of consistency effects are alternative ON- or KUN-reading errors, in which a participant initially produced either a KUN-reading (or different ON-reading) of the first character of a target word, then realized his or her mistake and pronounced the whole word correctly. As noted earlier, only two errors of this type occurred in Experiment 1: one each in the Intermediate-2 and inconsistent groups. Under conditions of speeded naming, the corresponding numbers of ON-KUN errors for the Intermediate-2 and inconsistent conditions were five and three. Although this is still an almost vanishingly small number of such errors, the fact that they occurred at all, and more under conditions of speeded naming, provides the first hint of character-level processing in this series of experiments.

### Discussion

Mean latencies in Experiment 2 were on average 82 ms shorter than those in Experiment 1, suggesting the general effectiveness of the speeded naming instructions. Initial analyses of Experiment 2 suggested that the inconsistent words were named more quickly than the other types of words (Intermediate-1, -2, and consistent words). However, this difference disappeared under ANCOVA. In other words, once again, naming latencies of two-character Kanji words did not differ significantly as a function of degree of consistency of their component characters. Significantly more errors were made to the inconsistent words than to the consistent words; however, this apparent small consistency effect itself was not very consistent because the error rate in the inconsistent group was very similar to that for the fairly consistent words of the Intermediate-1 condition. Thus, with the exception of a very few ON-KUN confusion errors in the Intermediate-2 and inconsistent conditions, both latency and error analyses of speeded naming of Kanji words have once again failed to reveal any prominent character-to-sound consistency effects.

### Experiment 3

In Experiments 1 and 2, word familiarity ranged over a relatively wide spectrum; apart from the best possible matching of familiarity across conditions, however, we made no attempt to manipulate these variables. Consistency effects in English (e.g., Jared et al., 1990), and also in other writing systems such as Italian (Colombo, 1992) and Chinese (Hue, 1992), are more prominent for lower frequency words; we

could thus argue that any consistency effects in Kanji will have been diluted by use of word sets ranging from low to high frequency. In Experiment 3, therefore, word familiarity was treated as a dichotomous variable. Furthermore, given our previous failure to find significant differences even between the most extreme conditions of consistent and inconsistent words, much less any differences as a function of more subtle gradations of consistency, in Experiment 3 the manipulation of consistency was also achieved with a dichotomous classification. If there is any tendency toward consistency effects in Kanji naming, this should be revealed in a Frequency  $\times$  Consistency interaction, as seen in other languages. In one additional effort to uncover consistency effects, in this experiment we presented consistent and inconsistent stimulus words in separate blocks on the hypothesis that blocked presentation of stimuli might make character-to-sound consistency-inconsistency more salient.

### Method

#### Participants

Eleven Japanese male and female native speakers (aged between 18 and 46 years), who were brought up in Japan until at least age 18 and had lived in the United Kingdom less than 20 months, served as participants in this experiment. All had normal or corrected-to-normal vision. Each participant was paid a small fee for participating.

#### Apparatus

The apparatus was identical to that used in Experiments 1-3.

#### Stimuli

There were 80 ON-reading two-character Kanji nouns: 40 with consistent pronunciations (as in the consistent condition of Experiments 1 and 2) and 40 with inconsistent pronunciations (as in the inconsistent group in the previous experiments). Half of the words in each consistency group were relatively common words (high familiarity: above 4.45 on a scale of 1 to 7), and half were relatively uncommon (low familiarity: below 3.55 on a scale of 1 to 7). Also, half of the items in each consistency-familiarity group had three-morae pronunciations, and half four-morae. Word familiarity and initial phoneme were matched, pairwise, between the two consistency conditions.

#### Procedure

Presentation of stimulus items was blocked according to condition, that is, participants named all of the consistent words in one block and all of the inconsistent words in another. Half of the participants started with consistent words, and half with inconsistent words. The order of stimulus words within a block was randomized by the computer before presentation such that each participant received a different word order. The rest of the procedure was the same as that in Experiment 1.

### Results

Table 3 shows mean RTs for correct responses (in milliseconds percentage of error rates) per condition.

RTs

RTs were submitted to a two-way ANOVA, with independent variables of consistency and familiarity, and with repeated measures over subjects and items. The results showed a significant main effect of familiarity:  $F_1(1, 10) = 47.90, MSE = 910, p < .001; F_2(1, 76) = 21.16, MSE = 4,187, p < .001$ . Although there was a small RT advantage of consistent over inconsistent words, neither the main effect of consistency nor the interaction between familiarity and consistency were significant in either subject or item analyses (all  $F_s < 1$ ).

Errors

Error scores were square root transformed after addition of a constant of 0.1 and were submitted to a two-way ANOVA with the same variables as the RT analyses. Subject analyses showed that the main effects of consistency,  $F_1(1, 10) = 7.84, MSE = 0.040, p < .018$ , and familiarity,  $F_1(1, 10) = 74.64, MSE = .008, p < .001$ , as well as the interaction between consistency and familiarity,  $F_1(1, 10) = 14.41, MSE = 0.011, p < .003$ , were all significant. More errors were made with inconsistent than with consistent words and with low-familiarity than high-familiarity words. Further, a priori mean comparison tests revealed that the effect of consistency was significant only for the low-familiar words and that there was no significant difference between consistent and inconsistent high-familiarity words. Item analyses, however, showed only a significant effect of familiarity,  $F_1(1, 76) = 6.31, MSE = 0.002, p < .01$ ; the consistency effect was only approaching significance,  $F_1(1, 76) = 3.40, p = .069$ . The interaction between consistency and familiarity was not significant ( $F < 1$ ).

Errors were categorized according to the usual scheme and are shown in Table 4. Here it can be seen that, despite the higher error rate for inconsistent than for consistent words, (a) the actual number of errors was still very low, (b) half of the errors on inconsistent words (9 out of 17) were of the start-stop variety that—as indicated before—are not exactly incorrect responses, and (c) there were once again rather few of the alternative ON- or KUN-reading errors that would be most indicative of character-level consistency effects.

Table 3  
Mean Reaction Times (in Milliseconds) and Error Rates (in %) in Kanji Naming With Blocked Stimuli Presentation (Consistency and Familiarity) in Experiment 3

Variable	Consistent		Inconsistent	
	High familiarity	Low familiarity	High familiarity	Low familiarity
Familiarity				
<i>M</i>	4.95	3.05	5.10	3.20
<i>SD</i>	0.32	0.46	0.46	0.29
Consistency				
<i>M</i>	640	700	646	713
<i>SD</i>	92.73	102.37	83.72	106.40
ER	0.5	2	0.9	7

Note. ER = error rate.

Table 4

Number of Different Types of Errors for the Two Conditions in Experiments 3 and 4

Error type	Experiment 3		Experiment 4	
	Condition		Condition	
	1	4	1	4
ON-KUN reading	0	5	0	4
Start-Stop	4	9	7	2
Visual	1	0	1	2
Semantic	0	1	1	2
Other	1	2	8	1
Total	6	17	17	11

Note. 1 = consistent condition; 4 = inconsistent condition.

Discussion

In Experiment 3, naming latency for two-character Kanji nouns was significantly affected by word familiarity but not by pronunciation consistency of component characters; there was also no significant Familiarity × Consistency interaction. Virtually all of the relatively few errors occurred on lower familiarity words; although more errors were made on inconsistent than on consistent words, the ANOVA produced a significant main effect of consistency and a significant Familiarity × Consistency interaction for errors over subjects but not over items. Once again, then, we have failed to uncover any RT effects of character-sound consistency in naming Kanji words and have detected only a tiny trend toward such effects in the accuracy data.

Experiment 4

In Experiment 4, pronounceable two-character Kanji nonwords were included along with the word stimuli used in Experiment 3. These nonwords have unambiguous pronunciations because they were constructed by combining into nonlexical two-character strings characters from consistent words in which constituent characters have only one ON-reading each. Like the frequency and blocking manipulations of Experiment 3, we hypothesized that the inclusion of nonwords in Experiment 4 might coax out the consistency effects that we have thus far failed to observe. The basis for this hypothesis is that because the Kanji nonwords are unfamiliar (indeed, illegitimate) combinations of characters, the only way in which participants can pronounce the nonwords is to process them at the level of the individual component characters. If this character-by-character reading for the nonword stimuli were to induce a similar strategy for the words, then—because both word and character levels would yield only one possible pronunciation for the consistent words but could yield conflicting pronunciations for the inconsistent words—consistency effects in Kanji reading might emerge. The parallel to this predicted phenomenon in English has been demonstrated by Monsell, Patterson, Graham, Hughes, and Milroy (1992), who found a significantly higher error rate (particularly regularization errors) to exceptional inconsistent English words when nonwords were mixed in with the words.

### Method

#### Participants

Ten Japanese male and female native speakers (aged between 18 and 38 years), who were brought up in Japan until at least age 18 and had lived in the United Kingdom for less than 18 months, served as participants in this experiment. All had normal or corrected-to-normal vision. Each participant was paid a small fee for participating.

#### Apparatus

The apparatus was the same as in Experiment 1.

#### Stimuli

The stimulus list consisted of 80 Kanji words and 80 nonwords. The word stimuli were the same as those used in Experiment 3 (40 consistent and 40 inconsistent words, further divided into 20 each of high- and low-familiarity words). The nonwords were constructed from two-character Kanji words in which each constituent character has only a single ON-reading. One character from such a word was combined with another character from a different word, resulting in two-character Kanji nonwords that were unambiguously pronounceable but meaningless. The words used to construct the nonwords items were different from those used as word stimuli, thus preventing any intralist repetition of characters.

#### Procedure

All stimulus words and nonwords were randomized by the computer so that each participant saw a different randomization order of the stimuli. The rest of the procedure was the same as in Experiments 1-3 by using a straightforward naming technique.

### Results

Mean RTs for correct responses (in milliseconds) and percentage of error rates per condition are shown in Table 5.

#### Analyses With Nonwords

*RTs.* Mean RTs for correct responses were submitted to a one-way ANOVA. The main effect of stimulus type was significant,  $F_1(2, 9) = 102.10$ ,  $MSE = 1,055$ ,  $p < .001$ . A priori mean comparisons between three different stimulus types revealed that there were significant differences between consistent words and nonwords,  $F_1(1, 18) = 134.30$ ,  $p < .001$ , and between inconsistent words and nonwords,  $F_2(1, 18) = 170.16$ ,  $p < .001$ . However, the difference between consistent and

inconsistent words was not significant. Item analyses were not carried out because the number of stimulus items differed between conditions, consistent and inconsistent words (40 each), and nonwords (80).

*Errors.* Error scores (square root transformed after a constant of 0.1 was added) were submitted to a one-way ANOVA. As with the RT analyses, the main effect of stimulus type was significant,  $F_1(2, 9) = 44.50$ ,  $MSE = 0.105$ ,  $p < .001$ . A priori mean comparisons between different stimulus types revealed that there were significant differences between consistent words and nonwords,  $F_1(1, 18) = 54.06$ ,  $p < .001$ , and between inconsistent words and nonwords,  $F_2(1, 18) = 77.55$ ,  $p < .001$ . However, the difference between consistent and inconsistent words was not significant ( $F < 1$ ).

#### Analyses Without Nonwords

Further analyses were carried out on word stimuli only, with word type (consistent or inconsistent) and familiarity (either high or low) as variables.

*RTs.* Mean RTs for correct responses were submitted to a two-way ANOVA. There was a significant main effect of familiarity, with shorter RTs to high- than to low-familiarity words,  $F_1(1, 9) = 85.09$ ,  $MSE = 956$ ,  $p < .001$ . There was also a significant main effect of consistency,  $F_1(1, 9) = 7.80$ ,  $MSE = 555$ ,  $p < .02$ , but in the opposite direction to that predicted, with faster RTs to inconsistent than to consistent words. Item analyses, however, revealed that only familiarity was significant,  $F_2(1, 76) = 26$ ,  $MSE = 6,531$ ,  $p < .001$ , and neither the main effect of consistency nor the interaction between familiarity and consistency were reliable ( $F_s < 1$ ).

As with Experiment 1, we carried out an ANCOVA to estimate the differences between consistent and inconsistent words, had the variable of character frequency been properly matched between conditions. Character frequencies might be especially important in this experiment because the Kanji word stimuli were mixed with Kanji nonwords, which can only be named character-by-character. The adjusted means were 740 ms for the consistent words and 735 ms for the inconsistent words, a nonsignificant difference ( $F < 1$ ).

*Errors.* Error scores (transformed in the usual way) were submitted to a two-way ANOVA with word type and familiarity as variables. The results showed only a significant main effect of familiarity:  $F_1(1, 9) = 14.65$ ,  $MSE = 0.041$ ,  $p < .004$ ;  $F_2(1, 76) = 4.73$ ,  $MSE = 0.177$ ,  $p < .03$  (all other  $F_s < 1$ ). More errors were made with low-familiarity words than with high-familiarity words.

Table 5  
Mean Reaction Times (RTs; in Milliseconds) and Error Rates (ERs; in %) in Kanji Naming of Words With Nonwords in Experiment 4

Condition	High familiarity			Low familiarity			Total mean		Adj. M
	M	SD	ER	M	SD	ER	RT	%	
Words									
Consistent	698	81.15	2	795	102.50	7	745	5	740
Inconsistent	685	75.12	2	767	88.69	4	725	3	735
Nonwords							914	10	

Note. Adj. = adjusted.

The errors were tabulated according to type and are summarized in Table 4. As in most of the preceding experiments, there were few errors in no discernible pattern and essentially no alternative ON- or KUN-reading errors even on low-frequency inconsistent words.

### Discussion

As in experiments on word-nonword naming in English, there was an advantage in both speed and accuracy for naming of Kanji real words compared with Kanji nonwords. In fact, the RT difference between words and nonwords (an average of about 180 ms) was extremely large compared with the lexicality advantage typically observed in English naming studies, a point that will be taken up in the General Discussion section. Relative to previous experiments in this series (especially Experiment 3, in which we used exactly the same words), the presence of nonwords also seems to have slowed the latencies for words: The mean RT for the words in Experiment 4 was 735 ms, compared with 676 ms in Experiment 3 (though of course it must be recalled that the two experiments had different individuals as participants). This apparent RT difference might be an indication that participants shifted their attention more toward character-by-character reading rather than whole-word processing; however, as this slowing was not accompanied by an advantage for consistent over inconsistent words, nor by a Familiarity  $\times$  Consistency interaction for either accuracy or RT, it certainly does not provide any convincing indication of character-level influences in Kanji word naming. The manipulation of mixing nonwords into a list containing inconsistent Kanji words has failed to produce an effect parallel to that seen in English, in which the addition of nonwords yields a significantly augmented rate of regularization errors to exception-inconsistent words (Monsell et al., 1992), an effect interpreted in terms of processing at subword levels.

In all four experiments presented thus far, the critical stimulus materials have been ON-reading Kanji words, composed either of characters with a single unambiguous pronunciation or of characters that—in other contexts—can take one or more alternative pronunciations. As explained toward the end of the introduction, there were good reasons for concentrating on ON-reading words; however, as explained there also, it is just possible that the ON-reading of an inconsistent character in a multicharacter word is in some sense the more dominant pronunciation, making the impact of an alternative but less dominant pronunciation hard to observe. In Experiments 5 and 6, therefore, we turn our attention to words with KUN-readings.

### Experiment 5

In this experiment, single-character KUN-reading words were used as stimuli. Our rationale was to see whether consistency effects might be revealed by a manipulation of the number of pronunciation enemies possessed by a given character. The KUN-reading one-character words selected for this experiment had either a single alternative ON-reading or two ON-readings: these two word classes were labeled *single-enemy*

and *double-enemy* words. If alternative pronunciations of a character are computed, and the correct pronunciation must be chosen amongst these alternatives, in the course of naming a Kanji word, then the double-enemy words should perhaps yield longer RTs or higher error rates than the single-enemy words.

### Method

#### Participants

Twenty Japanese male and female native speakers (aged between 19 and 43 years), who were brought up in Japan until at least age 18 and had lived in the United Kingdom for less than 18 months, served as participants in this experiment. All had normal or corrected-to-normal vision. Each participant was paid a small fee for participating.

#### Apparatus

The apparatus was the same as in Experiments 1-4.

#### Stimuli

There were 42 KUN-reading single Kanji nouns: half with single alternative ON-readings (the single-enemy condition) and half with two ON-readings (the double-enemy condition). Words from the two conditions were matched as closely as possible on an item-by-item basis for initial phoneme, number of morae (ranging between two and four), rated word frequency, familiarity, and imageability (mean word frequency, familiarity, and imageability are 3.76, 4.84, and 6.25 for single-enemy and 3.76, 4.98, and 6.27 for double-enemy, respectively). It was not possible to match the words for character frequency because, as explained earlier, there is an inevitable confounding in which characters with multiple pronunciations occur more commonly than those with few or only one pronunciation.

#### Procedure

The procedure was straightforward word naming as in Experiment 1.

### Results

Table 6 shows mean RTs (in milliseconds) for correct responses and percentage of error rates for the two conditions.

#### RTs

A one-way ANOVA on mean RTs for correct responses showed that the difference between conditions was approach-

Table 6  
Mean Reaction Times (in Milliseconds) and Error Rates (ERs; in %) in Kanji Naming With Single Versus Double Enemies in Experiment 5

Condition	Correct response			
	<i>M</i>	<i>SD</i>	Adj. <i>M</i>	ER
Inconsistent				
Single enemy	649	56.34	632	5
Double enemy	634	59.93	634	3

Note. Adj. = adjusted.

ing significance over subjects,  $F_1(1, 19) = 4.15$ ,  $MSE = 571$ ,  $p = .055$ , but in the opposite direction to that predicted and was not significant over items,  $F_1(1, 40) = 1.59$ ,  $MSE = 1,991$ ,  $p = .21$ .

Because the character frequency match between the two conditions was far from perfect, an ANCOVA was carried out with character frequency as covariate; however, this also failed to reveal any difference between conditions,  $F(1, 36) = 2.63$ ,  $MSE = 1,628$ ,  $p = .11$ . The ANCOVA-adjusted means for the two conditions are shown in Table 6, as are the error rates that were again in the opposite direction to that predicted by a hypothesis that number of enemies should have an adverse effect on naming.

### Errors

Neither subject nor item analyses on errors were significant ( $F_s < 1$ ). The number of errors was too small to warrant classification into different types.

### Discussion

The result of this experiment was quite clear: There was no difference in RT or error rate between KUN-reading one-character Kanji words with a single ON-reading enemy and those with two ON-reading enemies. Once again, it appears that character-to-pronunciation ambiguity in Kanji word naming has little effect on naming. There is, however, one final line of investigation to be explored, in a comparison of ON- and KUN-reading two-character words, which formed the content of Experiment 6.

### Experiment 6

Experiments 1–4 compared, under various conditions, the naming of consistent and inconsistent two-character ON-reading words. Experiment 5 compared single-enemy and double-enemy one-character KUN-reading words. None of these experiments has involved a contrast between ON- and KUN-reading words. If both readings of an inconsistent character are activated at some stage in the process of computing the pronunciation of a word containing it, then a contrast between words with ON- and KUN-readings might at long last provide some evidence of consistency effects in Kanji. For this final experiment, the stimulus words selected were pairs of two-character words sharing the identical first character, in which one word of each pair takes an ON-reading and the other a KUN-reading. We call this a final hunt for consistency effects because, of course, any word composed of characters with both ON- and KUN-readings must by definition be an inconsistent word. That is, we did not compare consistent and inconsistent words but rather two groups of inconsistent words; the contrast is therefore something more like that between regular inconsistent and exception inconsistent words in English or perhaps like exception words characterized as having different degrees of irregularity in English (Shallice et al., 1983). As expressed in the general introduction, it is not even entirely clear whether it is the ON- or the KUN-reading of a character with both readings that should be

considered the more typical: the latter is almost always appropriate when the character forms a word on its own, but the former is much more common in words of two-plus characters. Because the modal uninflected word (noun) in Kanji has two characters, and because on average far more two-character words take ON- than KUN-readings, the obvious assumption is that the ON-reading would be more dominant. The prediction for Experiment 6 is that if consistency has some impact on computation of pronunciation in Kanji, KUN-reading two-character words should yield longer RTs, a higher error rate, or both, than ON-reading words.

This set of stimulus materials also offers the opportunity to address, though only in an approximate fashion, the issue of relative frequency of friends and enemies in the neighborhoods of inconsistent Kanji words. As explained in the *Method* section of Experiment 1, because there are no word frequency norms in Japanese (only character frequency norms exist), it is not possible to do the kind of detailed statistics on summed token frequency of friends and enemies that has illuminated the study of consistency effects in English (Jared et al., 1990). Because the Japanese dictionary *Toyo Kanji jiten* (1962) lists all of the multicharacter words containing particular characters, however, one can calculate type frequencies—that is, as applied to the present set of stimulus words, we can specify the mean number of two-character words containing each component character that has ON- versus KUN-readings. If there are substantially more ON- than KUN-reading words containing these characters, then despite the fact that we cannot calculate token frequencies, it seems safe to assume that the ON-reading words of the matched pairs would have, on average, far higher summed-frequency friends than enemies, whereas the reverse would be true for the KUN-reading stimulus words. This is indeed the case for type frequency: With respect to the first (matched) character in the 26 pairs of words used in this experiment, there are on average 35.8 two-character words containing each of these characters that have ON-readings but only 6.8 words with KUN-readings. There is a similar imbalance favoring ON-readings for the second characters of our stimuli.

### Method

#### Participants

Twenty male and female Japanese speakers (aged between 19 and 25 years), who were brought up in Japan until at least age 18 and had lived in the United Kingdom for less than 18 months, served as participants in this experiment. All had normal or corrected-to-normal vision. Each participant was paid a small fee for participating.

#### Apparatus

The apparatus was the same as in Experiments 1–5.

#### Stimuli

There were 52 two-character Kanji nouns: half of which were ON-reading words and the other half were KUN-reading words. These 52 words consisted of 26 word pairs with the identical first character in common; an example is given below:

ON-reading stimulus word: 相場 *souba* (meaning stock market)  
 KUN-reading stimulus word: 相手 *aite* (meaning opponent).

Both stimulus words have the first character 相 in common. Each pair of words had the same number of morae (three in the example given above, *so-u-ba* and *a-i-te*). The words in each pair are, by definition, perfectly matched for first-character frequency. Also by definition, the words in each pair are unmatched and unmatchable for initial phoneme because the same first character has two different pronunciations in the two words of each pair. Items in the two conditions were matched as closely as possible on whole-word rated familiarity (4.5 vs. 4.4).

Each participant named both ON- and KUN-reading words, but the 52 words were divided into two stimulus lists such that any given participant named either the ON- or the KUN-reading word of each pair but not both. For example, if a participant received *souba* (from the example indicated above) as a target word, then this participant did not see *aite*. We intended for this design feature to avoid any possible effect, either positive or negative, of naming two words with the same first character but two different pronunciations. Therefore, each individual word was named by 10 participants, and each participant saw 26 items.

### Procedure

The procedure was the same as in Experiment 1, except that the participants were given 30 practice trials before the experimental trials.

### Results

Table 7 shows correct mean RTs (in milliseconds) and percentage of error rates for the two conditions.

#### RTs

A one-way ANOVA on mean RTs for correct responses showed no reliable differences between the two conditions:  $F_1(1, 19) = 2.20, MSE = 880, p = .15; F_2 < 1$ .

A multiple regression analysis revealed that the variable of rated word familiarity was at borderline significance,  $t(44) = -1.99, p = .053$ , and that no other variable approached significance. An ANCOVA with second-character frequency, word frequency, familiarity, and imageability as covariates produced adjusted means for the two conditions that were virtually identical: 646 ms for the KUN-reading condition and 642 ms for the ON-reading condition.

#### Errors

A one-way ANOVA on error scores (transformed in the usual way) also showed no reliable differences between the two conditions:  $F_1(1, 19) = 2.51, MSE = 0.224, p = .12; F_2(1, 50) =$

Table 7  
 Mean Reaction Times (in Milliseconds) and Error Rates (ERs; in %) in ON-Reading Versus KUN-Reading Kanji Naming in Experiment 6

Condition	Correct response			
	<i>M</i>	<i>SD</i>	Adj. <i>M</i>	% ER
ON-reading	636	81.64	642	7
KUN-reading	650	86.68	646	4

Note. Adj. = adjusted.

$1.0, MSE = 0.26, p = .31$ . In any case, the nonsignificant numerical difference in errors was in the direction opposite to our prediction. There were too few errors for an informative classification.

### Discussion

This experiment produced essentially identical patterns of performance for ON- and KUN-reading inconsistent nouns matched for initial character. If there were a major role for character-level computation of phonology in Kanji, one would expect the inconsistent words with more typical component pronunciations to have an advantage in speed, accuracy, or both, of naming. The ON-readings of the constituent characters used in this experiment have substantially higher type frequencies (so much greater that they certainly also have higher summed frequencies) and so must be considered more typical; yet there was absolutely no advantage for the ON- over the KUN-reading words. This pattern gives no support to a role for character-level translation from orthography to phonology in Kanji.

### General Discussion

A simplified but reasonably accurate summary of the six experiments reported in this article is that we observed no significant character-to-pronunciation consistency effects in Kanji word naming. This was true (a) for both RT and accuracy measures of naming performance; (b) for ON- and KUN-reading two-character words and one-character KUN-reading words; (c) for lower frequency words as well as common words; and (d) even under several manipulations such as speeded naming, and the inclusion of nonwords, that have been shown to enhance consistency effects in English. There were a few tantalizing hints of the consistency effects that we had expected to observe: In Experiment 3, for example, with blocked presentation of consistent and inconsistent words and with a dichotomized manipulation of target-word familiarity, there was at least a numerical RT disadvantage for the low-familiarity inconsistent words. However, this "effect" failed to reach significance either as a main effect of consistency or as a Familiarity  $\times$  Consistency interaction; the only statistically reliable effect on RT in this as in other experiments was familiarity.

These results in Kanji differ from those for English discussed in the introduction (e.g., Glushko, 1979; Jared et al., 1990, and many others). In computation of phonology for an English target word, the degree of consistency of pronunciation across words sharing orthographic segments with the target influences the efficiency of the computation. Moreover, this impact is appropriately "graded": Inconsistent words with many high-frequency pronunciation enemies suffer more than words with few or only uncommon enemies with many friends to support their particular spelling-to-sound relationships. In Kanji, however, whether each constituent character has one or more other pronunciations, and also whether the alternative pronunciations apply to other two-character Kanji words, had no significant impact on naming latencies for the stimulus words. As noteworthy as the absence of an RT effect was the lack of word-naming errors that would, according to our predictions, have been a clear signature of consistency effects. In English word-naming experiments, the difference between regular and 'irregular' words often reveals itself in error rates

as well as in RTs. The most common error to an irregular English word is the assignment of an alternative, more common pronunciation of the inconsistent segment (typically the body or rime); and the likelihood of such regularization errors can be augmented by manipulations such as speeded naming instructions (Strain et al., 1995) or inclusion of nonwords in the stimulus set (Midgley-West, 1979; Monsell et al., 1992). In contrast, our experiments yielded almost no Kanji naming errors that would be the equivalent of assigning an alternative pronunciation to an inconsistent segment, that is, alternative ON-KUN reading errors. Across all of the stimulus materials in all of the experiments that would permit ON-KUN substitutions, such errors accounted for fewer than 1% of responses, whereas typical rates of regularization errors to low-frequency inconsistent English words were in the range of 6%–15% (see for example, Glushko, 1979; Paap & Noel, 1991).

The lack of significant consistency effects for Kanji is different not only from English but from word-naming data in other languages as well, for example, Italian (e.g., Colombo, 1992) and, more strikingly, Chinese (Hue, 1992; Yin, Butterworth, & Patterson, 1995). Most Chinese characters are phonograms with two components: the radical, which represents a broad semantic field (i.e., a clue to the meaning of the character), and the phonetic component, which typically provides a clue to the pronunciation of the character. Within these phonograms, one can distinguish—somewhat as in English—variables of both regularity (or congruency) and consistency. Regular or congruent characters are those for which the pronunciation of the character as a whole corresponds exactly to the pronunciation of the phonetic component; for irregular or incongruent characters, the pronunciation of the whole character and of its phonetic component do not match. Consistency, as in other writing systems, refers more to the characteristics of pronunciation across an orthographic neighborhood. Consistent characters are those for which all characters containing a particular phonetic component have the same pronunciation; for inconsistent characters, one or more exemplars sharing the same phonetic component have discrepant pronunciations. There is a further group of characters, sometimes called *unique*, that do not contain phonetic stems.

Hue (1992, Experiment 2) demonstrated that, within the set of low-frequency experimental stimuli, regular (congruent) characters were named more rapidly and more accurately than either irregular or unique characters. Even more germane to the present article, both Hue (1992, Experiment 3) and Yin et al. (1994) obtained marked effects of consistency, with inconsistent characters yielding slower RTs and more errors than consistent characters. The errors were mainly pronunciations appropriate to a different character containing the same radical. Note that this is a true neighborhood effect: In both experiments, the correct pronunciation of an inconsistent character was always congruent with that of its phonetic component, so there was no conflicting alternative pronunciation within the character itself; the conflict is attributable solely to the fact that other characters containing that same phonetic component have different pronunciations. In other words, this is an approximate equivalent of a disadvantage for the naming of regular inconsistent words in English (e.g., *bone*, which is regular but has the pronunciation enemies *done* and

*gone*). The fact that these consistency effects were not only statistically reliable but of substantial magnitude in both Chinese experiments makes the contrast with our nonresults for Japanese Kanji even more noteworthy and surprising. It must be remembered, however, that characteristics of the spelling-sound relationship in Kanji are different not only from alphabetic writing systems but also from Chinese. For example, although there may be pronunciation ambiguity in Chinese characters either within a character (if the whole character and its phonetic component are incongruent) or across neighborhoods of characters sharing a phonetic component, each whole character has only one correct pronunciation that is invariant across the set of intraword contexts in which the character may occur.

It should be noted that, in contrast to the multimoraic Kanji words used here, the majority of studies on consistency effects in English have used monosyllabic words as stimuli, possibly because consistency in English is often defined with respect to “body” neighborhoods, which are much easier to manipulate in monosyllabic words. We consider it unlikely that our failure to observe consistency effects for Kanji is attributable to the use of multimoraic (syllabic) words because standard consistency effects established for monosyllabic words have also been obtained for multisyllabic words in both English (Monsell et al., 1992; Strain et al., 1995) and Italian (Colombo, 1992). We acknowledge, however, that evidence on multisyllabic words (at least of more than two syllables) is sparse for English and that word length might be one of several important differences preventing an unambiguous interpretation of apparent differences in the impact of consistency manipulations in Kanji as compared with English (and other alphabetic orthographies).<sup>2</sup>

We turn now to a consideration of what factor(s) in the characteristics of Kanji may be responsible for its lack of susceptibility to consistency effects. The first important thing to establish in this regard is that we do not propose, as some theorists have speculated, that words written in Kanji might be named by a different set of processes from those applying to alphabetic words, that is, primarily by accessing meaning from orthography and then phonology from meaning. This is very unlikely to be the case. Japanese children learn to compute phonology directly from Kanji orthography, and indeed this direct transcoding continues to play an essential role in adult Japanese reading, as indicated by various forms of experimental evidence including (a) reliable homophone effects in semantic classification of Kanji words (Wyde et al., 1993) paralleling homophone effects shown for English by Van Orden (1987) and others and (b) adequate naming of Kanji words by neurological patients with impaired comprehension of the same words (Sasanuma, Sakuma, & Kitano, 1992). We assert that the difference between reading in Kanji and other writing systems is not in the role played by phonological codes but rather in the details of the process by which phonological codes are computed (see Perfetti, Zhang, & Berent, 1992, for a similar conclusion).

The obvious interpretation is that although readers of other languages learn orthographic-to-phonological computations

<sup>2</sup> We are grateful to an anonymous reviewer for raising this point.

that generalize over commonly recurring subword patterns, readers of Japanese Kanji learn that the only reliable level of generalization is the whole word. As emphasized throughout this article, the correct pronunciation of a subword component in Kanji (the individual character) is highly dependent on intraword context. Of course, to some extent, the same is true in English; but two factors about English (and other alphabetic scripts) might still encourage generalization at the level of subword units. The first is that, for most of the inconsistent spelling patterns in English, there is one dominant spelling-sound correspondence. Setting weights on connections to take account of the dominant pattern may hurt the occasional word (pronunciation of *caste* will be slowed or even incorrect because the participant computes the pronunciation "kaist" as well as "cast" as the result of the strength of the correspondence between *aste* and "aist" across the family of words including *haste*, *waste*, etc.); however, this organization will, across the vocabulary as a whole, provide more benefit than harm simply because the dominant spelling-sound correspondence does most often apply. Across single- and multiple-character words, neither the ON nor the KUN pronunciation of a Kanji character can really be considered dominant; therefore, perhaps there would be no comparable benefit to subword generalization. The second factor is one often emphasized in discussions about subword levels of computation in English (see, for example, Norris & Brown, 1985): The great majority of irregular English words deviate from a more typical, rule-governed pronunciation only in one or two phonemes, typically the vowel (e.g., the regularized pronunciations of *caste* and *pint* are only one phoneme different from the correct pronunciations). If much, even though not all, of what the reader computes at the subword level is applicable to the final pronunciation, then once again it may be more helpful than harmful. By contrast, the two (or more) alternative pronunciations of a Kanji character typically have no phonological elements in common (cf. the ON- and KUN-readings of the character 親, which are "oya" and "shin"). In this case, it may be more harmful than helpful to compute both pronunciations of a component character on the path to a phonological representation for the whole word, which is the ultimate goal of this computation.

The fact that several of the naming experiments reported in this article produced reliable effects of individual-character frequency, with faster naming latencies to words containing more commonly encountered Kanji characters, might seem to be at odds with our conclusion that the computation of phonology from orthography is dominated by the word level. We do not, however, think that this result is in conflict with our interpretation for the simple reason that there is another plausible locus of the character frequency effect at the stage of orthographic analysis. There is only a small number of alphabetic characters (26 in English, even fewer in Italian, for example), all of which are relatively simple visual forms and—despite differences in frequency of occurrence (compare *s* and *z*, for example)—every one of which would be encountered in reading almost any text. By contrast, there are approximately 3,000 commonly used Kanji characters, some with rather complex visual forms. On the reasonable assumption that the efficiency of early orthographic analysis of a Kanji word will be modulated by the familiarity and complexity of its

characters as orthographic patterns, the character frequency effect need have no direct implications for a model of phonological processing of Kanji words.

One aspect of our results that seems to fit the proposal of a mainly whole-word basis for computation of Kanji phonology is the massive lexicality advantage for RTs observed in Experiment 4. It would of course be impossible to offer a single value for a lexicality effect in experiments with English readers, as the effect will clearly vary as a function of the characteristics of the words, the nonwords, the reading skill of the participants, and so forth. Several studies in the literature, however, in which English words and matched nonwords were named in mixed lists (as in the design of Experiment 4 for Kanji) suggest a typical RT difference in the region of 30–50 ms (see, for example, Glushko, 1979; Monsell et al., 1992). In our Experiment 4, the corresponding difference was about 180 ms. It is also pertinent that latencies to name English nonwords are predicted by the degree of pronunciation ambiguity: Seidenberg, Plaut, Petersen, McClelland, and McRae (in press) have demonstrated that nonwords like *zale* and *pift*, which are given one consistent pronunciation by all participants, yield significantly faster RTs than even the most common pronunciation of nonwords given two or three different readings by various participants. The Kanji nonwords in Experiment 4, composed of characters with only a single ON-reading, had unambiguous pronunciations and yet yielded much slower naming RTs than comparable words. This result can be interpreted as evidence that computing the phonology of each component character independently, as is required to name a Kanji nonword, plays relatively little role in translating familiar Kanji words from orthography to phonology.

The idea that processing in Kanji relies primarily on a whole-word level finds some general support in a study by Morton, Sasanuma, Patterson, and Sakuma (1992), who investigated recognition units for Kanji words by using a different paradigm to the current naming experiments. In that study, target words to be identified tachistoscopically were preceded, about half an hour earlier and in a different task, by words in one of several critical conditions: (a) an identity prime (the prime and target words were identical), (b) a single-character prime (the target was a two-character word, and the prime was one of its component characters), and (c) a two-character prime (in which the target was either a single-character word corresponding to one of the components of the prime or a different two-character word sharing a character with the prime). Relative to appropriate control conditions, Morton et al. found a strong identity priming effect but no cross-facilitation between compound words sharing a single character or from a single character to a compound word that contained it or from compound words to their component characters. From this pattern of data, Morton et al. (1992) concluded that Kanji words are recognized at the word rather than the character level.

The transcoding between phonology and meaning must be dominated by whole-word level representations in all languages because, except for morphological variations, there is no useful level of subword generalization in the correspondences between phonology and meaning: Words that sound similar do not have similar meanings. Although we had not

predicted these results, our findings for the computation between orthography and phonology in Kanji suggest a similar picture to that between phonology and meaning. In English and other alphabetic writing systems, words that are orthographically similar almost always have similar pronunciations; this is not true in Kanji, and the lack of consistency effects in Kanji word naming may be a logical consequence of this difference.

### References

- Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition*, *10*, 565-575.
- Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 802-814.
- Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 234-254.
- Brown, G. D. A. (1987). Resolving inconsistency: A computational model of word naming. *Journal of Memory and Language*, *26*, 1-23.
- Colombo, L. (1992). Lexical stress effects and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 987-1003.
- Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of human information processing* (pp. 151-215). London: Academic Press.
- Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 674-691.
- Gum, T., & Bub, D. (1988). *PsychLab software*. Montreal, Quebec, Canada: Montreal Neurological Institute.
- Hue, C. W. (1992). Recognition processes in character naming. In H.-C. Chen & Ovid J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 93-107). Amsterdam: North-Holland.
- Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency: Effects in word naming. *Journal of Memory and Language*, *29*, 687-715.
- Midgley-West, L. (1979). *Strategies in reading*. Unpublished doctoral dissertation, Birkbeck College, London, England.
- Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118*, 43-71.
- Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H., & Milroy, R. (1992). Lexical and sub-lexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 452-467.
- Morton, J., Sasanuma, S., Patterson, K., & Sakuma, N. (1992). The organisation of the lexicon in Japanese: Single and compound Kanji. *British Journal of Psychology*, *83*, 517-531.
- National Language Institute. (1976). Gendai shinbun no Kanji [A study of uses of Kanji characters in modern newspapers] (Report No. 56). Tokyo: Shuei Shuppan.
- Norris, D., & Brown, G. (1985). Race models and analogy theories: A dead heat? *Cognition*, *20*, 155-168.
- Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research*, *53*, 13-24.
- Parkin, A. J. (1982). Phonological recoding in lexical decision: Effects of spelling-to-sound regularity depend on how regularity is defined. *Memory & Cognition*, *10*, 43-53.
- Patterson, K., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance: XII. The psychology of reading* (pp. 421-447). Hove, England: Erlbaum.
- Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a "universal" phonological principle. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 227-248). Amsterdam: North-Holland.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1994). *Understanding normal and impaired word reading: Computational principles in quasi-regular domains* (Tech. Rep. No. PDP.CNS.94.5). Pittsburgh, PA: Carnegie Mellon University, Department of Psychology.
- Sampson, G. (1985). *Writing systems*. Stanford, CA: Stanford University Press.
- Sasanuma, S. (1986). Universal and language-specific symptomatology and treatment of aphasia. *Folia Phoniatrica*, *38*, 121-175.
- Sasanuma, S., Sakuma, N., & Kitano, K. (1992). Reading Kanji without semantics: Evidence from a longitudinal study of dementia. *Cognitive Neuropsychology*, *9*, 465-486.
- Seidenberg, M. S., & McClelland, J. (1989). A distributed model of word recognition and naming. *Psychological Review*, *96*, 523-568.
- Seidenberg, M. S., Plaut, D. C., Petersen, A. S., McClelland, J. L., & McRae, K. (in press). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology: Human Perception and Performance*.
- 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*, 393-404.
- Shallice, T., Warrington, E. K., & McCarthy, R. (1983). Reading without semantics. *Quarterly Journal of Experimental Psychology*, *35A*, 111-138.
- Stanhope, N., & Parkin, A. J. (1987). Further explorations of the consistency effect in word and nonword pronunciation. *Memory & Cognition*, *15*(2), 169-179.
- Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140-1154.
- SweetJAM 4.5 [computer software]. (1990). Tokyo, Japan: A & A Company, Ltd.
- Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language*, *26*, 608-631.
- Toyo Kanji jiten* [Toyo Kanji dictionary]. Y. Abe (Ed.). (1962). Tokyo: Obunsha.
- Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound and reading. *Memory & Cognition*, *15*, 181-198.
- Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 664-680.
- Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition*, *13*, 557-572.
- Wyde, T. (1991). *Processes in the reading of Japanese: Comparative studies between English and Japanese orthographies*. Unpublished doctoral dissertation, University of London, London, England.
- Wyde, T. N., Patterson, K. E., & Humphreys, G. W. (1993). Phonologically mediated access to meaning for Kanji: Is a ROWS still a ROSE in Japanese Kanji? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 491-514.
- Wyde, T. N., Quinlan, P., & Butterworth, B. (1995). *Japanese lexical database: 2,357 Japanese nouns rated on frequency, familiarity and imageability and other language statistics on Kanji characters/words*. Manuscript submitted for publication.
- Yin, W., Butterworth, B., & Patterson, K. (1995). *Consistency and congruency in naming Chinese characters*. Manuscript submitted for publication.

Received April 12, 1994

Revision received November 16, 1994

Accepted November 23, 1994 ■