



Mora or more? The phonological unit of Japanese word production in the Stroop color naming task

Rinus G. Verdonschot¹ · Sachiko Kinoshita²

Published online: 6 December 2017
© Psychonomic Society, Inc. 2017

Abstract

In English, Dutch, and other European languages, it is well established that the fundamental phonological unit in word production is the phoneme; in contrast, recent studies have shown that in Chinese it is the (atonal) syllable and in Japanese the mora. The present study investigated whether this cross-language variation in the size of the unit of word production is due to the type of script used in the language (i.e., alphabetic, morphosyllabic, or moraic). Capitalizing on the multiscriptal nature of Japanese, and using the Stroop color naming task, we show that the overlap in the initial mora between the color name and the written distractor facilitates color naming independent of script type. These results confirm the mora as the phonological unit of word production in Japanese, and establish the Stroop color naming task as a useful task for investigating the fundamental (or “proximate”) phonological unit used in speech production.

Keywords Language production · Psycholinguistics

Many core processes underlying language production are presumed to be universal (e.g., Gómez et al., 2014) and should therefore not depend on any language. Models dealing with phonological encoding (i.e., creating abstract representations of the sound form of the word) often reflect this assumption and occasionally overlook variability across languages. Many models assume that the phoneme is the unit underlying word production (e.g., Dell, 1986; Caramazza, 1997; Levelt, Roelofs, & Meyer, 1999; but see Roelofs, 2015). One reason for this assumption is that most psycholinguistic research has concentrated on a few (European) languages (e.g., English, Dutch).

Within language production models, Levelt et al. (1999) is most explicit on how phonological encoding takes place. Specifically, during phonological encoding, two pieces of information are needed: (1) a “metrical frame” containing the number of syllables of a word as well as its main stress

position and (2) the units (phonemes) necessary to fill the syllabic structure. Subsequently, this information is combined by incrementally inserting the activated segments (phonemes) into the corresponding syllables outlined in the metrical frame in a segment-to-frame association process (prosodification). For example, to produce the word *tiger* the metrical frame would specify $\omega = \sigma' \sigma$ (i.e., one word [ω] with two syllables [σ] with stress [$'$] on the first syllable) in which six phonemes (taɪgəɹ) would have to be inserted to create /taɪ gəɹ/.

Most evidence supporting the assumption that the phoneme is the unit used in phonological encoding during speech production comes from experiments using the implicit priming (“form preparation”) paradigm (e.g., Meyer, 1990, 1991). In this paradigm, participants first learn a small set of semantically related word pairs, called prompt and response, respectively, for example: *fruit*–*pear*. After learning these word pairs, participants are asked to produce the response word (e.g., /pɛəɹ/) when a prompt word (e.g., *fruit*) is presented. The critical manipulation involves the word-pair groupings, such that in the homogeneous condition the response words overlap in certain features, such as the first phoneme (e.g., /pɛəɹ/, /pɒnd/, /pɒn/), and in the heterogeneous condition they do not (e.g., /pɛəɹ/, /reɪs/, /taʊn/). The two critical results observed with this task are that a segmental (phoneme) overlap between responses significantly facilitates response latencies (i.e., homogeneous < heterogeneous), and that this segmental overlap advantage is only observed for beginning-related

✉ Rinus G. Verdonschot
rinusverdonschot@gmail.com

¹ Department of Oral and Maxillofacial Radiology, Institute of Biomedical & Health Sciences, Hiroshima University, Hiroshima, Japan

² Department of Psychology and ARC Centre of Excellence for Cognition and its Disorders, Macquarie University, Sydney, Australia

and not end-related overlap (e.g., not /dip/, /lup/, /hip/), attesting to the unit size as well as the left-to-right incrementality.

Additional evidence for a phoneme-sized phonological unit comes from a different paradigm involving reading aloud words in combination with masked priming. Here, the task is to read aloud words presented on the screen. Just before the target word, another word (the “prime”) is presented very briefly (50 ms) such that participants are unaware of its presence. Forster and Davis (1991) showed that target words are read aloud faster if the prime shares just the initial letter with the target (e.g., prime *pear*, target *pond*) relative to an unrelated control prime (e.g., *town*). Like the implicit priming paradigm, when prime and target were end related (e.g., *door-pear*), facilitation does not occur (Kinoshita, 2000). The masked onset priming effect has been observed consistently across several European languages (English: Forster & Davis, 1991; Kinoshita & Woollams, 2002; Dutch: Schiller, 2004; and Spanish: Dimitropoulou, Dunabeitia, & Carreiras, 2010).

Cross-language variation in the phonological unit used in speech production

The brief review above indicated that for European languages there is an empirical consensus for the phoneme as the phonological unit in speech production. Recently, however, evidence has accumulated that this basic unit may differ between languages, with recent studies finding differently sized units for Mandarin Chinese (e.g., Chen, Chen, & Dell, 2002; O’Seaghdha, Chen, & Chen, 2010) and Japanese (e.g., Kureta, Fushimi, & Tatsumi, 2006; Verdonschot et al., 2011). Using the implicit priming paradigm, Chen et al. (2002) found that in Mandarin Chinese, significant preparation effects were only found when the whole (atonal) syllable overlapped. Specifically, for response words such as *da1ying4*, *de2guo2*, *du3buo2*, respectively (number indicates tone), the onset phoneme overlap did not result in significant facilitation, but response words such as *ke1xue2*, *ke2tan2*, *ke3lian2*, that is, words with overlapping initial syllable (irrespective of tone), did produce facilitation. Similarly, using the masked priming read-aloud task, it was shown that native Mandarin speakers do not show onset segment priming (e.g., Verdonschot, Lai, Feng, Tamaoka, & Schiller, 2015), but they do show initial syllable priming (e.g., Chen, Lin, & Ferrand, 2003; You, Zhang, & Verdonschot, 2012). Recently, Chen, O’Seaghdha, and Chen (2016) used masked primes with picture targets (thereby avoiding any reading processes involving the target) and found that pictures (e.g., 太陽 /tai4yang2/) were named significantly faster (compared to control) when preceded by a syllable-related prime (e.g., 台 /tai2/), though significant priming was not obtained for onset-related primes (e.g., 圖 /tu2/). Similar results indicating the absence of onset segment overlap effect in Cantonese Chinese were reported by

Wong and Chen (2008, 2009) using the picture-word interference task, and in the implicit priming task (Wong, Huang, & Chen, 2012).

In Japanese, the phonological unit for speech production is considered to be the mora (e.g., Kubozono, 1989). The mora is a rhythmical unit typically consisting of CV (e.g., /ka/) or V (e.g. /a/) but never C alone (except for the nasal coda /N/); also, vowel elongation and geminates are counted as one mora. For example, the word *Japan* in Japanese is /ni.ho.N/, which has three moras (a dot denotes the mora boundary). In the comprehension literature, the importance of the mora has been well established (see Otake, Hatano, Cutler, & Mehler, 1993; who obtained moraic segmentation effects for Japanese words). In the production literature, Kureta et al. (2006) used the implicit priming paradigm and found preparation effects for response words when the initial mora overlapped, for example: かつら, 歌舞伎 and 鞆 (i.e., /ka.tsu.ra/, /ka.bu.ki/, /ka.ba.N/, respectively), but not when the onset segment overlapped, as in かつら, くじら, 古墳 (i.e., /ka.tsu.ra/, /ku.dzi.ra/, /ko.φu.N/, respectively). Corroborating evidence comes from masked priming read-aloud experiments showing that targets (e.g., まんが /ma.n.ga/) preceded by onset segment-related primes (e.g., メイ /me.e/) did not show faster naming latencies; however, initial mora-related primes (e.g., マメ /ma.me/) did show priming compared to unrelated primes (Verdonschot et al. 2011).

Is the cross-language variation in the unit of speech production due to script type?

An important question arising from these studies is whether the cross-language variations in unit size reflect the difference in the script type. In alphabetic scripts, each letter, or letter cluster, maps onto a phoneme. In contrast, in Chinese a character maps onto a morpho-syllable, and in Japanese a kana maps onto a mora. Is the size of phonological unit of word production determined by the unit size of mapping between the script and phonology? That is, it is possible that phonemes constitute the unit in English, Dutch, and other European languages because they are written in the Latin alphabet; syllables constitute the unit in Chinese because Chinese characters correspond to a syllable, and mora constitute the unit in Japanese because each kana maps onto a mora.¹ As we discuss below, there is currently no clear answer.

There are two approaches to investigate this question. One is to ask whether the cross-language variation is a consequence of acquiring knowledge of the writing system used to write the language. This possibility is suggested by research

¹ Li, Wang and Idsardi (2015) asked exactly this question and concluded in the affirmative, based on an implicit priming study with Mandarin Chinese speakers.

on phoneme awareness—the ability to perceive and manipulate phoneme segments in spoken words—as tested, for example, in the phoneme deletion task (e.g., “Say what remains after you take away the sound /s/ from *sand*”). It is well-established that phonemic awareness develops with reading skill (see Castles & Coltheart, 2004). Although the direction of causality of this relationship is intensely debated, with the original idea being that phonemic awareness is a prerequisite for acquisition of literacy (see Castles & Coltheart, 2004, for discussion), there is evidence suggesting that it is the acquisition of literacy that leads to phoneme awareness. For example, Morais, Cary, Alegria, and Bertelson (1979) showed that Portuguese adults who were illiterate performed more poorly on phoneme awareness tasks than comparable people who had been illiterate as adults but had subsequently learned to read. Of relevance to the issue of cross-language variation; Read, Zhang, Nie, and Ding (1986) reported that Chinese adults who are only literate in Chinese characters (and not pinyin, the alphabetic transcription of Chinese) could not add or delete individual consonants in spoken Chinese words, and thus concluded that the phoneme awareness develops “in the process of learning to read and write *alphabetically*” (Read et al. 1986, p. 32). From this perspective, the findings of cross-language variation in unit size in the speech production tasks could be similarly interpreted in terms of the difference in the script type used in the language, because all previous studies involved literate adult participants.

From this perspective, the way to investigate whether the variation in script type is causally related to the cross-language variation in the proximate unit used in speech production is to compare illiterate adults (or preliterate children) who are speakers of languages with different proximate units (e.g., English vs. Chinese) on a speech-production task that does not require the presentation of written input, such as the implicit priming task. We are not aware of such studies,² which is perhaps not too surprising given the obvious difficulties associated with this approach.

The second approach is to investigate whether there is an influence of orthography on a spoken word-production task. For the implicit priming task, under some conditions, the answer is yes. Using English words, Damian and Bowers (2003) reported that when onset segments shared by the response words (e.g., /k/) were spelled differently (e.g., *camel*, *kidney*), the advantage in response latency relative to the heterogeneous set was eliminated. These authors found this effect of spelling even when the (prompt words and) response words were spoken in the learning phase and hence the orthography was not explicit, and concluded that “orthographic codes are

mandatorily activated in speech production by literate speakers” (p. 119). However, using Dutch words, Roelofs (2006, Experiment 3) found no spelling influence, contrasting with Damian and Bowers (2003). Alario, Perre, Castel, and Ziegler (2007) tested the possibility that this discrepancy between Damian and Bowers’ (2003) and Roelofs’ (2006) might have been due to cross-language differences in the degree of consistency of phonology-orthography mapping. Using picture stimuli named in French, they found no evidence for the effects of spelling differences and concluded that the discrepancy between the results reported by Damian and Bowers (2003) and Roelofs (2006) could not be explained in terms of language, because French, like English, is less consistent than Dutch. Instead, they concurred with Roelofs (2006) who suggested that the effect of spelling reported by Damian and Bowers (2003) might not bear directly on the speech production processes but rather to the memorization processes recruited by the prompt-response learning procedure (“Perhaps Damian and Bowers’s participants encoded the initial sounds of the responses in both orthographic and phonological forms to help maintain them in memory, thereby hampering response preparation in inconsistent sets,” Roelofs, 2006, p. 34).

In addition to the above experiment in which prompt and response words were not presented as written words, Roelofs (2006, Experiments 1 and 2) also used a design that did not require the learning of prompt-response pairs, and only the response words had to be produced at test. He found a spelling influence when the response words were read (Experiment 1) but not when the response words were presented as to be named pictures (Experiment 2). Roelofs thus concluded that “the spelling of a word constrains spoken word production in Dutch only when it is relevant for the word production task at hand” (p. 33). In sum, there may be an influence of orthography on the implicit priming task (1) when the response words are presented visually as written words and (2) even when the written form is not presented (which is not always the case), and when it is present, the effect of (implied) orthography seems to be on the learning of prompt-response pairs rather than on the speech production process per se.

It is relevant to note in this context that recently two separate studies have reported using the implicit priming task with alphabetic transcriptions of Chinese and Japanese words and showed the preparation effect with onset segment with (Mandarin) Chinese speakers and Japanese speakers. Li et al. (2015) showed the preparation effect for onset segments when pinyin is used. Similarly, Kureta, Fushimi, Sakuma, and Tatsumi (2015) reported that when the response words were presented in romaji (romanized Japanese) in the prompt-response learning phase, the preparation benefit was found for the shared onset segments, and they concluded that Japanese speakers “can prepare phonemic segments in special circumstances in which the orthographic representation of

² There are studies that tested preliterate Japanese children which showed that they can segment speech signals at the boundary of morae (e.g., Ito & Tatsumi, 1997). Although clearly relevant, these findings do not speak directly to the size of proximate unit used in speech production.

romanized Japanese is first activated” (p. 50). It should be noted that both pinyin and romaji are not native scripts, and not normally encountered in text. Note also that Kureta et al.’s (2015) finding is at odds with the absence of onset segment priming effect with romaji stimuli reported by Verdonschot et al. (2011) with native Japanese speakers. Li et al. (2015) also pointed out a caveat that their participants were all undergraduate or graduate students at an American university and had extensive exposure to English, and hence “their higher English language proficiency may have facilitated them to attend more to a small phonological unit such as the onset” (p. 575).³ In sum, although the two implicit priming studies show that in the implicit priming task Mandarin speakers and Japanese speakers can use the onset segments to prepare speech production when words are presented in alphabetic script, they do not speak to the key question at hand, namely, whether cross-language variation in the phonological unit used in speech production is due to the difference in native script type.

The above considerations with non-native scripts highlight yet another issue. Both European languages, like English and Dutch, and Chinese are associated with only one type of native script; thus, for these languages, there is no way of disentangling the role of native script type from the phonological unit of word production used in the language.

Japanese writing systems: Kanji and kana

Japanese, however, provides an interesting exception as it is multiscriptal—that is, Japanese is written using both a mora-based system (kana) and logographic kanji characters. There are two types of Japanese kana: hiragana and katakana (each consisting of 46 base characters). Hiragana is typically used to write inflections and other grammatical markers and Japanese words for which kanji are rare or unfamiliar (e.g., かぼち /ka.bo.tɕa/ *pumpkin*); whereas katakana is mostly used to write foreign or loan words (e.g., トマト /to.ma.to/ *tomato*) as well as scientific or technical terms. Japanese children are formally taught kana around the age of 6 years, in the first year of elementary school, though many children know (some) kana prior to starting school.

Japanese kanji are logographs typically representing content words such as nouns and (stems of) adjectives and verbs (e.g., 赤 /a.ka/ *red*). Starting from elementary school, all Japanese children must learn a fixed number of kanji during each school year resulting, in the accumulation of 1,006 kanji at the end of elementary school. The rest of the official (jōyō) kanji, prescribed

by the Japanese Ministry of Education (see Tamaoka, Makioka, Sanders, & Verdonschot, 2017), will be learned in high school, ultimately resulting in the knowledge of 2,136 kanji (though many Japanese would know more kanji than this).

Japanese kanji are derived from Chinese hanzi; however, unlike Chinese, where a character is a monosyllabic morpheme, a Japanese kanji character is a morpheme that may be multimoraic (i.e., of varying length). For example, the kanji character 菊 (meaning “chrysanthemum”) is pronounced /ki.ku/, the character 娘 (“daughter”) is pronounced /mu.su.me/. This means that it would be possible to test in Japanese empirically whether the status of the mora as the unit of word production is dependent on the nature of the (native) script used.

Kureta et al. (2006) considered this issue, and concluded that “the moraic structure is derived phonologically”, pointing out that “the emergence of the moraic structure—for example, learning to segment speech signals at the boundary of morae—is known to precede the onset of literacy” (p. 1112). However, they did not provide direct evidence that the mora as the unit of Japanese speech production is independent of the script used. Citing the work by Damian and Bowers (2003), in English, described above, which showed that the preparation effect disappeared when the onset segment shared by the response words (e.g., /k/) was spelled differently (e.g., “c” or “k”), Kureta et al. (2006) acknowledged that participants in the implicit priming paradigm may perform the task “by keeping orthographic information in mind, given that the to-be-reproduced materials are presented visually during the learning phase” (p. 1112). They further pointed out that the kanji words they used shared phonology but not spelling (like Damian & Bowers, 2003), and argued that “if our participants depended exclusively on the same orthographic code as actually presented in the memorization phase, there should have been no facilitation at all for all the homogenous contexts” for these stimuli. However, Kureta et al. used a mix of kanji stimuli and kana stimuli (i.e., their “homogenous” set might have been inconsistent), and they were not examined separately—thus, it is not known whether the initial mora overlap effect was observed with the kanji stimuli as well as the kana stimuli.

Recently, using the masked priming read aloud task, Yoshihara, Nakayama, Verdonschot, and Hino (2017) reported finding exactly this dissociation between kanji and kana. They used two-kanji compound words in which the first character was multimoraic, and manipulated the overlap in the initial mora (e.g., prime: 発案 /ha.tsu.a.n/ – target: 博物 /ha.ku.bu.tsu/ versus prime: 立案 /ri.tsu.a.n/ – target: 博物 /ha.ku.bu.tsu/). Unexpectedly, they found that an overlap in the initial mora (as in the first example above) was not sufficient to produce priming; only when the whole pronunciation of the first kanji character overlapped (e.g., /ha.ku/ in 迫害 /ha.ku.ga.i/ – 博物 /ha.ku.bu.tsu/) significant priming was observed. When stimuli were transcribed in kana (e.g., ハツアン /ha.tsu.a.n/ – はくぶつ /ha.ku.bu.tsu/), initial

³ Indeed, this possibility is supported by the finding of Nakayama, Kinoshita and Verdonschot (2016), in which an onset segment overlap effect (bark-BENCH < dark-BENCH) emerged in high-proficient but not low-proficient Japanese–English bilinguals reading aloud English words.

mora priming emerged, indicating that the nature of script modulated the unit size of phonological priming.

What is not clear, however, is whether this script type effect reflects the modulation of the phonological unit of speech production, or an effect on reading aloud a written word. Although generating a name from a concept (as occurs in spontaneous speech or in picture naming) and reading aloud words both involve speech production processes, how the phonological units comprising the name (be it phonemes, syllables, or morae) are generated from input are different. Specifically, while a picture name must be retrieved as a whole, the phonological form of a written word can be generated, in a piecemeal fashion, via script to sound mapping. In fact, the original account of masked onset priming effect (Forster & Davis, 1991) and, later, the dual route cascaded (DRC) model of reading (e.g., Mousikou, Rastle, Besner, & Coltheart, 2015) situates its origin solely in print-to-sound mapping, prior to the phonological encoding process. Yoshihara et al. (2017) thus noted that their finding “demonstrates that the difference in script type does matter at least when we read a word aloud,” (p.1303) leaving open the possibility that its effects may differ in other speech production tasks.

The present study

In summary, there are data indicating cross-language variation in the unit size of spoken word production, but in almost all cases, there is a perfect correspondence between the nature of script (alphabetic, morphosyllabic, moraic) and the phonological unit (phoneme, syllable, mora) used in speech production of the language. A priori, with literate adult participants, the role of script type cannot be teased apart from the phonological unit of speech production in languages that are written in only one type of script.

As stated, Japanese is unique in that words are written using both kana, which maps onto a mora, and kanji, which are morphemes that may be multimoraic. The question we ask is whether the mora-sized unit of speech production in Japanese is dependent on script type. The answer from a recent masked priming read-aloud study is positive (Yoshihara et al., 2017). However, the task here was to read aloud the target word, and it is unclear whether the influence of script type generalizes to other speech production tasks. Therefore, we turn to a different task, the Stroop color naming task using word distractors written in kana and kanji.

In the Stroop color naming task, participants are asked to name the color in which a word distractor is displayed (Stroop, 1935). The standard Stroop congruence effect is the finding that color naming is faster when the word distractor is congruent with the response color than when it is incongruent. In a variation of this task, it has been shown that color naming is faster when the word (and pseudoword) distractor overlaps

with the response color name in the onset segment (e.g., the word RUN or the pseudoword ROZZ presented in red) than when it does not (e.g., the word FUN or the pseudoword FOZZ presented in red (Coltheart, Woollams, Kinoshita, & Perry, 1999; Mousikou et al., 2015), an effect referred to as the phonological Stroop effect. In the present study, we will use the phonological Stroop effect as a tool to study the phonological unit of speech production.⁴

There are several reasons why we chose the Stroop color naming task for investigating the influence of script type on the phonological unit of speech production. One is that, here, the goal of the task is to simply name the color, and it does not require the learning of prompt–response pairs. This is important, because the orthographic properties of the response words could influence the learning of prompt–response pairs (see Alario, et al., 2007; Roelofs, 2006). Second, in the Stroop color naming task, because the written word distractor is never the target to be responded to, there is no reason to intentionally read the word. Recall that Roelofs (2006, Experiments 1 and 2) has shown that even in the absence of the requirement to learn the prompt–response pairs, the form preparation effect was disrupted by spelling inconsistency when the response words had to be read (but not when they were presented as pictures). Roelofs (2006) suggested that the spelling of a word constrains spoken word production “when it is relevant for the word production task at hand” (p. 36). Spelling is clearly relevant to the masked priming read aloud task, in which a written target word is to be read aloud. In contrast, in the Stroop color naming task, the target to be named is the color in which the word distractor is displayed; the word distractor is a priori irrelevant to the task at hand.

We start by using kana stimuli in Experiment 1, to confirm the finding that in Japanese speech production the initial mora overlap, but not the onset segment overlap facilitates naming. This establishes the Stroop color naming task as a suitable tool for investigating the phonological unit of speech production. In two subsequent experiments, we contrast the effects of initial mora overlap with word distractors written in multimoraic kanji and kana to test the effects of script type.

Experiment 1: Stroop task (C and CV overlap) using katakana distractors

As far as we know, the Stroop color naming task has not been used to investigate the phonological unit of speech production

⁴ We hasten to add that we are not the first to use this task to study speech production. For a major theory of Stroop color naming performance couched within the WEAVER++ model of speech production, see Roelofs (2003). We note in particular that Roelofs (2003) assumes the classic Stroop interference effect arises during the lemma retrieval process and the word-form (phonological) encoding (see Fig. 7, p. 98). For word distractors that are unrelated to color used in the investigation of the phonological Stroop effect, there is no reason to expect lemma retrieval to be involved, leaving the latter as the only possible locus.

in Japanese. Thus, the main aim of this first experiment was to establish that the overlap in the initial mora, but not the onset segment, between the word distractor and the color name produces facilitation. The distractors were all two-mora pseudowords written in kana.

Method

Participants Twenty-two students from Waseda University (12 males; average age 23 ± 2 years) in Tokyo participated in the experiment and were paid 500 yen each.

Design The experiment used the Stroop color naming task, and manipulated (1) onset congruence between the distractor and color name (congruent vs. incongruent) and (2) unit of onset (mora = CV vs. consonant phoneme = C). The dependent variables were color naming latency and error rate.

Materials The critical stimuli were 192 two-katakana pseudowords (e.g., パヤ , roman transcription “paya,” IPA: /pa.ja/) presented in pink, green or blue. Participants were instructed to use the color names “pinku,” “guriin,” and “buruu” (IPA: /pi.ŋ.ku/, /gu.ri.i.n/, /bu.ru.u/), which are commonly used loanword color names. We chose these color names because they have different consonant onsets, which is necessary to test for the onset segment effect, and the primary color names like red and blue in Japanese (red= アカ /a.ka/, blue= アヲ /a.o/) do not contain consonant onsets. There were four conditions resulting from the factorial combination of onset congruence and onset unit: (1) C-congruent; (2) C-incongruent; (3) CV-congruent; and (4) CV-incongruent. Forty-eight pseudowords contained the onset phoneme /p/, /b/ or /g/ (but not the initial mora /pi/, /gu/, or /bu/; e.g., パヤ , /pa.ja/ presented in pink) presented in the onset-congruent color; these constituted the congruent C stimuli. The incongruent C stimuli were 48 nonwords generated from the congruent C stimuli by replacing the onset phoneme /p/ with /t/ (e.g., from パヤ , /pa.ja/, タヤ , /ta.ja/ was generated), /b/ with /n/, and /g/ with /ci/. The congruent CV stimuli contained the initial mora /pi/, /gu/, or /bu/, presented in the onset-congruent color (e.g., ピヤ , /pi.ja/ presented in pink). The incongruent CV stimuli were 48 nonwords generated from the congruent CV stimuli by replacing the initial mora /pi/ with /ri/ (e.g., from ピヤ , /pi.ja/, リヤ , /ri.ja/ was generated), /bu/ with /nu/, and /gu/ with /tsu/. The stimuli are listed in the Appendix Table 4.

Apparatus and procedure Participants were tested individually, seated approximately 60 cm in front of a flat-screen monitor. Each participant completed 192 color naming trials, presented in four blocks, with each block containing 48 trials with a self-paced break between the blocks. Each block contained an equal number of trials from the four experimental conditions presented in either pink, green, or blue in equal proportion. A practice

block of nine trials containing each of the three colors occurring equally often preceded the test blocks.

Participants were instructed at the outset of the experiment that on each trial they would be presented with a katakana pseudoword presented in one of three colors—pink, green or blue—and their task was to name the color of the stimulus as fast and accurately as possible.

Stimulus presentation and data collection were achieved using the DMDX program (Forster & Forster, 2003). Stimulus display was synchronized to the screen refresh rate (16.7 ms).

Each trial started with the presentation of a fixation (+) for 816 ms, in the center of the screen. It was replaced by a blank screen for 400 ms, then by a word or five hash marks in one of three colors (pink, green, or blue) for a maximum of 2,000 ms, or until the participant named the color. After the participant’s response, the screen went blank for 816 ms, after which the next trial started. All stimuli were presented in MS P Gothic font. The experimenter sat next to the participant and noted the errors.

Results

In this and subsequent experiments, we report an analysis of correct RT and error rates using linear mixed-effects model, with subjects and items as crossed random factors (Baayen, 2008). The preliminary treatment of RT data for this analysis was as follows. First, we examined the shape of the RT distribution for correct trials and applied a log transformation (which best approximated a normal distribution) after excluding 84 data points faster than 200 ms (most of which were mistriggers of the voice key) to meet the distributional assumption of the linear mixed-effects model, resulting in 4,103 observations.

The logRT data were submitted to a linear mixed-effects model using the lme4 (Bates, Maechler, Bolker & Walker, 2015) package implemented in R 3.4.0 (R Core Team, 2016). Degrees of freedom (estimated using Satterthwaite’s approximation) and *p* values were estimated using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2016). We tested models that included random slopes, and where the model fit was no better, we report the simpler model with the subject and item intercepts (as crossed random factors). The mean correct RT and error rates are shown in Table 1.

In the analysis, the fixed factors were onset unit (C vs. CV) and congruence (congruent vs. incongruent). Both were dummy coded, and the congruence factor was referenced to the mismatch condition. For the onset unit factor, the C condition was used as the reference level to test the segment overlap effect, and the CV condition was used as the reference level to test the mora overlap effect. Using R syntax, the final model we report is: $\log RT \sim \text{onset unit} * \text{congruence} + (1 | \text{subject}) + (1 | \text{item})$, with 22 subjects and 192 items (distractors). With the C condition as the reference level, the model showed that the effect of congruence was nonsignificant: $t = -1.089$, $p = .277$ (i.e., there was no segment overlap effect). Referenced to the

Table 1 Mean color naming latencies (RT, in ms) and percentage error rates (%E) in Experiment 1 (katakana distractors)

Onset congruence with color name		Congruent		Incongruent		Difference	
Onset unit	Example of a congruent distractor (IPA) for /pi.ŋ.ku/ (pink)	RT	%E	RT	%E	RT	%E
C (phoneme)	パジャ (/pa.ja/)	504	0.9	511	0.7	7	−0.2
CV (mora)	ピャ (/pi.ja/)	483	1.1	520	0.8	37	−0.3

Note. Example is for “pink,” though the means reflect conditions for all three target colors

CV condition, the congruence effect was significant, $t = -5.658$, $p < .001$ (i.e., there was a mora overlap effect). The effect of onset unit was nonsignificant, $t = 1.408$, $p = .16$. The interaction between congruence and onset unit was significant, $t = -3.231$, $p < .002$, indicating that the effect of congruence in onset between the color name and the kana distractor depended on the onset unit (mora vs. segment). Error rate was not analyzed as they were too few errors to warrant an analysis.

Discussion

The results indicate clearly that for katakana stimuli (a moraic script) there is no benefit when only the onset segment of the distractor overlapped with the intended color name (e.g., /p/ between パジャ /pa.ja/ and the color name /piŋku/); in contrast, when the CV (i.e., the mora) overlapped (e.g., /pi/ between ピャ /pi.ja/ and /piŋku/), color naming was facilitated. These results replicate the previous findings using the implicit priming paradigm (Kureta et al., 2006) and the masked priming read-aloud task (Verdonschot et al., 2011). The absence of the onset segment effect here is all the more striking given that the participants were instructed to use loanword color names (/pi.ŋ.ku/, /gu.ri.i.n/, /bu.ru.u/, corresponding to pink, blue, green, respectively) that resemble the original English color names and hence may be considered even more conducive for finding an onset segment effect. In contrast to the nonsignificant onset segment effect, the initial mora effect of 37 ms is sizable, and comparable to the 40 ms form preparation effect reported by Kureta et al. (2006, Experiment 1).⁵ These results establish the phonological Stroop effect as a suitable vehicle to study the phonological unit in Japanese speech production.

Experiment 2: Stroop task (CV overlap) using kanji/hiragana distractors (separate blocks)

In the following two experiments, we turn to the main question of interest, namely, whether the phonological unit of speech production is dependent on script type. To this end, we use the Stroop color naming task and compare the effect of initial mora overlap with the target color name for kana versus multimoraic

kanji distractors. We first selected single-kanji words that have a multimoraic pronunciation, for example, 菊 (/ki.ku/), 兄 (/a.ni/), 娘 (/mu.su.me/); they were then presented either as a kanji character or as kana transcriptions (e.g., きく – /ki.ku/, あに – /a.ni/, むすめ – /mu.su.me/), in a color that shared the initial mora (e.g., /ki.ku/ in yellow /ki.i.ro/), or a color that did not (e.g., /ki.ku/ in purple, i.e., /mu.ra.sa.ki/).

Replicating Experiment 1, we expect to find the initial mora congruence effect for the kana-transcribed distractors. Note that native Japanese color names are typically written using kanji (i.e., not in katakana as in Experiment 1). Therefore, replication of the mora congruence effect here would serve to highlight that the effect is not dependent on the to-be-named colors usually being written in a moraic script. The critical prediction concerned the kanji distractors. Recall that Yoshihara et al. (2017) found no initial mora priming effect for two-character kanji words unless the whole pronunciation of the initial kanji character was the same in the prime and the target; kana transcription of the prime and target words restored the initial mora priming effect. If these results are due to the script-dependence of phonological unit of speech production, then similarly no initial mora congruence effect should be found in the present Stroop color naming task for the kanji distractors. On the other hand, if script type constrains speech production only when the target is a written word as in Yoshihara et al.’s read-aloud task, the initial mora congruence effect should be found irrespective of script type in the present Stroop color naming task.

In Experiment 2, we manipulated script type within participants, and presented the kanji distractors and kana distractors in separate blocks, with the kanji block always presented first. This was done to adhere to the design of Yoshihara et al. (2017) in which the script type was manipulated between experiments, and hence the kanji stimuli would not have been subject to carry over effects from the kana stimuli. In Experiment 3 we randomly mixed the kanji and kana distractors.

Method

Participants Thirty students from Waseda University in Tokyo (17 males; average age 22 ± 5 years) participated in the experiment; each participant was paid 500 yen.

Design The experiment used the Stroop color naming task and involved the factors onset congruence (congruent vs.

⁵ In contrast, the initial mora priming effect observed with the kana stimuli in Yoshihara et al.’s (2017) masked priming read-aloud experiment was 16 ms.

Table 2 Mean color naming latencies (RT, in ms) and percentage error rates (%E) in Experiment 2 (kanji and kana distractors presented in separate blocks, with the kanji block first)

Onset congruence with color name		Congruent		Incongruent		Difference	
Script type	Example of a congruent distractor (IPA) for /ki.i.ro/ (yellow)	RT	%E	RT	%E	RT	%E
Kanji	霧 (/ki.ri/)	589	1.0	623	1.2	34	0.2
Kana	きり (/ki.ri/)	582	0.4	630	0.9	48	0.5

Note. Example is for “yellow,” though the means are for all five target colors

incongruent) and distractor script type (kanji vs. kana). The dependent variables were color naming latency and error rate.

Materials The critical stimulus words were 45 single-character kanji words, selected using an online kanji database (Tamaoka et al., 2017). Most Japanese kanji characters have multiple readings, *on-* (“Chinese”) and *kun-* (“Japanese”) readings, and when a character is presented singly, the *kun-*reading predominates (see Verdonschot et al., 2013). The critical words contained, in their *kun-*readings, the initial mora /ki/, /a/, /mi/, /mu/, or /ci/; the initial mora of the color names used, /ki.i.ro/ (yellow), /a.ka/ (red), /mi.do.ri/ (green), /mu.ra.sa.ki/ (purple), and /ci.ro/ (white). The kanji words ranged in the number of strokes from three to 19 (mean 9.22), and frequency 2.49 to 1,826 per million (mean 194). Three of the words were three mora long; the remaining 42 words were two mora long.

The distractors were either presented as a kanji character, or transcribed and presented in hiragana. In both the kanji conditions and kana conditions, the distractor was presented either in a color that had the congruent initial mora (e.g., the word /ki.ku/, written either in kanji or kana, was presented in yellow (/ki.i.ro/) or in a color that had an incongruent initial mora (e.g., the word /ki.ku/ was presented in purple; /mu.ra.sa.ki/). Thus, in total there were 180 test trials (45 stimulus words presented in four different script type x onset congruence combinations). The stimuli are listed in the Appendix Table 5.

Apparatus and procedure Apparatus and the general procedure were identical to Experiment 1. Before the test trials, participants were presented with 10 practice trials in which a string of hash marks were presented in four different colors to familiarize with the color names. The presentation of kanji and kana distractors were blocked, and the kanji block was presented first. Five practice trials containing kanji- or kana distractors (not used in the test trials) preceded each of the kanji or kana blocks.

Results

The analysis procedure was identical to Experiment 1. The preliminary treatment of RT data for this experiment resulted in 5,310 data points, after excluding 42 data points (out of 5,352 trials) faster than 250 ms (most of which were

mistriggers of the voice key). The mean correct RT and error rates for Experiment 2 are shown in Table 2.

In the analysis, the fixed factors were script type (kanji vs. kana) and initial mora congruence (congruent vs. incongruent). Both were dummy coded, and the congruence factor was referenced to the incongruent condition. For the script type factor, the kanji condition and the kana condition was each used as the reference level. In this experiment, the model with intersubject and interitem variation in the sensitivity to congruence (i.e., subject and item random slopes on congruence) was preferred. Using R syntax, the model we report is: $\log RT \sim \text{script type} * \text{congruence} + (\text{congruence} | \text{subject}) + (\text{congruence} | \text{item})$, with 30 subjects and 90 items (distractors).

With kana as the reference level, the model showed that the effect of congruence was significant: $t = -4.889$, $p < .001$. Also with kanji as the reference level, the effect of congruence was significant, $t = -3.134$, $p < .003$. The effect of script type was non-significant, $t = .660$, $p = .511$.⁶ The interaction between onset congruence and script type was nonsignificant, $t = 1.291$, $p = .20$. Numerically the onset congruence effect was greater for the kana distractors (48 ms vs. 34 ms), and in the model with only the subject and item random intercepts ($\log RT \sim \text{script type} * \text{congruence} + (1 | \text{subject}) + (1 | \text{item})$)⁷, the interaction was statistically significant, $t = 3.205$, $p < .002$. This suggests that there was inter-subject and inter-item variability in the sensitivity to the manipulation of onset congruence, and the interaction largely reflects this variability. Error rate was not analyzed as they were too few errors to warrant analysis.

⁶ Given that kanji words are read aloud more slowly than kana transcriptions (Tamaoka, 2005), the absence of script type effect here may at first appear surprising. Note however that the task here is to name the color, not to read aloud the kanji and kana distractors. In the Stroop color naming task, it is well-established that the lexical factors that have large effects on the latency to read aloud the distractor (e.g., word frequency, lexicality) have little effects on the Stroop color naming latency (e.g., Kinoshita, De Wit, & Norris, 2017; Monsell, Taylor & Murphy, 2001). Thus, the absence of kanji–kana difference here is not unexpected.

⁷ The comparison of the two models showed that the AIC (Akaike information criterion) for the random-intercept only model was -3671.2 and for the random-slope model -3965.6 ; the BIC (Bayesian information criterion) was -3625.2 and -3893.3 ; log likelihood was 1842.6 and 1993.8, and the deviance was -3685.2 and -3987.6 , respectively. The chi square value was 302.36, $p(\text{chi square}) < .001$.

Table 3 Mean color naming latencies (RT, in ms) and percentage error rates (%E) in Experiment 3 (kanji and kana distractors mixed randomly)

Onset congruence with color name (5 colors used)		Congruent		Incongruent		Difference	
Script type	Example of congruent distractor (IPA) for /ki.i.ro/ (yellow)	RT	%E	RT	%E	RT	%E
Kanji	霧 (/ki.ri/)	526	1.0	565	1.8	39	0.8
Kana	きり (/ki.ri/)	522	0.8	563	1.9	41	1.1

Note. Example is for “yellow,” though the means reflect conditions for all five target colors

Discussion

Experiment 2 showed that the kanji distractors and their kana transcriptions produced statistically equivalent Stroop facilitation in color naming resulting from the overlap in the initial mora with the target color. The results for the kana distractors replicated Experiment 1 and previous findings of mora overlap effects found with the implicit priming procedure (Kureta et al., 2006) and the masked onset priming read-aloud experiments (Verdonschot et al., 2011). In contrast, the results for the kanji distractors stand in contrast to the results reported by Yoshihara et al. (2017), particularly with respect to the fact that their multimoraic kanji stimuli did not show any mora priming in the masked priming read-aloud task. Before turning to a discussion of this dissociation, we note that here numerically the initial mora congruence effect was smaller for the kanji stimuli (34 ms) than the kana stimuli (48 ms), and the interaction with script type was statistically significant when the item and subject variability in the sensitivity to the congruence was disregarded. Because the kana and kanji distractors were presented in separate blocks, and the script type manipulation was completely confounded with block order, it is unclear how to interpret the apparent interaction. In Experiment 3, therefore, we randomly mixed the kana and kanji distractors.

Experiment 3: Stroop task (CV overlap) using kanji/hiragana distractors (mixed blocks)

This experiment is essentially the same as Experiment 2, except the kanji and kana distractors are mixed randomly.

Method

Participants Twenty students from Waseda University in Tokyo (seven males; average age 19 ± 1 years) participated in the experiment.

Design The experimental design was identical to Experiment 2; the only difference is that the kanji and kana distractors were mixed randomly, in an unpredictable order.

Materials, apparatus, and procedure They were identical to Experiment 2, except that the presentation of kanji and kana distractors was mixed, and following practice with color naming with the string of hash marks, 10 practice trials containing kanji or kana distractors preceded the test trials.

Results

The same analysis procedure as Experiment 2 was followed. The preliminary treatment of RT data for this experiment resulted in 3487 data points, after excluding 64 data points (out of 3551 trials) faster than 250 ms (mostly voice key mistriggers). The mean correct RT and error rates for Experiment 3 are shown in Table 3.

The analysis was identical to Experiment 2, with script type (kanji vs. kana) and initial mora congruence (congruent vs. incongruent) as fixed factors. Both were dummy coded, and the congruence factor was referenced to the incongruent condition. For the script type factor, the kanji condition and the kana condition was each used as the reference level. As in Experiment 2, the model with intersubject and interitem variation in the sensitivity to onset congruence (i.e., subject and item random slopes on onset congruence) was preferred. Using R syntax, the model we report is: $\log RT \sim \text{script type} * \text{congruence} + (\text{congruence} | \text{subject}) + (\text{congruence} | \text{item})$, with 20 subjects and 90 items (distractors).

With kana as the reference level, the model showed that the effect of congruence was significant: $t = -4.281$, $p < .001$. Also with kanji as the reference level, the congruence effect was significant, $t = -4.157$, $p < .001$. The effect of script type was nonsignificant, $t = -0.283$, $p = .778$. Importantly, and replicating Experiment 2, the interaction between congruence and script type was also nonsignificant, $t = -0.091$, $p = .928$, indicating that initial mora congruence facilitated color naming equally for kanji or kana distractors. Error rate was not analyzed as they were too few errors to warrant analysis.

Discussion

We replicated the findings of Experiment 2 indicating that color naming in the Stroop task benefits from the overlap

in the initial mora between the color name and the word distractor, whether the word distractors were written in kana or kanji. Here, the kanji and kana distractors were mixed randomly, and the numerical trend toward a smaller initial mora overlap effect for the kanji stimuli observed in Experiment 2 disappeared. As in Experiment 2, the initial mora overlap effect was sizable (~40 ms) and very similar in size to the effect observed in Experiment 1 (and Kureta et al., 2006). These results strengthen our conclusion that the phonological unit of Japanese speech production, as indexed by the phonological Stroop effect, is independent of script type.

General discussion

The phonological unit underlying speech production has received ample attention in recent literature (e.g., Chen et al., 2016; O’Seaghdha, 2015; O’Seaghdha et al., 2010; Roelofs, 2015; Verdonschot et al., 2011; Verdonschot et al., 2013; Nakayama et al., 2016). It has been suggested that this unit is different between languages: Specifically, it has been proposed to be the phoneme in English and most other Indo-European languages, but the syllable in Mandarin Chinese and the mora in Japanese.

In the present study, we asked whether the cross-language variation in the phonological unit of speech production reflects the difference in script type used to write the language. This factor cannot be disentangled from the phonological unit of speech production in languages that are written using only one type of script. Here, we capitalized on the multiscripted nature of Japanese in which a word may be written in moraic kana, or logographic (morphemic) kanji. Using the Stroop color naming task, we showed that the overlap in the initial mora between the color name and the word distractor (e.g., the word /ki.ku/ presented in yellow /ki.i.ro/) produced a sizable facilitation, and crucially, the facilitation did not depend on script type: The multimoraic kanji distractors (e.g., 菊) produced an equally sized initial mora congruency effect as the kana distractors (e.g., きく).

The present results stand in marked contrast to a masked priming study reported by Yoshihara et al. (2017), in which the mora priming effects were absent for kanji compound words containing a multimoraic initial character. This difference was anticipated on the basis that in Yoshihara et al., the target was a written (two-character compound kanji) word to be read aloud, whereas here it was a (multimoraic) color name. In reading

aloud a written word, unlike a color name, the phonological form can be generated in a piecemeal fashion, via the process of mapping script to sound. The masked priming effect observed by Yoshihara et al. may have reflected this script-to-sound mapping process rather than (or as well as) the phonological encoding process.⁸ Chen et al. (2016), who used masked priming to study the proximate unit used in Mandarin Chinese speech production, made a similar point. They used pictures as targets (preceded by Mandarin *hànzì* primes) thereby avoiding “perceptual priming of characters in the writing system” (p. 827). One point of concern with the present Stroop task is that it allows for a limited number of potential response options because the number of available colors is limited. The fact that Chen et al. (2016) used picture targets (they used 42 pictures) and showed a similar absence of onset segment overlap effect (with Mandarin Chinese speakers) is therefore reassuring.

It should also be pointed out that the dissociation between the present study and Yoshihara et al. (2017) is at odds with accounts based on the dual-route cascaded (DRC) model of reading (e.g., Coltheart et al., 1999; Mousikou et al., 2015), according to which the phonological Stroop effect and the masked onset priming effect have the same locus (Mousikou et al., 2015, p. 1078). The DRC model assumes that both effects originate in the reading of the prime/distractor. More specifically, it attributes the serial, left-to-right nature of both effects (i.e., beginning overlap but not end overlap produces facilitation) to the serial operation of the nonlexical route of print-to-sound mapping: This route is assumed to apply grapheme–phoneme correspondence rules, letter by letter, to generate a sequence of phonemes that serves as the input to the speech planning process. Although the DRC model has yet not been applied to Japanese, for kana stimuli, a similar serial application of kana-to-mora correspondence rules could plausibly produce an initial mora priming effect in reading aloud (as observed by Yoshihara et al., 2017) and an initial mora overlap effect in the Stroop color naming task observed here. In contrast, on

⁸ In addition, the two studies differed in that Yoshihara et al. (2017) used two-character compound kanji words (e.g., 博物 /ha.ku.bu.tsu/), whereas we used single-character kanji words (e.g., 娘 /mu.su.me/). As noted in the Method section of Experiment 2, many kanji characters have multiple pronunciations (cf. Verdonschot et al., 2013), and single-character kanji words usually take on the “kun” (“Japanese”) reading whereas compound kanji words usually take on the “on” (Chinese) reading. It is possible that a mora priming effect may emerge for kanji characters with different whole-character pronunciation (e.g., 無 /mu/–娘 /mu.su.me/) when single-character kanji words are read aloud. This possibility remains to be investigated.

the assumption that the pronunciation of kanji words is retrieved “lexically” as a whole (e.g., Feldman & Turvey, 1980; Wydell, Butterworth, & Patterson, 1995), no initial mora overlap effect is expected for these stimuli, contrary to the present results. Moreover, assuming a single, common locus of the masked onset priming effect and the phonological Stroop effect, the dissociation between Yoshihara et al. (2017) and the present Stroop color naming task is unexpected by the DRC.

DRC’s assumption of serial operation of the nonlexical route has explained many (serial) phenomena observed with English (and other European languages) in the read-aloud task, such as the position-of-irregularity effect (see Rastle & Coltheart, 2006, for a review). The DRC model offers detailed, computationally implemented process accounts of these phenomena (see Mousikou et al., 2015, for the masked onset priming effect and the phonological Stroop effect). However, these phenomena have been found using the read-aloud task, which involves speech production, and it has been suggested that the serial nature of reading aloud could be subsumed under the left-to-right incremental nature of phonological encoding process in speech production (Kinoshita, 2000; Roelofs, 2004). To date, the DRC model has not incorporated this suggestion, and the serial nature of these phenomena is entirely situated in the grapheme-phoneme mapping process, which is specific to alphabetic scripts and is assumed to be completed before speech planning commences. This approach is difficult to reconcile with the fact that the facilitation in naming due to the overlap in the initial phonological unit is found at the level of mora in Japanese and at the level of segments in European languages, as well as the fact that the initial mora congruence effect is independent of script type within Japanese.

The alphabet/phoneme-centric approach contrasts with the “proximate unit principle.” The term proximate unit was proposed by O’Seaghdha and colleagues (e.g., O’Seaghdha, 2015; O’Seaghdha & Chen, 2009) to refer to the unit used in the phonological encoding of speech production, and particularly the cross-language variation in the unit. O’Seaghdha (2015) criticized the extant theories of language production which regard phonemes as the essential building blocks of lexical sound systems, noting that “because theories of language production have been shaped by the properties of the languages in which they emerged, the properties of those languages are over-represented in theories” (p. 5). In contrast, in the proximate unit approach full recognition is given to the diversity of form—the cross-language variation in phonology—and the proximate units (i.e., the primary units mediating the initial transition from lexical to phonological

representation) are suggested to vary across languages (phonemes/segments in European languages, atonal syllables in Chinese, mora in Japanese).

In the proximate unit approach, the dissociation between Yoshihara et al.’s (2017) finding and the present study is less puzzling. In the review of the literature, O’Seaghdha (2015) discussed how the implicit priming paradigm and the masked onset priming task have provided the evidence for the proximate unit, but in different contexts of speech production. In principle, the proximate unit addresses the retrieval of phonology from lexical memory, as in normal speaking, and not the mapping of phonology from orthographic input (i.e., reading aloud). This is not to deny the usefulness of masked priming to investigate speech production, but as O’Seaghdha (2015) noted, “masked primes could influence many word production processes, and so caution must be exercised in applying the proximate units’ approach to masked priming data” (p. 15). As pointed out by Roelofs (2006) in explaining the orthographic influence on the implicit priming paradigm, orthography (and script type) is expected to constrain spoken word production when it is relevant to the word production task at hand. In this regard, an important difference between the masked priming read-aloud task and the Stroop color naming task is that in the former the target is a written word to be read aloud, and in the latter the target is a color to be named. The phonological form of a written word can be generated in a piecemeal fashion, script-by-script, which is not possible for a color or a picture, thus as noted by Chen et al. (2016), priming in a read-aloud task may reflect “perceptual priming of characters in the writing system rather than by production-specific processes” (p. 827).

In conclusion, the present Stroop color naming study showed that color naming in Japanese benefitted from an overlap with the distractor in the initial mora, but not in the onset segment, and that the initial mora effect was independent of script type (moraic kana vs. multimoraic kanji). This script independence contrasts with the result reported by Yoshihara et al. (2017) in a recent masked priming read-aloud study. These results, taken together with the previous demonstrations of the onset segment congruence effect in English (Coltheart et al., 1999; Mousikou et al., 2015) establish the Stroop color naming task as a useful task for investigating the proximate unit in speech production.

Acknowledgements This research is supported by a JSPS Grant-in-Aid for Scientific Research (C) (17K02748) as well as a Waseda University Grant for Promotion of International Joint Research to Rinus G. Verdonschot. We are grateful to Akira Yoshida for his assistance in carrying out the experiments.

Appendix

Table 4 Critical stimuli used in Experiment 1 (Note: /pin_ku/ = pink; /gu_{ri}:n/ = green; /bu_{ru}:/ = blue)

C congruent			C incongruent		
IPA transcription	Distractor	IPA color name	IPA transcription	Distractor	IPA color name
/pa.ja/	パヤ	/pi.ŋ.ku/	/ra.ja/	ラヤ	/pi.ŋ.ku/
/pa.u/	パウ	/pi.ŋ.ku/	/ra.u/	ラウ	/pi.ŋ.ku/
/pa.ha/	パハ	/pi.ŋ.ku/	/ra.ha/	ラハ	/pi.ŋ.ku/
/pa.ta/	パタ	/pi.ŋ.ku/	/ra.ta/	ラタ	/pi.ŋ.ku/
/pa.mi/	パミ	/pi.ŋ.ku/	/ra.mi/	ラミ	/pi.ŋ.ku/
/pa.ni/	パニ	/pi.ŋ.ku/	/ra.ni/	ラニ	/pi.ŋ.ku/
/pa.çi/	パシ	/pi.ŋ.ku/	/ra.çi/	ラシ	/pi.ŋ.ku/
/pa.ki/	パキ	/pi.ŋ.ku/	/ra.ki/	ラキ	/pi.ŋ.ku/
/pa.ju/	パユ	/pi.ŋ.ku/	/ra.ju/	ラユ	/pi.ŋ.ku/
/pa.ru/	パル	/pi.ŋ.ku/	/ra.ru/	ラル	/pi.ŋ.ku/
/pa.he/	パヘ	/pi.ŋ.ku/	/ra.he/	ラヘ	/pi.ŋ.ku/
/pa.re/	パレ	/pi.ŋ.ku/	/ra.re/	ラレ	/pi.ŋ.ku/
/pa.me/	パメ	/pi.ŋ.ku/	/ra.me/	ラメ	/pi.ŋ.ku/
/pa.ne/	パネ	/pi.ŋ.ku/	/ra.ne/	ラネ	/pi.ŋ.ku/
/pa.ro/	パロ	/pi.ŋ.ku/	/ra.ro/	ラロ	/pi.ŋ.ku/
/pa.to/	パト	/pi.ŋ.ku/	/ra.to/	ラト	/pi.ŋ.ku/
/bo.ɯa/	ボワ	/bu.ru.u/	/no.ɯa/	ノワ	/bu.ru.u/
/bo.ja/	ボヤ	/bu.ru.u/	/no.ja/	ノヤ	/bu.ru.u/
/bo.ha/	ボハ	/bu.ru.u/	/no.ha/	ノハ	/bu.ru.u/
/bo.na/	ボナ	/bu.ru.u/	/no.na/	ノナ	/bu.ru.u/
/bo.sa/	ボサ	/bu.ru.u/	/no.sa/	ノサ	/bu.ru.u/
/bo.ka/	ボカ	/bu.ru.u/	/no.ka/	ノカ	/bu.ru.u/
/bo.çi/	ボヒ	/bu.ru.u/	/no.çi/	ノヒ	/bu.ru.u/
/bo.ni/	ボニ	/bu.ru.u/	/no.ni/	ノニ	/bu.ru.u/
/bo.tçi/	ボチ	/bu.ru.u/	/no.tçi/	ノチ	/bu.ru.u/
/bo.ju/	ボユ	/bu.ru.u/	/no.ju/	ノユ	/bu.ru.u/
/bo.ɸu/	ボフ	/bu.ru.u/	/no.ɸu/	ノフ	/bu.ru.u/
/bo.re/	ボレ	/bu.ru.u/	/no.re/	ノレ	/bu.ru.u/
/bo.me/	ボメ	/bu.ru.u/	/no.me/	ノメ	/bu.ru.u/
/bo.he/	ボヘ	/bu.ru.u/	/no.he/	ノヘ	/bu.ru.u/
/bo.a/	ボア	/bu.ru.u/	/no.a/	ノア	/bu.ru.u/
/bo.ho/	ボホ	/bu.ru.u/	/no.ho/	ノホ	/bu.ru.u/
/gi.ja/	ギヤ	/gu.ri.i.n/	/çi.ja/	シヤ	/gu.ri.i.n/
/gi.ma/	ギマ	/gu.ri.i.n/	/çi.u/	シウ	/gu.ri.i.n/
/gi.ha/	ギハ	/gu.ri.i.n/	/çi.ha/	シハ	/gu.ri.i.n/
/gi.na/	ギナ	/gu.ri.i.n/	/çi.na/	シナ	/gu.ri.i.n/
/gi.sa/	ギサ	/gu.ri.i.n/	/çi.sa/	シサ	/gu.ri.i.n/
/gi.çi/	ギヒ	/gu.ri.i.n/	/çi.çi/	シヒ	/gu.ri.i.n/
/gi.ju/	ギユ	/gu.ri.i.n/	/çi.ju/	シユ	/gu.ri.i.n/
/gi.nu/	ギヌ	/gu.ri.i.n/	/çi.su/	シス	/gu.ri.i.n/
/gi.tsu/	ギツ	/gu.ri.i.n/	/çi.tsu/	シツ	/gu.ri.i.n/
/gi.ku/	ギク	/gu.ri.i.n/	/çi.ku/	シク	/gu.ri.i.n/
/gi.re/	ギレ	/gu.ri.i.n/	/çi.re/	シレ	/gu.ri.i.n/
/gi.me/	ギメ	/gu.ri.i.n/	/çi.me/	シメ	/gu.ri.i.n/
/gi.he/	ギヘ	/gu.ri.i.n/	/çi.he/	シヘ	/gu.ri.i.n/
/gi.jo/	ギヨ	/gu.ri.i.n/	/çi.jo/	シヨ	/gu.ri.i.n/

Table 4 (continued)

/gi.mo/ /gi.to/	ギモ ギト	/gu.ri.i.n/ /gu.ri.i.n/	/ci.mo/ /ci.to/	シモ シト	/gu.ri.i.n/ /gu.ri.i.n/
CV congruent			CV incongruent		
IPA transcription	Distractor	IPA color name	IPA transcription	Distractor	IPA color name
/pi.ja/	ピヤ	/pi.ŋ.ku/	/ri.ja/	リヤ	/pi.ŋ.ku/
/pi.u/	ピウ	/pi.ŋ.ku/	/ri.u/	リウ	/pi.ŋ.ku/
/pi.ha/	ピハ	/pi.ŋ.ku/	/ri.ha/	リハ	/pi.ŋ.ku/
/pi.ta/	ピタ	/pi.ŋ.ku/	/ri.ta/	リタ	/pi.ŋ.ku/
/pi.mi/	ピミ	/pi.ŋ.ku/	/ri.mi/	リミ	/pi.ŋ.ku/
/pi.ni/	ピニ	/pi.ŋ.ku/	/ri.ni/	リニ	/pi.ŋ.ku/
/pi.ci/	ピシ	/pi.ŋ.ku/	/ri.ci/	リシ	/pi.ŋ.ku/
/pi.ki/	ピキ	/pi.ŋ.ku/	/ri.ki/	リキ	/pi.ŋ.ku/
/pi.ju/	ピユ	/pi.ŋ.ku/	/ri.ju/	リユ	/pi.ŋ.ku/
/pi.ru/	ピル	/pi.ŋ.ku/	/ri.ru/	リル	/pi.ŋ.ku/
/pi.he/	ピヘ	/pi.ŋ.ku/	/ri.he/	リヘ	/pi.ŋ.ku/
/pi.re/	ピレ	/pi.ŋ.ku/	/ri.re/	リレ	/pi.ŋ.ku/
/pi.me/	ピメ	/pi.ŋ.ku/	/ri.me/	リメ	/pi.ŋ.ku/
/pi.ne/	ピネ	/pi.ŋ.ku/	/ri.ne/	リネ	/pi.ŋ.ku/
/pi.ro/	ピロ	/pi.ŋ.ku/	/ri.ro/	リロ	/pi.ŋ.ku/
/pi.to/	ピト	/pi.ŋ.ku/	/ri.to/	リト	/pi.ŋ.ku/
/bu.u.ɥa/	ブウ	/bu.ru.u/	/nu.u.ɥa/	ヌウ	/bu.ru.u/
/bu.ja/	ブヤ	/bu.ru.u/	/nu.ja/	ヌヤ	/bu.ru.u/
/bu.ha/	ブハ	/bu.ru.u/	/nu.ha/	ヌハ	/bu.ru.u/
/bu.ni/	ブニ	/bu.ru.u/	/nu.ni/	ヌニ	/bu.ru.u/
/bu.sa/	ブサ	/bu.ru.u/	/nu.sa/	ヌサ	/bu.ru.u/
/bu.ka/	ブカ	/bu.ru.u/	/nu.ka/	ヌカ	/bu.ru.u/
/bu.çi/	ブヒ	/bu.ru.u/	/nu.çi/	ヌヒ	/bu.ru.u/
/bu.ni/	ブニ	/bu.ru.u/	/nu.ni/	ヌニ	/bu.ru.u/
/bu.tçi/	ブチ	/bu.ru.u/	/nu.tçi/	ヌチ	/bu.ru.u/
/bu.ju/	ブユ	/bu.ru.u/	/nu.ju/	ヌユ	/bu.ru.u/
/bu.ɸu/	ブフ	/bu.ru.u/	/nu.ɸu/	ヌフ	/bu.ru.u/
/bu.re/	ブレ	/bu.ru.u/	/nu.re/	ヌレ	/bu.ru.u/
/bu.me/	ブレ	/bu.ru.u/	/nu.me/	ヌメ	/bu.ru.u/
/bu.he/	ブヘ	/bu.ru.u/	/nu.he/	ヌヘ	/bu.ru.u/
/bu.a/	ブア	/bu.ru.u/	/nu.a/	ヌア	/bu.ru.u/
/bu.ho/	ブホ	/bu.ru.u/	/nu.ho/	ヌホ	/bu.ru.u/
/gu.ja/	グヤ	/gu.ri.i.n/	/tsu.sa/	ツサ	/gu.ri.i.n/
/gu.ma/	グマ	/gu.ri.i.n/	/tsu.nu/	ツヌ	/gu.ri.i.n/
/gu.ha/	グハ	/gu.ri.i.n/	/tsu.ha/	ツハ	/gu.ri.i.n/
/gu.na/	グナ	/gu.ri.i.n/	/tsu.ɥa/	ツウ	/gu.ri.i.n/
/gu.sa/	グサ	/gu.ri.i.n/	/tsu.sa/	ツサ	/gu.ri.i.n/
/gu.çi/	グヒ	/gu.ri.i.n/	/tsu.çi/	ツヒ	/gu.ri.i.n/
/gu.ju/	グユ	/gu.ri.i.n/	/tsu.su/	ツス	/gu.ri.i.n/
/gu.nu/	グヌ	/gu.ri.i.n/	/tsu.nu/	ツヌ	/gu.ri.i.n/
/gu.tsu/	グツ	/gu.ri.i.n/	/tsu.tsu/	ツツ	/gu.ri.i.n/
/gu.ku/	グク	/gu.ri.i.n/	/tsu.ku/	ツク	/gu.ri.i.n/
/gu.re/	グレ	/gu.ri.i.n/	/tsu.re/	ツレ	/gu.ri.i.n/
/gu.me/	グメ	/gu.ri.i.n/	/tsu.se/	ツセ	/gu.ri.i.n/
/gu.he/	グヘ	/gu.ri.i.n/	/tsu.he/	ツヘ	/gu.ri.i.n/
/gu.jo/	グヨ	/gu.ri.i.n/	/tsu.jo/	ツヨ	/gu.ri.i.n/
/gu.mo/	グモ	/gu.ri.i.n/	/tsu.mo/	ツモ	/gu.ri.i.n/
/gu.to/	グト	/gu.ri.i.n/	/tsu.to/	ツト	/gu.ri.i.n/

Table 5 Critical stimuli used in Experiments 2 and 3 (Note. /kiro/ = yellow, /aka/= red, /midori/ = green, /murasaki/= purple, /giro/ = white)

IPA transcription	Translation	Kanji	Hiragana	IPA congruent color name	IPA incongruent color name
/ki.ri/	mist	霧	きり	/ki.i.ro/	/muu.ra.sa.ki/
/ki.çi/	(river) bank	岸	きし	/ki.i.ro/	/muu.ra.sa.ki/
/ki.kuu/	chrysanthemum	菊	きく	/ki.i.ro/	/muu.ra.sa.ki/
/ki.ta/	north	北	きた	/ki.i.ro/	/muu.ra.sa.ki/
/ki.nuu/	silk	絹	きぬ	/ki.i.ro/	/muu.ra.sa.ki/
/ki.mo/	liver	肝	きも	/ki.i.ro/	/mi.do.ri/
/ki.N/	gold	金	きん	/ki.i.ro/	/mi.do.ri/
/ki.tsuu.ne/	fox	狐	きつね	/ki.i.ro/	/mi.do.ri/
/ki.dzuu/	injury	傷	きず	/ki.i.ro/	/mi.do.ri/
/a.çi/	foot	足	あし	/a.ka/	/ki.i.ro/
/a.sa/	morning	朝	あさ	/a.ka/	/ki.i.ro/
/a.sa/	flax (hemp)	麻	あさ	/a.ka/	/ki.i.ro/
/a.me/	rain	雨	あめ	/a.ka/	/ki.i.ro/
/a.na/	hole	穴	あな	/a.ka/	/ki.i.ro/
/a.ni/	elder brother	兄	あに	/a.ka/	/ki.i.ro/
/a.se/	perspiration	汗	あせ	/a.ka/	/ki.i.ro/
/a.to/	mark	跡	あと	/a.ka/	/ki.i.ro/
/a.ne/	elder sister	姉	あね	/a.ka/	/ki.i.ro/
/a.mi/	net	網	あに	/a.ka/	/mi.do.ri/
/a.uqa/	bubble	泡	あわ	/a.ka/	/mi.do.ri/
/a.ki/	autumn	秋	あき	/a.ka/	/mi.do.ri/
/a.ra.çi/	storm	嵐	あらし	/a.ka/	/mi.do.ri/
/a.dzi/	taste	味	あじ	/a.ka/	/mi.do.ri/
/mi.mi/	ear	耳	みみ	/mi.do.ri/	/a.ka/
/mi.se/	shop	店	みせ	/mi.do.ri/	/a.ka/
/mi.ki/	(tree) trunk	幹	みき	/mi.do.ri/	/a.ka/
/mi.tçi/	road	道	みち	/mi.do.ri/	/a.ka/
/mi.ne/	peak	峰	みね	/mi.do.ri/	/a.ka/
/mi.ja/	Shinto shrine	宮	みや	/mi.do.ri/	/çi.ro/
/mi.dzo/	ditch	溝	みぞ	/mi.do.ri/	/çi.ro/
/mi.dzuu/	water	水	みず	/mi.do.ri/	/çi.ro/
/mi.gi/	right	右	みぎ	/mi.do.ri/	/çi.ro/
/muu.ne/	chest	胸	むね	/muu.ra.sa.ki/	/çi.ro/
/muu.gi/	wheat	麦	むぎ	/muu.ra.sa.ki/	/çi.ro/
/muu.çi/	worm	虫	むし	/muu.ra.sa.ki/	/çi.ro/
/muu.ko/	son-in-law	婿	むこ	/muu.ra.sa.ki/	/çi.ro/
/muu.suu.me/	daughter	娘	むすめ	/muu.ra.sa.ki/	/a.ka/
/çi.o/	salt	塩	しお	/çi.ro/	/a.ka/
/çi.ka/	deer	鹿	しか	/çi.ro/	/a.ka/
/çi.ta/	tongue	舌	した	/çi.ro/	/a.ka/
/çi.mo/	frost	霜	しも	/çi.ro/	/a.ka/
/çi.ma/	island	島	しま	/çi.ro/	/a.ka/
/çi.ta/	down	下	した	/çi.ro/	/a.ka/
/çi.ri/	buttocks	尻	しり	/çi.ro/	/a.ka/
/çi.ba/	lawn	芝	しば	/çi.ro/	/a.ka/

References

- Alario, X. L., Perre, L., Castel, C., & Ziegler, J.C. (2007). The role of orthography in speech production revisited. *Cognition*, *102*, 464–45.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: <https://doi.org/10.18637/jss.v067.i01>.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177–208.
- Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, *91*, 77–111.
- Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language*, *46*, 751–781.
- Chen, J.-Y., Lin, W.-C., & Ferrand, L. (2003). Masked priming of the syllable in Mandarin Chinese speech production. *Chinese Journal of Psychology*, *45*, 107–120.
- Chen, J.-Y., O'Seaghdha, P. G., & Chen, T.-M. (2016). The primacy of abstract syllables in Chinese word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 825–836. doi:<https://doi.org/10.1037/a0039911>
- Coltheart, M., Woollams, A., Kinoshita, S., & Perry, C. (1999). A position-sensitive Stroop effect: Further evidence for a left-to-right component in print-to-speech conversion. *Psychonomic Bulletin & Review*, *6*, 456–463.
- Damian, M., & Bowers, J. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language*, *49*, 119–132.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283.
- Dimitropoulou, M., Dunabeitia, J. A., & Carreiras, M. (2010). Influence of prime lexicality, frequency and pronounceability on the masked onset priming effect. *Quarterly Journal of Experimental Psychology*, *63*, 1813–1837.
- 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*(2), 141–147.
- Forster, K. I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language*, *30*, 1–25.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers*, *35*, 116–124.
- Gómez, D. M., Berent, I., Benavides-Varela, S., Bion, R. A., Cattarossi, L., Nespor, M., & Mehler, J. (2014). Language universals at birth. *Proceedings of the National Academy of Sciences*, *111*(16), 5837–5841.
- Ito, T., & Tatsumi, I. (1997). Tokusyu-haku ni taisuru meta-gengo-chisiki no hattatsu [The development of metalinguistic awareness of mora-phonemes in Japanese young children]. *The Japan Journal of Logopedics and Phoniatrics*, *38*, 196–203.
- Kinoshita, S. (2000). The left-to-right nature of the masked onset priming effect in naming. *Psychonomic Bulletin & Review*, *7*, 133–141.
- Kinoshita, S., De Wit, B., & Norris, D. (2017). The magic of words reconsidered: Investigating the automaticity of reading color-neutral words in the Stroop task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *43*, 369–384.
- Kinoshita, S., & Woollams, A. (2002). The masked onset priming effect in naming: Computation of phonology or speech-planning? *Memory & Cognition*, *30*, 237–245.
- Kubozono, H. (1989). The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech*, *32*, 249–278.
- Kureta, Y., Fushimi, T., Sakuma, N., & Tatsumi, I. F. (2015). Orthographic influences on the word-onset phoneme preparation effect in native Japanese speakers: Evidence from the word-form preparation paradigm. *Japanese Psychological Research*, *57*, 50–60. doi:<https://doi.org/10.1111/jpr.12067>
- 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. doi:<https://doi.org/10.1037/0278-7393.32.5.1102>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in linear mixed effects models (R Package Version 2.0–30) [Computer software]. Retrieved from <http://cran.r-project.org/package=lmerTest>
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.
- Li, C., Wang, M., & Idsardi, W. (2015). The effect of orthographic form-cuing on the phonological preparation unit in spoken word production. *Memory & Cognition*, *43*, 563–578.
- Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*, 524–45.
- Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, *30*, 69–89.
- Monsell, S., Taylor, T. J., & Murphy, K. (2001). Naming the color of a word: Is it response or task sets that compete? *Memory & Cognition*, *29*, 137–151.
- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*, 323–331.
- Mousikou, P., Rastle, K., Besner, D., & Coltheart, M. (2015). The locus of serial processing in reading aloud: Orthography-to-phonology computation or speech planning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1076–1099. doi:<https://doi.org/10.1037/xlm0000090>
- Nakayama, M., Kinoshita, S., & Verdonschot, R. G. (2016). The emergence of a phoneme-sized unit of speech planning in Japanese-English bilinguals. *Frontiers in Psychology*, *7*, 175.
- O'Seaghdha, P. (2015). Across the great divide: Proximate units at the lexical-phonological interface. *Japanese Psychological Research*, *57*, 4–21.
- O'Seaghdha, P. G., & Chen, J. Y. (2009). Toward a language-general account of word production: The proximate units principle. In: CogSci... Annual Conference of the Cognitive Science Society. Cognitive Science Society (US). Conference (Vol. 2009, p. 68). NIH Public Access.
- O'Seaghdha, P. G., Chen, J.-Y., & Chen, T.-M. (2010). Proximate units in word production: Phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition*, *115*, 282–302.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, *32*, 258–278.
- R Core Team. (2016). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria: R Foundation for statistical computing. Retrieved from <http://www.R-project.org/>
- Rastle, K., & Coltheart, M. (2006). Is there serial processing in the reading system; and are there local representations? In S. Andrews (Ed.), *From inkmark to ideas: Current issues in lexical processing* (pp. 3–24). Hove: Psychology Press.

- Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic spelling. *Cognition*, 24, 31–44.
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110, 88–125.
- Roelofs, A. (2004). Seriality of phonological encoding in naming objects and reading their names. *Memory & Cognition*, 32, 212–222.
- Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review*, 13, 33–37.
- Roelofs, A. (2015). Modeling of phonological encoding in spoken word production: From Germanic languages to Mandarin Chinese and Japanese. *Japanese Psychological Research*, 57, 22–37.
- Schiller, N. O. (2004). The onset effect in word naming. *Journal of Memory & Language*, 50, 477–490.
- Stroop, J. R. (1935) Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662.
- Tamaoka, K. (2005). Is an orthographic unit of a single Japanese kanji equivalent to a single kanji phonological unit in a naming task? *Cognitive Studies*, 12, 47–73. (in Japanese)
- Tamaoka, K., Makioka, S., Sanders, S., & Verdonschot, R. G. (2017). A new interactive online database for psychological and linguistic research on Japanese kanji and their compound words. *Psychological Research*, 81, 696–708. Available at www.kanjidatabase.com
- Verdonschot, R. G., Kiyama, S., Tamaoka, K., Kinoshita, S., La Heij, W., & 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.
- Verdonschot, R. G., La Heij, W., Tamaoka, K., Kiyama, S., You, W.-P., & Schiller, N. O. (2013). The multiple pronunciations of Japanese kanji: A masked priming investigation. *Quarterly Journal of Experimental Psychology*, 66, 2023–2038.
- Verdonschot, R. G., Lai, J., Feng, C., Tamaoka, K., & Schiller, N. O. (2015). Constructing initial phonology in Mandarin Chinese: Syllabic or sub-syllabic? A masked priming investigation. *Japanese Psychological Research*, 57, 61–68.
- Wong, A. W.-K., & Chen, H.-C. (2008). Processing segmental and prosodic information in Cantonese word production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34, 1172–1190.
- Wong, A. W.-K., Huang, J., & Chen, H.-C. (2012). Phonological units in spoken word production: Insights from Cantonese. *PLOS ONE*, 7(11), e48776
- Wong, A. W.-K., & Chen, H.-C. (2009). What are effective phonological units in Cantonese spoken word planning? *Psychonomic Bulletin & Review*, 16, 888–892.
- 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(5), 1155–1168.
- Yoshihara, M., Nakayama, M., Verdonschot, R. G., & Hino, Y. (2017). The phonological unit of Japanese kanji compounds: A masked priming investigation. *Journal of Experimental Psychology: Human Perception and Performance*, 43(7), 1303–1328. doi: <https://doi.org/10.1037/xhp0000374>
- You, W.-P., Zhang, Q.-F., & Verdonschot, R. G. (2012). Masked syllable priming effects in word and picture naming in Chinese. *PLOS ONE*, 8, e46595.