Effects of orthography in the picture-word task: Evidence from Japanese scripts

Alexandra S. Dylman^{1,2} · Mariko Kikutani³ · Miho Sasaki⁴ · Christopher Barry⁵

Accepted: 30 May 2021 / Published online: 13 June 2021
© The Author(s) 2021

Abstract

The picture-word task presents participants with a number of pictured objects together with a written distractor word superimposed upon each picture, and their task is to name the depicted object while ignoring the distractor word. Depending on the specific picture and word combination, various effects, including the identity facilitation effect (e.g., DOG + dog) and the semantic interference effect (e.g., GOAT + cow), are often observed. The response patterns of the picture-word task in terms of naming latencies reflect the mechanisms underlying lexical selection in speech production. Research using this method, however, has typically focused on alphabetic languages, or involved bilingual populations, making it difficult to specifically investigate orthographic effects in isolation. In this paper, we report five experiments investigating the role of orthography in the picture-word task by varying distractor script (using the multiscriptal language Japanese, and pseudohomophonic spellings in English) across three different populations (Japanese monolinguals, Japanese-English bilinguals, and English monolinguals), investigating both the identity facilitation effect and the semantic interference effect. The results generally show that the magnitude of facilitation is affected by orthography even within a single language. The findings and specific patterns of results are discussed in relation to current theories on speech production.

Keywords Japanese · Orthography · Picture-word task · Speech production

✉ Alexandra S. Dylman
asdylman@gmail.com

¹ Department of Special Education, Stockholm University, 10691 Stockholm, Sweden

² Mälardalen University, Västerås, Sweden

³ Kanazawa University, Kanazawa, Japan

⁴ Keio University, Minato City, Japan

⁵ University of Essex, Colchester, UK

An essential part of speech production is lexical selection, whereby the lexical form of a to-be-produced word is retrieved. A communicative intent (or a given stimulus to name) will activate a number of relevant lexical representations out of which one will be chosen for production. What this selection process entails is a matter of debate. Some theories of speech production have proposed a competition account, suggesting that a set of activated concepts compete for selection, until, finally, the most appropriate lexical representation is selected (e.g., Levelt et al., 1999). Others have challenged this notion of competition, instead arguing that non-intended lexical representations are excluded at a post-lexical stage, without involving a competitive process per se (e.g., Mahon et al., 2007). Yet others have proposed that while speech production must involve some level of competition, it may not be at the level of lexical selection (Spalek et al., 2013).

The picture-word task¹ is a frequently used method of investigating various such questions regarding speech production. In the picture-word task, participants are presented with a number of pictured objects together with a written distractor word superimposed upon each picture. The task is to name the depicted object while ignoring the distractor word. The relationship between the target object and the distractor picture can be varied in numerous ways, and changes in naming latencies depending on this manipulation have been informative of the underlying mechanisms involved in spoken word production.

One classic phenomenon observed in this task is the *identity facilitation effect*; naming latencies are faster when the presented distractor word is the name of the target picture (e.g., DOG + dog²) compared to cases where the word does not specify the picture. Another recurrently found phenomenon in the picture-word task is the *semantic interference effect* whereby naming latencies are slower when the distractor word is semantically related to the target picture (e.g., DOG + cat, Roelofs, 1997; Schriefers et al., 1990; Starreveld & La Heij, 1996). Traditionally, both the facilitation and the interference effect have been explained in terms of activated lexical representations competing to be selected for production, and this notion is referred to as *lexical selection by competition* (e.g., Levelt et al., 1999; Roelofs, 2003). This hypothesis suggests that the time taken to select the lexical representation of the target word depends on the level of activation of the distractor word, in such a way that the target naming gets slower as the activation level of the distractor word gets higher, leading to more distraction from the competitor (Lupker, 1979). When the semantic distance between the target and distractor is very close, the lexical representation of the distractor can receive signals both from the written word and the picture via the semantic store. The lexical node of the distractor is, therefore, activated strongly, which interferes with the selection of the target word to be produced.

¹ While many refer to this task as the picture-word *interference* task, we prefer the more neutral term picture-word task, as the task has also been used in studies examining (and finding) *facilitation* (e.g., Costa, Miozzo & Caramazza, 1999; Dylman & Barry, 2018).

² Throughout this paper, target object names are denoted in upper case, and printed distractor words are underlined (e.g., DOG + dog). Lexical representations are denoted as ‘single quotes’ (e.g., ‘dog’), and spoken responses as “double quotes” (e.g., “dog”).

The identity facilitation effect can be explained by a similar mechanism. The target picture sends activation to the lexical representation of the to-be-produced word, and the distractor word sends additional activation to the same lexical representation, leading to a more unambiguous competitor (Lupker, 1979).

More recently, however, the competition account has been challenged, as a number of studies have failed to observe semantic interference effects even when the semantic distance between the target and distractor has been close (see details in Janssen et al., 2008). Importantly, these studies have often observed semantic *facilitation* effects (target naming is shortened by semantically related distractor words) instead of interference, which cannot be explained as easily by the lexical competition account.

A proposal for explaining the discrepant findings of the interference effect is, as mentioned, the *response exclusion hypothesis* (Dhooge & Hartsuiker, 2010; Janssen et al., 2008; Mahon et al., 2007; Marsh et al., 2017), stating that the locus of the interference effect is at a post lexical level of processing rather than at the lexical selection stage. Since only one word can be produced at any given time, the assumption is that there is a single channel output buffer which a single representation of a word occupies and awaits being produced. The response exclusion hypothesis explains the semantic interference effect in terms of how efficiently distractor words are excluded from the buffer (Mahon et al., 2007). As in the classic Stroop interference task, a written distractor word (production-ready representation) occupies the single-channel buffer more easily than a target word represented in other formats, such as a colour or a picture (Glaser & Glaser, 1989). Thus, the distractor has to be excluded from the buffer to make room for the target word, and certain (response relevant) criteria, such as the provenance (picture to be named) or the coarse semantic category of the distractor, determine how quickly it can be excluded (Janssen et al., 2008). This theory also explains the facilitation effect in terms of the magnitude of target-distractor priming, and the target naming time is not affected by the level of lexical activation of any non-target words (Mahon et al., 2007).

Effects of orthography on picture-naming

Although both the facilitation and the interference effects are examined in various different languages, the impact of orthography on these effects has not been thoroughly examined yet, even though this may have important implications for the understanding of speech production. Most theoretical models of speech production work under the assumption that activated units at each stage of processing (e.g., lexemes, morphemes, phonemes) get passed on to the next stage (e.g., Levelt, 1989; Levelt et al., 1999). However, there may be considerable differences between alphabetic and logographic languages in terms of, for example, the functional size of phonological units (such as phonemes vs. syllables), how these are segmented (Verdonschot et al. 2011a, b), and how they are represented orthographically. Such differences can be reflected in speech production behaviour. For instance, in an implicit priming task, it was found that speakers of languages with alphabetic orthographies (such as Dutch) were faster at reading aloud words that shared the

same first phoneme compared to words with dissimilar first phonemes (e.g., Meyer & Schriefers, 1991; Roelofs, 2006). In contrast, Mandarin speaking Chinese participants showed no such effect at the phoneme-level, but did show a facilitation effect for words that shared the same first syllable (Chen et al., 2002).

The effect of orthography in the picture-word task was observed over three decades ago by Rusted (1988) who compared the logographic script Chinese and the alphabetic script English. Rusted found that a group of Chinese-English bilinguals who performed the picture-word task showed different degrees of interference depending on the orthography of the distractor words (Chinese vs. English). Specifically, there was larger interference from Chinese than English distractors, regardless of whether the participants named in English or in Chinese. Rusted suggested that these results were due to the characteristic of the orthographic differences between the two languages. Logographic scripts are likely to have a more direct access to semantics than alphabetic scripts, and so, written Chinese words produced stronger competition at lexical selection stage than written English words did.

More recent studies have further shown clear differences between alphabetic languages and logographic languages (such as Chinese) in terms of the additivity of semantic, phonologic and orthographic effects in the picture-word task. The magnitude of facilitation/interference effects caused by the semantic characteristics of distractors and form-based characteristics of distractors (i.e., orthographically similar) are not typically additive in alphabetic languages, but have been found to be clearly additive in Mandarin (e.g., Zhu et al., 2015, 2016). This is likely because phonology and orthography in alphabetic languages tend to overlap, but are considerably easier to separate in logographic scripts (Zhu et al., 2016). Supporting this notion, Zhang and Weekes (2009) investigated the role of orthography and phonology in the picture-word task with Mandarin speakers by using distractor words that were either orthographically or phonologically similar to the target picture, while varying the stimulus onset asynchrony. The authors found that the temporal locus of orthographic facilitation occurred earlier than that of phonological facilitation (Zhang & Weekes, 2009), suggesting that the two processes do not intrinsically overlap in Mandarin. They also suggested that these processes are likely to be separate for most languages, but that they tend to be more difficult to separate in alphabetic languages.

Thus, there is clear indication that the effects that are found when using alphabetic scripts do not necessarily translate perfectly to other orthographies, even though orthography is an important component in any model explaining effects found in the picture-word paradigm. That, in turn, must also mean that the study of alphabetic languages is an inconclusive method of investigating orthographic effects in the picture-word task. Comparing logographic and phonological (typically alphabetic) scripts, however, is often associated with a fundamental problem, which is that the two types of script typically belong to different languages. This also limits participants to bilinguals, and therefore, numerous variables, such as language proficiency, the specific combination of native and second language, and the educational experience of the speaker, can impact the results. Unbalanced proficiency of the two languages was, indeed, offered as one explanation for the results found in Rusted's (1988) study (where the Chinese language was likely to have been more dominant than English for those bilingual participants). In order to eliminate the use of two

different languages to investigate the role of orthography in lexical facilitation and interference in the picture-word task, the current study implemented the multiscriptal language Japanese. The Japanese writing system is composed of both a logographic script, Kanji (the Japanese word for Japanese, ‘nihongo’ written in Kanji: 日本語), and two phonological orthographies Hiragana (‘nihongo’ written in Hiragana: にほんご) and Katakana (‘nihongo’ written in Katakana: ニホンゴ), collectively called Kana scripts. Further, Japanese can be typed out phonologically using the Roman alphabet, referred to as Romaji (the Japanese word for Japanese would in Romaji be typed using the Latin alphabet, i.e., nihongo). Crucially, different orthographic forms of the same word (as written using the different Japanese scripts) all map onto the same phonological form in speech, within the same language.

The traditional view of the Japanese writing system has been that the logographic Kanji and the phonological scripts Hiragana and Katakana are processed differently. More specifically, it has been argued that Kanji uses a more whole-word approach and is read via a semantic-lexical route, and in a more semantics-to-phonology manner, while Hiragana and Katakana are read in a more direct phonological fashion (e.g., Feldman & Turvey, 1980; Morton & Sasanuma, 1984; Sasanuma, 1975; Sasanuma & Fujimura, 1971; Wydell & Kondo, 2015). Recently, however, Dylman and Kikutani (2018) asked native Japanese speakers to switch between reading the different Japanese scripts, and found qualitative differences between Hiragana and Katakana processing, where Hiragana processing more resembled Kanji processing. Across four experiments, they asked participants to either read aloud or to make semantic decisions about pairs of Japanese words written in Kanji, Hiragana or Katakana. Any two words presented were either typed in the same Japanese script (e.g., Kanji-Kanji) or in different scripts (e.g., Kanji-Hiragana), and the processing cost was measured for when there was a switch from one script to another. One major finding was that Kanji reading was significantly more effortful in the reading aloud task, but quicker in the semantic decision task when the Kanji words were not required to be produced in speech, but processing remained at the semantic level. Another major finding was that a clear processing cost (observed as a delay in reaction times) was found when switching between the two phonological scripts Hiragana and Katakana, indicating that different mechanisms may be involved in reading the two scripts, and that the underlying mechanisms are not as similar as previous research had suggested.

The current study

The current study aimed to investigate the role of orthography in the picture-word task by varying the distractor script, using the multiscriptal language Japanese. The present research consists of five experiments. Experiment 1 investigated the magnitude of identity facilitation from the different Japanese scripts. Native Japanese speakers named pictures superimposed with written distractor words which were either the name of the object (identity condition) in the different Japanese scripts (Kanji, Hiragana, Katakana, and Romaji), or unrelated control words in each of these scripts (unrelated condition). Although the primary aim of this experiment

is to examine whether the magnitude of facilitation differs across scripts, the comparison of naming times among the scripts in each condition also provides insight into the underlying mechanisms behind the facilitation/interference effect in the picture naming task. As Rusted (1988) argued, naming a picture accompanied with its name written in the script with a more direct access to semantics should be easy. In the case of Japanese, Kanji is semantically much more unambiguous than the Kana scripts. In contrast, words written in Kana cannot distinguish between homophones without contextual information (Dylman & Kikutani, 2018). As the Japanese language contains a large number of homophones, this would, generally speaking, lead to higher ambiguity in terms of assessing the meaning of words written in Kana.³

In Experiment 2, Japanese-English bilinguals named pictures, once in their native Japanese, and once in their second language, English. The distractors were the name of the target picture typed in English, or in each of the Japanese scripts, Kanji, Hiragana, and Romaji (again, with unrelated distractors in each script). Lexical selection is thought to be harder for bilinguals than for monolinguals as concepts may map onto different words in the two languages (indeed referred to as the “hard problem”, Finkbeiner et al., 2006). For example, a picture of a dog will activate both ‘dog’ and ‘hund’ for a bilingual speaking English and Swedish, and a control mechanism for choosing the intended word in the intended language is required. A number of hypotheses have been proposed regarding the process of lexical selection in bilinguals, including the Inhibition Control model (Green, 1998) suggesting that translation equivalents in the non-target language are actively inhibited. However, this model cannot fully account for findings of lexical facilitation from cross-language translation distractors (e.g., Costa et al., 1999), and subsequent models have suggested that translation equivalents in both languages will be activated, and that the most highly activated lexical representation (regardless of language) will ultimately be selected for production (e.g., Finkbeiner et al., 2006).

Both the lexical selection by competition account and the response exclusion hypotheses have difficulties explaining certain empirical findings where bilinguals have performed the picture-word task (see details in Hall, 2011). For example, in the case of bilinguals naming a picture in one language, say in English, a distractor word identical to the name of the picture but written in another language, say in Spanish (e.g., DOG + perro), reduces the time to produce the accurate response (DOG) instead of slowing it down (e.g., Costa et al., 1999). At the same time, translation distractors that are not identical to pictures but semantically related to them (e.g., DOG + gato) interfere with target naming. According to the competition hypothesis, the direct translation of the name of the picture should be a strong competitor in terms of lexical activation and, thus, should produce interference. The response exclusion hypothesis can explain this facilitation from cross-language

³ Of course, there are instances where individual non-homophones are written in Kana, and as such are semantically unambiguous (this may be especially true for longer, three- or four-mora words). As such, any strategy based on the general nature of the Japanese language will not necessarily encompass all individual words, but the argument here is that the high frequency of homophones in the Japanese language has necessitated the use of a script where homophones can be distinguished clearly and unambiguously (namely, Kanji), and which is more unambiguous than words written in Kana.

translation distractors by using the mechanism of response relevant criteria. When naming in English, Spanish words do not meet the requirements for accurate responses, and so they will be excluded from the articulatory buffer fairly quickly. However, Spanish distractors still produce semantic priming for target English words, which leads to facilitation. This hypothesis fails to account for the findings that semantically related Spanish distractors interfere with, rather than facilitate, English naming.

Recently, Dylman and Barry (2018) presented a framework for understanding the processes involved in spoken word production as investigated through the picture-word task, combining established models of object naming (e.g., Humphreys et al., 1988) and reading (e.g., Coltheart et al., 2001), highlighting specifically the interaction between these picture naming and word reading processes. This framework firstly proposes that both the picture itself and the written distractor word (via the semantic-lexical route) send activation to the semantic representations. Secondly, the written distractor word activates a phonological lexical representation of the word (via the direct lexical reading route). Thirdly, for bilingual speakers, there are direct and bi-directional connections between translation-equivalents (inter-lexical translation connections), ensuring co-activation (if to varying degrees) of the target picture's name in both languages. Finally, the model incorporates a supervisory control mechanism (C-system) that sends additional activation to the target in the intended language, as well as ensuring that the task requirement of naming the picture and ignoring the written word is met. The lexical node that finally receives the highest amount of activation (out of a cohort of activated semantic representations with differing levels of activation) is subsequently selected for production. Although this model assumes that the level of activation of each lexical node plays significant role on speech production, it also parallels the response exclusion hypotheses in some aspects. For example, the C-system of this model is supposed to govern response relevant criteria.

Dylman and Barry's model (2018) was supported by the results from three experiments testing different groups of bilinguals and two experiments testing monolingual groups using distractor words that were translation-equivalents of the target pictures (common names of the pictures and their alternative synonyms were used for monolinguals, e.g., glasses-spectacles). Across all five experiments, the most important pattern found was that distractors written in L2 (or alternative synonyms for the monolinguals) had little or no effect on naming latencies when naming the picture in L1 (or common names for the monolinguals), while significant facilitation from the L1 distractors when naming in L2 was consistently found. This reflects a crucial aspect of the model's inter-lexical translation connections, which is that the net flow of activation is always greater from L1 to L2 than from L2 to L1. The strength of the inter-lexical translation connections is, therefore, asymmetric.

When investigating bilingual lexical selection using the picture-word task, the majority of studies have tested languages with the same or similar orthographies (e.g., Catalan-Spanish, Costa et al., 1999; Dutch-English, Hermans et al., 1998; Spanish-English, Costa & Caramazza, 1999; Swedish-English, Dylman & Barry, 2018), despite reading being an integral part of the processes involved during the picture-word task. Therefore, the second aim of the current study is to investigate the

replicability of Dylman and Barry's (2018) findings of asymmetric identity facilitation in bilingual naming in an additional (and non-alphabetic) language by targeting two languages differing drastically in orthography, namely Japanese and English.

Experiment 3 investigated the semantic interference effect with Japanese monolingual speakers, by using distractor words that are semantically related to the target picture (e.g., DOG + cat). In this experiment, three types of the semantic distance between the target and the distractor were employed; close (e.g., GOAT + cow), distant (e.g., GOAT + whale) and unrelated (e.g., GOAT + flower), and distractor words were written in Kanji, Hiragana, or Romaji. In addition to these three script conditions, this experiment presented a distractor as a picture rather than as a written word. The lexical selection by competition hypothesis would predict that target naming time increases as the semantic distance between the target and distractor increases, and a number of studies support this account (e.g., Schriefers, et al., 1990). Other studies, however, have found that certain semantically related distractors either do not produce interference, or, in fact, facilitate naming (e.g., semantically related verbs such as BED + sleep, or semantically close distractors of the same category, such as HORSE + zebra; see Mahon et al., 2007), supporting the response exclusion hypothesis (the distractors produce semantic priming but are then efficiently excluded from the articulatory buffer in accordance with the response relevance criteria applied during the task).

Kanji distractors would activate the lexical node of the distractor more quickly and strongly than Hiragana distractors (due to more speedy semantic access), and so Kanji should be more prone to interference according to the competition account. Semantically close distractors in particular should have a large impact on target naming time. In contrast, the response exclusion hypothesis would predict faster naming times in the Kanji condition, in general, as well as smaller interference. The picture condition in this experiment, where the distractors are small pictures presented within target pictures, would give a good indication about how Kanji is processed in this task.

The picture condition of Experiment 3 derives from Damian and Bowers (2003) who compared picture naming when printed words or smaller pictures were presented as distractors, in order to show that the picture-word interference effect reflects a lexical rather than a purely semantic or conceptual level process. They found an interference effect when the distractors were semantically related printed words, but no effect when they were semantically related smaller pictures. Other studies have also failed to observe semantic interference with picture-picture stimuli. However, the absence of interference does not necessarily mean that the distractor picture is not lexicalized. For example, Morsella and Miozzo (2002), who presented participants with two partially overlapping line drawings (one in green ink and one in red ink), and asked them to name the object in green (and ignore the one in red), found that naming times were faster when the distractor picture's name was phonologically related to the target (e.g., BED + *BELL*), which suggests that the distractor pictures, indeed, were lexicalized (and activated the phonology of their names). According to the response exclusion hypothesis, distractor pictures do not slow down target naming times not because they are not lexicalized, but because their representations (which are not in production-ready formats) do not occupy the

articulation buffer in a same way as written words. The naming time for the pictures with Kanji distractors, therefore, imply whether Kanji characters, being logograms, are similar to small pictures and, as such, are processed more like pictures (e.g., Park & Arbuckle, 1977). Experiment 4 will further investigate this in a group of Japanese-English bilinguals.

Experiment 5 aims to more closely understand the case of Romaji in the Japanese writing system. While Romaji is often used by Japanese speakers, it is characteristically different from the other Japanese scripts. Words written in Romaji can be read easily (like Kana words) as Romaji is a phonological script. However, it seems to have weaker access to semantics, possibly because the other Japanese scripts are more strongly associated with their respective semantic representations. These characteristics resemble pseudohomophones in English, such as “caik” (cake) or “sheap” (sheep), which are pronounceable words if read out phonologically, but do not match the exact orthographic format of the semantic representations of said word. Experiment 5, thus, tested English monolinguals performing a picture-word task, to examine whether the extent of semantic facilitation obtained from pseudohomophones (e.g., SHEEP + sheap) resembles the facilitation observed for Japanese participants in Experiment 1. The aim of this experiment was to increase our understanding of the nature of Romaji processing mechanisms, by attempting to simulate the effects observed for Romaji in a group of English monolinguals.

To sum up, the current study attempts to investigate the role that orthography plays in the picture-word task, whilst taking out methodological issues that have limited the conclusions of previous studies investigating various orthographic effects (such as predominantly studying alphabetic languages, or studying bilinguals with languages using different orthographies, making it difficult to control for the various aspects that may be at play, including but not limited to relative language proficiencies in each language, language exposure, educational aspects, and cultural context). By using the multiscriptal language Japanese, we can directly measure orthographic effects, for both the identity facilitation and the semantic interference effect (both of which are commonly studied phenomena in the speech production literature), within the same language. Additionally, we aimed to test the generalisability of the underlying mechanisms of these processes by attempting to simulate the findings from the Romaji condition in a group of native English speakers by manipulating the spelling of real English words. Further, we aimed investigate the applicability of Dylman and Barry’s (2018) model of speech production for bilinguals with a non-alphabetic language as their native language. Finally, we attempt to adjudicate between two opposing theoretical accounts of speech production, namely the lexical selection by competition account and the response exclusion hypothesis.

Experiment 1: Lexical facilitation in Japanese monolinguals

Experiment 1 investigated lexical facilitation by using as distractors the name of the target object to be produced. These identity distractors were printed in Kanji, Hiragana, Katakana and Romaji, and were coupled with unrelated control distractors in the same four scripts. The participants’ naming latencies were measured

in order to investigate the magnitude of potential identity facilitation. Differences in facilitation between the different scripts will be informative of the underlying processing mechanisms of the different Japanese orthographies.

The lexical selection by competition account and the response exclusion hypothesis would predict different patterns of result for Experiment 1. As laid out in the Introduction, based on previous studies, accessing the semantics of a distractor word typed in Kanji should be faster than the same distractor word written in Hiragana or Katakana. Thus, target naming times in the identity condition should be shorter in the Kanji condition. In contrast, when the word is not related to the picture, this advantage for Kanji may disappear. The unrelated distractor word would activate a different lexical node than the target node, creating a certain level of competition at the lexical selection stage, but it would not activate the target name. The activation of the unrelated distractor word might be stronger for Kanji than for Kana, due to the semantic unambiguousness of Kanji words, thus leading to a larger delay for Kanji, as the unrelated Kanji distractor would more quickly activate the semantics of said distractor, compared to the same unrelated word written in a Kana script. Indeed, it has been found that unrelated real-word distractors delay target naming more than unrelated non-word distractors (Klein, 1964), suggesting that quick and precise access to semantic information contributes to a stronger activation of the lexical node. When the written word does not match the picture, such robust access to semantics may work detrimentally to the target naming (compared to the same distractor word typed in Kana). While the unrelated Kana words in this Experiment are real words, words written in Kana are, generally speaking, more semantically ambiguous (due to the number of homophones in the Japanese language), and may therefore not send as strong and unambiguous activation to the semantic level (as compared to Kanji), and thus the unrelated Kana distractors should be weaker competitors than the unrelated Kanji distractors. These predictions are in line with the lexical selection by competition hypothesis. The response exclusion hypothesis, however, may predict a different pattern of results, which we will lay out below.

Phonological scripts can be produced more easily than logographic scripts because they already are production-ready representations. It is often found that reading the logographic script Kanji is more effortful than reading the phonological scripts Hiragana or Katakana (e.g., Dylman & Kikutani, 2018). This has been attributed to the unique characteristic of Kanji, where each character can take multiple phonological forms, so-called ON- and KUN-reading (e.g., Verdonschot et al., 2011a, b, 2013). Therefore, in the identity condition, a Kana distractor word may occupy the single-channel articulation buffer more quickly than a Kanji word, leading to faster naming of the target picture. The opposite will occur in the unrelated condition. Since the logographic script Kanji is not a production-ready representation, any words written in Kanji may be processed more like pictures than written words (Park & Arbuckle, 1977). If so, due to the provenance of the representation, the Kanji unrelated distractor should be swiftly excluded from the buffer, while Kana distractor would take longer to be excluded. Experiment 1, therefore, could not only specify the impact of orthography on the degree of facilitation effect, but could also shed light on the locus of the effect.

Method

Participants

Thirty native Japanese speakers (24 women and 6 men) living in Japan participated in Experiment 1. All were undergraduate university students. Their mean age was 20.10 years ($SD = 1.19$). All participants had normal or corrected-to-normal vision, and none reported having any reading difficulties.

Stimulus materials

Thirty pictures were selected from Snodgrass and Vanderwart (1980), chosen on the basis of having high name agreements ($M = 90\%$), derived from a pilot study with 17 (other) native Japanese speakers living in Japan, and having short and relatively common Kanji words consisting of only one or two Kanji characters. All the Kanji characters used in this experiment were very familiar to adult readers and were taken from the list of “regularly-used Kanji characters” called the Jouyou list issued by Japanese Ministry of Education in 1981. Each target picture was paired with its name as a distractor word, and these words were rearranged to create unrelated control words for each target picture. Each distractor was then presented in the four different scripts Kanji, Hiragana, Katakana, and Romaji (see Appendix 1 for a full list of stimuli).

Procedure

The participants were tested individually on a MacBook Air laptop (13" screen) using the software program SuperLab version 4.5. Participants were asked to name the pictures as quickly and as accurately as possible, whilst trying to ignore the superimposed printed words. The pictures, when presented on the computer screen, were between 8×5 cm and 4×3 cm large. In order to familiarize the participants with the target pictures prior to the experiment, the participants were first shown all the target pictures together with their names. Following this familiarization, they received four practice trials, followed by the 240 experimental trials (30 pictures $\times$ 8 distractors) presented in a randomized order. The pictures and distractors remained on the screen until the participant produced a vocal response, which was detected by an Elecom usb-headset connected to SuperLab's voice-key. The experimenter then classified each response as being correct or incorrect, and the next trial began 1 s later.

Results and discussion

Responses that were incorrect, hesitant or non-verbal (e.g., “er”, “uh”, etc.), or failed to stop the timer were excluded, as were responses below 250 ms or above 3 s. Harmonic means of non-excluded responses were calculated for each participant

Table 1 Results from Japanese monolinguals in Experiment 1. Mean object naming latencies (in milliseconds), standard deviations, and rates of excluded responses (%E) by distractor script and distractor type. Also shown are the identity facilitation effects observed for each script

	Identity			Unrelated			Facilitation
	M	SD	%E	M	SD	%E	
Kanji	757	113	1.3	855	113	4.2	98**
Hiragana	767	111	0.7	887	114	3.1	120**
Katakana	775	113	1.7	886	113	2.8	111**
Romaji	789	116	1.0	846	122	1.6	57**
Combined mean	772			869			

** $p < .001$

In order to examine whether the RTs for the identity and unrelated conditions statistically differ for each script, a two-way ANOVA with the factors of script (4 levels) and distractor type (2 levels) were performed. The p-values were derived from a test of simple main effect examining the impact of distractor type for each script. This procedure was chosen over performing multiple paired sample t-tests with unadjusted error rates

and for each item, in each condition. Harmonic means were chosen because they involve a transformation which reduces the skew of naming latency distributions towards slower responses (see Ratcliff, 1993), avoiding using various cut-offs that often involve excluding data. Table 1 shows the results.

The reaction times were analysed using linear mixed effects (LME) models fit by REML. The calculations of the p-values were based on the Satterthwaite's approximation for the degrees of freedom. R version 3.3.0 (R Development Core Team, 2008) with the *lme4* package (Bates et al., 2015) was used. We entered fixed effects of relatedness (identity vs. unrelated) and script (Kanji, Hiragana, Katakana, Romaji), as well as the possible two-way interaction into the model. Participants, pictures and distractor words were treated as random factors. As random effects, we had intercepts for those three factors, as well as by-participant, by-picture, and by-word random slopes for the fixed effects. The analysis revealed a main effect of relatedness ($\beta = 55.67$, $SE = 9.17$, $t(6930) = 6.07$, $p < .001$); naming times for the identity distractors were faster than the control distractors (772 vs. 869 ms). There was also a significant main effect of script ($\beta = 51.99$, $SE = 5.48$, $t(704) = 9.49$, $p < .001$); naming times were the fastest for Kanji distractors (806 ms), slightly slower for Romaji distractors (817 ms), even slower for Hiragana distractors (827 ms), and the slowest for Katakana distractors (830 ms). However, these effects were modified by the significant interaction between relatedness and script ($\beta = -19.15$, $SE = 3.46$, $t(1020) = 5.53$, $p < .001$).

Simple main effects examined lexical facilitation by comparing the identity and the control distractor, at all four levels of script. Significant facilitation effects were found for all the scripts [Kanji trials (98 ms), $\beta = 21.34$, $SE = 6.12$, $t(1404) = 3.49$, $p < .001$, $d = 0.87$; Hiragana trials (120 ms), $\beta = 24.70$, $SE = 8.03$, $t(1221) = 3.08$, $p = .002$, $d = 1.05$; Katakana trials (112 ms), $\beta = 24.65$, $SE = 6.02$, $t(1271) = 4.09$, $p < .001$, $d = 0.99$; Romaji trials (57 ms), $\beta = -42.37$, $SE = 8.20$, $t(1698) = 5.17$, $p < .001$, $d = 0.47$].

An analysis of the proportion of excluded responses, using the equivalent LME model as was used for the naming latencies, found no significant effects.

As the LME analysis for reaction times showed no impact of participants or items, we calculated the magnitude of the facilitation for each participant for the following analysis. A one-way repeated-measures ANOVA on the facilitation (measured as RT difference between the identity and unrelated condition for each script) showed a significant difference between the scripts, $F(3, 87) = 21.53$, $p < .001$, $\eta_p^2 = 0.43$. Pairwise comparisons with Bonferroni corrections showed that the facilitation effect in the Hiragana condition (120 ms) was significantly larger than in both the Kanji (98 ms), $p = .020$; and the Romaji conditions (57 ms), $p < .001$; but was not significantly different from the Katakana condition (112 ms), $p = 0.157$. Further, the facilitation effect in the Katakana condition was significantly larger than the Romaji condition, $p < .001$; but not significantly different from the Kanji condition, $p = .105$. Finally, the facilitation effect was significantly larger for Kanji than for Romaji, $p < .001$.

When naming time for the identity and unrelated conditions is concerned, the lexical selection by competition account predicts that the time for the identity condition is faster for Kanji distractors than Kana, while the opposite pattern is predicted by the response exclusion account. For the unrelated condition, the competition account predicts shorter naming time for Kana distractors, while the exclusion account predicts shorter time for Kanji distractors. In order to assess which pattern is present, two separate one-way ANOVAs were conducted for the effect of script for each relatedness condition, but no effect was found for either condition ($F_s < 1$).

The results from Experiment 1 found clear facilitation effects from identity distractors in all four of the Japanese scripts with Hiragana and Katakana distractors showing the largest facilitation, followed by the Kanji distractors, with the Romaji distractors producing the smallest (but still significant) facilitation relative to the unrelated control distractors in each respective script. Stronger facilitation for the phonological script Hiragana compared to for the logographic script Kanji highlights the importance of phonological processing in the picture naming task. Although Kanji should have quicker access to semantics as compared to Hiragana, the reaction times for Kanji in the identity condition was only 10 ms quicker than those for Hiragana. Interestingly, the reaction time for the unrelated condition showed a larger difference between the two script, where Kanji was 30 ms quicker than Hiragana. This may reflect that unrelated distractors written in Kanji were dismissed from the articulatory buffer more quickly than Hiragana distractors. However, these differences are not statistically significant and so this interpretation requires some level of caution. Nonetheless, the general tendency of the results possibly resembles the prediction based on the response exclusion account.

Experiment 2: Lexical facilitation in Japanese-English bilinguals

In Experiment 2, native Japanese speakers with English as a second language were asked to name pictures once in Japanese, and once in English. The superimposed distractor words used were the name of the object (identity distractors) in English,

and in Japanese Hiragana, Kanji and Romaji, and unrelated controls for each of these four conditions. As Experiment 1 did not find any significant difference in the magnitude of the facilitation between Hiragana and Katakana, Experiment 2 did not include a Katakana condition.

According to Dylman and Barry's model, distractor words written in English (L2) should not impact naming latencies in Japanese (L1), while Japanese words should influence naming in English. The further questions regarding orthographies are whether the findings of Experiment 1 will be replicated, and whether the pattern found when naming in Japanese (L1) is also observed when naming in English (L2). Specifically, based on the results from Experiment 1, we expect different magnitudes of facilitation from the different Japanese scripts, such that the facilitation effect from the Kana distractors are expected to be larger than from Kanji distractors. Interestingly, Romaji distractors in this experiment share perceptual similarity with English words (using the same alphabet), but are phonetically dissimilar to the actual to-be produced words, being cross-language distractors, when naming in English. The Romaji distractors, then, may offer insight into the level of impact using the same orthographical script (albeit in different languages) will have on naming latencies in producing speech.

Method

Participants

Sixteen Japanese-English bilinguals (14 female, 2 male) participated in this experiment. Their mean age was 26.9 years ($SD=5.2$ years), ranging from 20 to 41 years. All participants had normal or corrected-to-normal vision. All were native Japanese speakers, but had been living and studying at a university in the UK for at least one year at the time of participation, which was conducted in the UK. Prior to the experiment, all participants completed the reading component of the Dialang English language assessment. The results revealed that 14/16 agreed with the statement that "I can understand articles and reports concerned with contemporary problems in which the writers adopt particular stances or viewpoints", 14/16 agreed with the statement "I can understand any correspondence with an occasional use of dictionary", and 10/16 agreed with "I can understand in detail a wide range of long, complex texts provided I can re-read difficult sections". This indicates a high level of L2 English proficiency. None of the participants reported having any reading difficulties in either language.

Stimulus materials

In order to test the reliability of the findings of Experiment 1, we used a different set of pictures and words for this experiment. Twenty-four pictures were newly chosen on the basis of having high name agreements (> 90%) derived from the same pilot study with native Japanese people (living in Japan) as was used in Experiment 1. Each target picture was paired with two different types of words: identity (e.g.,

CLOUD + cloud), and a matched unrelated control word (e.g., CLOUD + goat). The Japanese distractor words were chosen on the basis of having short Kanji versions of a maximum of two Kanji characters, and having common, or easily recognizable Kanji characters. The Kanji characters used for 18 words out of 24 were listed in the Jouyou list, and a small pilot study with native Japanese people living in Japan confirmed that the non-listed characters were as familiar as the listed ones for Japanese readers.

Each distractor for each target picture was presented in four ways: in Japanese Kanji (e.g., 雲); Japanese Hiragana (e.g., くも); Japanese Romaji (e.g., kumo); and in English (e.g., cloud). Each target picture was thus presented 8 times (see Appendix 2 for a complete list of all the distractors). Four additional pictures were used for practice trials. The English and Romaji distractor words were written in Verdana size 30, and the Kanji and Hiragana distractors were written in MS Gothic size 30.

Procedure

The experiment was run on the same equipment as used in Experiments 1, using the same general procedure. The participants completed two blocks of trials: once naming all the stimuli in Japanese, and once in English (with the order counterbalanced). Before each block, the participants were made familiar with all the pictures that would be presented, together with their names written in the target language of the subsequent block, and were asked to name each picture in that language. After having completed four practice trials, the participants were presented with the 192 experimental trials, presented in a randomized order. Participants were given a short break after one third of the trials and again after two thirds of the trials.

Results and discussion

The data were treated in the same way as in Experiment 1, using the same criteria for exclusion. Table 2 shows mean (SD, and % error) naming times for each condition, as well as the lexical facilitation effects calculated from these means for each script.

The naming latencies were analysed in the same way as in Experiment 1, using linear mixed effects (LME) models fit by REML, with the same specifications. Separate analyses were conducted for naming in Japanese and naming in English. The fixed effects of script in this experiment included Kanji, Hiragana, Romaji and English.

When naming in Japanese, there was a significant main effect of relatedness ($\beta = 217.92$, $SE = 23.93$, $t(127) = 9.11$, $p < .001$); naming times were faster with same name (758 ms) than unrelated control distractors (848 ms). The main effect of distractor script was also significant, ($\beta = 87.44$, $SE = 12.43$, $t(431) = 7.03$, $p < .001$). Pairwise comparisons showed that naming times were faster with Kanji distractors (773 ms) than with Hiragana (807 ms), $p = .017$, Romaji (800 ms), $p = .005$, and English distractors (832 ms), $p < .001$, and that naming times were faster with Romaji than English distractors, $p = .002$. However,

Table 2 Results from Japanese-English bilinguals in Experiment 2. Mean object naming latencies (in milliseconds), standard deviations, and rates of excluded responses (%E) by naming language, distractor script and distractor type. Also shown are the identity facilitation effects for each condition

	Identity			Unrelated			Facilitation
	M	SD	%E	M	SD	%E	
<i>Naming in Japanese</i>							
Kanji	722	112	1.8	825	144	3.1	103**
Hiragana	720	113	1.0	894	123	3.1	174**
Romaji	762	131	3.3	839	152	2.3	77*
English	830	142	3.6	835	131	2.6	5
Combined mean	759			848			
<i>Naming in English</i>							
Kanji	797	86	1.8	800	85	1.8	3
Hiragana	808	96	2.6	840	92	3.1	32*
Romaji	815	95	2.6	812	83	2.9	-3
English	739	96	1.8	822	89	1.8	83**
Combined mean	790			819			

* $p = .002$, ** $p < .001$

The RTs for the identity and unrelated conditions were compared using a two-way ANOVA with the factors of script (4 levels) and distractor type (2 levels) performed separately for the Japanese naming data and English naming data. The p-values were derived from a test of simple main effect examining the impact of distractor type for each script

these effects were again modified by the significant interaction between relatedness and script ($\beta = -44.38$, $SE = 7.37$, $t(2910) = 6.02$, $p < .001$). An analysis of the simple main effects of this interaction revealed a significant facilitation effect (faster reaction times for identity over unrelated condition) for Kanji, $\beta = 126.38$, $SE = 17.73$, $t(23.09) = 7.13$, $p < .001$, $d = 0.80$, Hiragana, $\beta = 203.20$, $SE = 15.95$, $t(23.26) = 12.74$, $p < .001$, $d = 1.48$, and Romaji, $\beta = 81.32$, $SE = 19.59$, $t(22.94) = 4.51$, $p < .001$, $d = 0.55$, but not for English words, $p = .301$.

The impact of script on the magnitude of facilitation (measured as RT difference between the identity and unrelated condition for each script) was examined by a one-way repeated-measures ANOVA as in Experiment 1. The effect of script was significant, $F(3, 45) = 28.05$, $p < .001$, $\eta_p^2 = 0.65$. Bonferroni pairwise tests revealed that the facilitation magnitude for Hiragana was significantly larger than for all the other scripts. The impact of Kanji and Romaji did not differ from each other and both were larger than the facilitation caused by English script.

When naming in English, there was no significant main effect of relatedness ($\beta = -32.88$, $SE = 18.83$, $t(935) = 1.75$, $p > .05$), but there was a significant main effect of distractor script ($\beta = -47.62$, $SE = 10.89$, $t(1002) = 4.37$, $p < .001$). Pairwise comparisons showed that naming times were faster with English (780 ms) than Kanji (799 ms), $p = .047$, Hiragana (824 ms), $p < .001$, and Romaji distractors (813 ms), $p < .001$, faster with Kanji than Hiragana distractors, $p = .010$, and Romaji distractors, $p = .043$. There was also a significant script by relatedness interaction ($\beta = 27.78$, $SE = 6.89$, $t(1760) = 4.03$, $p < .001$). Simple main effects

revealed that the lexical facilitation effect was significant for English distractors, $\beta = 104.33$, $SE = 16.53$, $t(22.95) = 6.31$, $p < .001$, $d = 0.90$, and for Hiragana distractors, $\beta = 45.23$, $SE = 20.23$, $t(22.89) = 2.23$, $p = .035$, $d = 0.34$, but not for Kanji and Romaji distractors.

A one-way repeated-measures ANOVA on the magnitude of facilitation revealed a significant effect of script, $F(3, 45) = 15.17$, $p < .001$, $\eta_p^2 = 0.50$. Bonferroni pairwise tests revealed that the English script yielded the largest facilitation, but the facilitation in the Hiragana condition was also significantly larger than Kanji and Romaji condition, which did not differ from each other.

The proportion of excluded responses were analysed using the equivalent LME model as was used for the naming latencies. When naming in Japanese, there was a significant main effect of relatedness ($\beta = -0.03$, $SE = 0.01$, $t(2405) = -2.03$, $p < .01$): there were more excluded responses in unrelated conditions. There was also a significant main effect of script ($\beta = -0.02$, $SE = 0.01$, $t(401) = -2.14$, $p < .01$): Hiragana trials had significantly lower number of excluded responses than other three scripts. There was no significant interaction. Finally, the analysis for the trials when naming in English found no significant results.

In order to see whether the pattern of naming time for identity and unrelated conditions predicted by the lexical competition or the response exclusion accounts were present, one-way ANOVAs for the effect of script were run for each level of relatedness. When naming in Japanese, no significant effect of script was found for the identity condition, $F(3, 60) = 2.48$, $p = .069$; or the unrelated condition, $F < 1$. Likewise, when naming in English, no significant effect of script was found for the identity condition, $F(3, 60) = 2.21$, $p = .097$; or the unrelated condition, $F < 1$.

The results from Experiment 2 replicate the phenomenon of within-language identity facilitation (e.g., Dylman & Barry, 2018) from all three Japanese scripts when naming in Japanese (L1), and from English distractors when naming in English (L2). When naming in Japanese, Hiragana distractors produced the largest facilitation effect, followed by Kanji and then Romaji, which is an identical pattern to the findings in Experiment 1. The results further replicated the asymmetrical cross-language effects of no facilitation from L2 (English) distractors when naming in L1 (Japanese), but significant facilitation from L1 (Japanese) distractors when naming in L2 (English). Interestingly, cross-language facilitation when naming in L2 was only found when the distractors were written in Hiragana, and no cross-language facilitation was found from distractors written in Kanji or Romaji, suggesting that the processing of these two scripts either does not send strong enough activation to lead to cross-language facilitation, or has a different temporal locus and does not overlap with crucial points in time for lexical facilitation to emerge.

Comparing Kanji and Hiragana, the reaction times for the identity condition show almost no difference while those for unrelated condition consistently show faster naming latencies for Kanji (40 vs. 50 ms) regardless of the naming language. This is exactly the same pattern as found in Experiment 1 and the results appear to support the response exclusion hypothesis. Again, however,

these differences failed to reach statistical significance, and thus, this is not conclusive.

Experiment 3: Semantic interference in Japanese monolinguals

Experiment 3 aimed to further investigate the role of orthography in the picture-word task by using semantically related distractor words. A different group of Japanese monolinguals were asked to name pictures with superimposed distractors, and the distractors were semantically close (e.g., GOAT + cow), distant (e.g., GOAT + whale) and unrelated (e.g., GOAT + flower), written in Kanji, Hiragana, and Romaji, and presented as a smaller picture superimposed on the target picture. Mahon et al. (2007) found that very closely related distractor (e.g., HORSE + zebra) words either did not produce an interference effect, or produced less interference than “less” closely related words (e.g., HORSE + whale), results which contributed to the proposal of the response exclusion hypothesis. These results were both surprising and conflicted with the results of Vigliocco et al. (2002) who found larger semantic priming as well as larger semantic interference effects in the more closely related condition. To investigate the semantic relatedness effect more closely, and avoid risking null effects due to the level of semantic relatedness, we decided to test both semantically close and distant distractors in Experiment 3.

The reasoning behind adding the condition with smaller pictures derives, as mentioned in the Introduction, from Damian and Bowers (2003) who compared picture naming when printed words or smaller pictures were presented as distractors in order to show that the picture-word interference effect reflects a lexical rather than purely semantic or conceptual level process. While some have suggested that Kanji characters, being logograms, are similar to small pictures and may, as such, be processed more like pictures (e.g., Park & Arbuckle, 1977), Flaherty (1993) argued that this is not the case. She found that Japanese readers, just as English readers, were slower to name pictures than to read both Kanji and Kana words, and that they were not significantly slower at reading Kanji words than Kana words. (Although Kana reading times were 59 ms faster than Kanji reading times, this effect was reported to be nonsignificant.)

As mentioned in the Introduction, again, the lexical selection by competition account and the response exclusion hypothesis would expect different patterns of results for Experiment 3. According to lexical selection by competition, Kanji distractors would activate the lexical node of the distractor more quickly and more strongly than Hiragana distractors (due to more speedy semantic access), and Kanji should, therefore, lead to more interference, particularly for semantically close distractors. The response exclusion hypothesis, however, would predict smaller interference (and faster naming times for Kanji distractors, in general). Following the results from Damian and Bowers (2003), we expect to find effects for the printed distractors but not the pictures. More importantly, based on the results from Experiments 1 and 2, we should expect a larger effect from Hiragana than from Kanji, and finally followed by Romaji.

Method

Participants

Twenty-two Japanese monolinguals participated in Experiment 3 (5 females and 17 males). All were native Japanese speaking undergraduate university students living in Japan. The mean age was 20.9 years ($SD=3.3$ years), ranging from 18 to 30 years. All participants had normal or corrected-to-normal vision.

Stimulus materials

Twenty pictures were chosen on the basis of high name agreements ($>90\%$) derived from a small pilot study with native Japanese people all living in Japan. Each target picture was then paired with three different types of items: a very closely semantically related word (e.g., CAT + tiger); a less close (or “distant”) but still semantically related word (e.g., CAT + crab); and a semantically unrelated word (e.g., CAT + eye). These distractor words were chosen on the basis of having short and relatively common (and thus easily read) Kanji versions consisting of only one or two Kanji characters. Each distractor was then presented in four different ways: in Japanese Kanji; Japanese Hiragana; Japanese Romaji; and finally, as a smaller picture.

Each target picture was thus presented 12 times. Appendix 3 gives a list of all the words written in Kanji, Hiragana, and Romaji, together with their English translations. The Romaji distractors were written in Verdana size 30, and the Kanji and Hiragana distractors were written in MS Gothic size 30.

Procedure

This experiment was run on an Apple MacBook computer, with a 13" screen, using the software program SuperLab. The procedure was largely the same as in Experiment 1, including instructions given to the participants, and familiarization with the target pictures prior to the experiment.

After the familiarization stage, the participants were given four practice trials, followed by 240 experimental trials presented in a randomized order. In each trial, the picture to be named appeared and remained on the screen until the participant gave a vocal response, which was detected by the computer's built-in microphone. The experimenter then classified each naming response as being correct or incorrect by pressing predetermined keys on the keyboard. The next picture appeared 1 s after the experimenter's response. The participants were given two breaks during the experiment: once after 80 pictures had been presented, and again after another 80 pictures had been presented.

Table 3 Results from Japanese monolinguals in Experiment 3. Mean object naming latencies (in milliseconds), standard deviations, and rates of excluded responses (%E) by relatedness and distractor type. Also shown are the semantic interference effects for each distractor type at both levels of semantic distance

	Close		Distant		Unrelated		Interference	
	M (SD)	%E	M (SD)	%E	M (SD)	%E	Close	Distant
Kanji	693 (68)	4.1	704 (87)	3.0	692 (62)	62	1	12
Hiragana	762 (70)	6.1	763 (84)	4.8	732 (69)	4.5	30*	31*
Romaji	722 (77)	5.9	715 (84)	3.9	720 (73)	3.0	2	-5
Picture	674 (71)	2.7	677 (67)	1.6	665 (64)	1.8	9	12
Combined mean	713		715		702			

* $p < .05$

Since the overall effect of distractor type was not significant in this experiment, the RT comparisons for the close against the unrelated, as well as for the distant against the unrelated conditions, were performed only for Hiragana, using paired-sample t-tests

Results and discussion

The data were treated in the same way as in Experiments 1 and 2, using the same criteria for exclusion. Table 3 shows mean (SD, and % error) naming times for each condition, as well as the semantic interference effects calculated from these means for each distractor condition.

The naming latencies were analysed in the same way as in Experiments 1 and 2, using linear mixed effects (LME) models fit by REML, with the same specifications. The fixed effects of script in this experiment included Kanji, Hiragana, Romaji and Picture.

There was no significant main effect of relatedness ($\beta = -7.89$, $SE = 6.78$, $t(220) = 9.11$, $p > .05$); but there was a significant main effect of distractor script, ($\beta = -14.11$, $SE = 5.07$, $t(4974) = 2.79$, $p < .001$). Post-hoc tests showed that naming times were fastest with the picture distractors (672 ms), followed by the Kanji distractors (696 ms), followed by the Romaji distractors (719 ms), and the slowest to name with the Hiragana distractors (752 ms). There was no significant interaction between relatedness and script ($\beta = 0.68$, $SE = 2.39$, $t(881) = 0.28$, $p > .05$).

Although no significant interference effect was apparent in the LME results, the Hiragana conditions show a tendency indicating a difference in naming time between unrelated and close conditions as well as between unrelated and distant conditions (see Table 3). Therefore, we compared each of the close and distant conditions against the unrelated condition using paired-samples t-tests. The differences were statistically significant for both comparisons; close: $t(21) = 2.66$, $p = .015$, $d = 0.43$; distant: $t(21) = 4.27$, $p < .001$, $d = 0.45$.

Semantic interference was present only for the Hiragana condition in this experiment, which nonetheless follows the pattern found in Experiments 1 and 2 that the picture naming is most strongly impacted by distractors typed in Hiragana. No semantic interference was found for Kanji and Romaji. Likewise, the smaller distractor pictures

did not produce any semantic interference. Regarding the semantic distance manipulation, there was no difference between the close and distant distractors; both (when typed in Hiragana) produced semantic facilitation of equal magnitude (close $d=0.43$; distant $d=0.45$, both indicating medium sized effect sizes).

The present experiment found consistently faster reaction time for Kanji than for Hiragana, and this is against the prediction by the lexical competition account, which assumes that semantically unambiguous Kanji distractors would trigger stronger activation of non-target representations than Hiragana distractors would manage. Furthermore, the reaction time for the Kanji condition resembles that of the picture condition, indicating that Kanji distractors were treated somewhat similar to pictures. The response exclusion hypothesis assumes that speakers apply response relevance criteria to judge whether distractors should be excluded from the articulatory buffer. Whether the distractors are easily producible or not might be such a criterion used, and thus the picture distractors (which are not represented in a production-ready format) can be excluded quickly, resulting in the absence of interference during target picture naming. The results of Experiment 3 suggest that the same criterion might be applied for Kanji distractors. Importantly, the response relevant criteria should, then, be the underlying mechanism of the semantic interference, and criteria related to semantic distance between the distractors and target should alter the timing of when the distractor is excluded from the articulatory buffer. However, the concept of response relevance criteria is rather vague, and it is not clear whether different sets of criteria can operate at the same time. Using the criteria related to production ease and to semantic distance simultaneously might be confusing and cognitively demanding. If so, there is a possibility that the implementation of the criteria related to semantic distance was weakened in this experiment, resulted in only a small semantic interference.

Experiment 4: Semantic interference in Japanese-English bilinguals

Experiment 4 attempted to replicate Experiment 3 in Japanese-English bilinguals. The participants were asked to name pictures once in their native Japanese and once in their L2 English. The target pictures were superimposed with semantically related distractors, which were either typed in Kanji, Hiragana, Romaji, or English. As in Experiment 3, there were two semantically related conditions, one close and one distant.

The aim of Experiment 4 was partly to investigate the semantic interference effect in a different bilingual group (and as such further explore issues raised in Dylman & Barry's model as well as in the bilingual literature in general), but mainly to investigate the role of orthography for semantic interference effects in the picture-word task, using the multiscript language of Japanese.

Method

Participants

Twenty Japanese bilinguals participated in this experiment (15 male, 5 female), with a mean age of 20.2 years ($SD=2.6$ years). All were native Japanese speakers living in Japan, and all had normal or corrected-to-normal vision. The participants completed the English language assessment Dialang, the results of which found that 17/20 participants agreed with advanced level statements such as “I can understand any correspondence with an occasional use of dictionary” and “I can understand in detail a wide range of long, complex texts provided I can reread difficult sections”, indicating good levels of English proficiency.

Stimulus materials

The stimuli used in Experiment 4 consisted of the same target pictures and distractors as used in Experiment 3, with the exception that the distractors that were a smaller picture in Experiment 3 were changed to the English name of those pictures.

Procedure

The procedure of Experiment 4 was essentially identical to that of Experiment 3, with the exception that the participants in Experiment 4 were asked to name the pictures twice; once in Japanese, and once in English (the order of which was counterbalanced). They were given four practice trials, followed by 240 experimental trials presented in a randomized order. The participants were given two breaks during each naming block (after 80 trials). After the completion of the experiment, the participants were presented with all the Kanji characters that had been used as distractor words during the experiment, and were asked to rate them for familiarity on a scale from 1—*Never seen/Cannot read the Kanji* to 5—*Very familiar*. They were asked to do the same rating for all the English words that had been presented during the experiment. Out of the 20 participants, 18 were familiar with over 75% of the Kanji characters, and the remaining two were familiar with the majority of the Kanji characters. For the English words, 19 (out of 20) were familiar with over 75% of the English words, with the remaining participant being familiar with the majority of the English words.

Results and discussion

The data were treated in the same way as in the previous experiments. Descriptive statistics are shown in Table 4.

The naming latencies were analysed in the same way as in Experiments 1 and 2, using linear mixed effects (LME) models fit by REML, with the same specifications. The fixed effects of script in this experiment included Kanji, Hiragana, Romaji and

Table 4 Results from Japanese-English bilinguals in Experiment 4. Mean object naming latencies (in milliseconds), standard deviations, and rates of excluded responses (%E) by relatedness and distractor type. Also shown are the semantic interference effects for each distractor type at both levels of semantic distance

	Close		Distant		Unrelated		Interference	
	M (SD)	%E	M (SD)	%E	M (SD)	%E	Close	Distant
<i>Naming in Japanese</i>								
Kanji	786 (98)	3.5	780 (103)	2.3	773 (105)	5.3	13	7
Hiragana	840 (105)	7.3	836 (90)	3.3	816 (88)	5.0	24	20
Romaji	792 (96)	5.0	786 (97)	5.3	791 (73)	4.5	1	-5
English	804 (95)	5.8	797 (101)	5.3	792 (93)	4.0	12	5
Combined mean	806		800		793			
<i>Naming in English</i>								
Kanji	822 (71)	2.0	822 (79)	2.3	817 (84)	4.8	5	5
Hiragana	847 (96)	5.3	860 (78)	3.0	845 (70)	5.0	2	15
Romaji	828 (70)	5.3	832 (82)	4.5	842 (74)	2.5	-14	-10
English	842 (80)	5.0	834 (88)	4.0	837 (84)	3.8	5	-3
Combined mean	835		837		835			

English. The data were analysed separately for naming in Japanese and naming in English. The analyses revealed no significant effects, neither for naming in Japanese nor for naming in English. An inspection of the means (see Table 4) shows a possible trend of interference from the Hiragana trials when naming in Japanese (close: 20 ms interference; distant: 24 ms interference). While there were no significant effects found in Experiment 4, there is nonetheless a trend mirroring the same general pattern as found in the previous experiments, with the most robust or noticeable effect from Hiragana distractors.

Although it is entirely speculative, the absence of semantic interference can be attributed to multiple sets of response criteria that participants had to consider. In addition to the criteria related to ease of production and semantic distance (as in the previous experiment), the participants in the current experiment had to consider an extra criterion of language. These multiple layers of response criteria could have been too demanding for the participants, which may have eliminated any noticeable effects for this experiment. Note that the participants in this bilingual experiment were all living in Japan (their native language country) whereas the participants in Experiment 2 all lived in England (their second language country) and were thus immersed in their second language and had higher levels of exposure to their L2. This may have altered the general level of resting activation of English for the participants in Exp. 2 in comparison to the participants in Exp. 4. Consequently, this may have allowed orthographic effects to be more clearly observable, especially for the identity facilitation effect, which is already a more prominent effects than the more intricate semantic interference effect. However, given the non-significant results from this experiment, this explanation is speculative, and so this remains an empirical question for now.

Experiment 5: Pseudohomophones in English monolinguals

Experiment 5 aimed to increase our understanding of the nature of Romaji and its processing mechanisms, by attempting to simulate the effects observed for Romaji (in Experiment 1 and 2) in a group of English monolinguals. This experiment investigated the identity facilitation effect by using pseudohomophonic (nonwords which, when read aloud, are phonologically identical to real words) distractors which look orthographically different from the real words but map onto the same phonological form (e.g., SHEEP + *sheap*, GOAT + *gote*). There were two conditions (equivalent of scripts in the previous experiments), depending on whether the distractor words in pictures were real words or pseudohomophones. The printed words matched the names of the pictures in the identity condition and not in the unrelated condition. Experiment 1 and 2 showed that Romaji distractors had the smallest facilitation among other Japanese scripts, suggesting that Romaji has weaker access to semantics than the other scripts, possibly due to its orthographically unique characteristics. We, therefore, argue that the pseudohomophones for English monolinguals and Romaji scripts for Japanese share certain similarities. We acknowledge that this experiment by no means is a perfect simulation of the Japanese case, as Romaji is a real orthography that Japanese speakers are taught and use in their daily lives, while pseudohomophones are not strictly speaking real words. However, it is meaningful to understand whether the processes involved in reading Romaji is completely unique to Japanese or if similar processing mechanisms exist for other languages, in this case English. Experiment 1 and 2 found that picture naming times for the Romaji identity condition was slower than for Hiragana and Kanji. The naming times for the unrelated Romaji condition, however, were slightly faster than for Hiragana, but were nearly identical to that for Kanji. The present experiment, then, tests whether a similar pattern is observed between the real word and the pseudohomophone conditions in English, by testing English monolinguals. The results from this experiment will inform our understanding of the generalisability of these types of orthographic effects in the picture-word task. Specifically, it will help us understand whether these mechanisms are limited to multiscript languages such as Japanese (and specifically, Romaji), or whether the same mechanisms can be found in other languages. If the results from Experiment 5 finds a similar pattern of result from pseudohomophones in English as was found for Romaji distractors in Experiments 1 and 2, this would suggest that these orthographic effects may be global mechanisms underlying not only Japanese, but all languages.

Method

Participants

Twenty-four native English monolingual speakers (20 female and four male) participated in this experiment. Their mean age was 19.6 years ($SD=2.4$ years), ranging from 18 to 27 years. All participants were university students, had normal or

Table 5 Results from English monolinguals in Experiment 5. Mean object naming latencies (in milliseconds), standard deviations, and rates of excluded responses (%E) by distractor script and distractor type. Also shown are the identity facilitation effects for each script.

	Identity			Unrelated			Facilitation
	M	SD	%E	M	SD	%E	
Real word	707	99	1.7	830	92	4.2	123**
Pseudohomophone	746	99	4.8	827	109	5.5	81**
Combined mean	727			829			

** $p < .001$

The RTs for the identity and unrelated conditions were compared using a two-way ANOVA with the factors of lexicality (2 levels) and distractor type (2 levels). The p-values were derived from a test of simple main effect examining the impact of distractor type for each lexicality condition

corrected-to-normal vision and none of the participants reported having any reading difficulties.

Stimulus materials

Twenty-two pictures were chosen based on having high name agreements (mean = 96%, Barry et al., 1997). Further, a small pilot study was conducted asking native English speakers to read a list of pseudohomophones out loud, and the stimulus pseudohomophones that were most often read like their corresponding real words were selected for this experiment.

Each target picture was then paired with four different types of distractor: identity (e.g., TRAIN + train); a matched control word for this name (e.g., TRAIN + floor); a pseudohomophone of the object name (e.g., TRAIN + trane), and an unrelated pseudohomophone control (e.g., TRAIN + flore). The real words and their controls were matched on word frequency counts provided by Brysbaert and New (2009; $M = 29.7$ and 30.3 per million respectively, $t = 0.91$, $p > .05$) word length ($M = 4.36$ and 4.41 respectively, $t = 1$, $p > .05$), number of syllables ($M = 1.09$ and 1.09 respectively, $t = 0$, $p > .05$), and grammatical class. The controls for the critical pseudohomophone distractors were the pseudohomophones of the control words (e.g., flore for floor). Each target picture was presented 4 times, making up a total of 88 experimental trials (See Appendix 4 for a complete list of all the distractors). Two different pictures were used for four practice trials. All distractors were written in Verdana size 36.

Procedure

The general procedure was the same as in previous studies. The participants were given four practice trials, followed by 88 experimental trials presented in a randomized order. Participants were given a short break halfway through the trials.

Results and discussion

The data were treated in the same way as in the previous experiments. Descriptive statistics are shown in Table 5.

The naming latencies were analysed in the same way as in the previous experiments, using linear mixed effects (LME) models fit by REML, with the same specifications. The fixed effects of script in this experiment were identity and pseudohomophone.

The analysis revealed a main effect of lexicality ($\beta = 505.24$, $SE = 40.66$, $t(131) = 12.43$, $p < .001$); naming times for the real word distractors were faster than the pseudohomophone distractors (769 vs. 786 ms). There was also a significant main effect of relatedness ($\beta = 78.78$, $SE = 24.16$, $t(47) = 3.26$, $p < .01$); naming times were faster for identity than unrelated distractors (727 vs. 829 ms). These effects, however, were modified by the significant interaction between relatedness and script ($\beta = -39.72$, $SE = 15.28$, $t(58) = 2.60$, $p < .05$). An analysis of the simple main effects showed that naming times were faster when the identity distractor was a real word (TRAIN + train, 707 ms) than when it was the corresponding pseudohomophone of the target picture (TRAIN + trane, 746 ms), $F(1, 23) = 7.19$, $p = .013$; but there was no significant difference between the two control conditions (TRAIN + floor, 830 ms; and TRAIN + flore, 827 ms), $F < 1$.

The impact of lexicality on the magnitude of facilitation (measured as RT difference between the identity and unrelated condition for each script) was examined by a one-way repeated-measures ANOVA as in Experiment 1. The effect of lexicality was significant, $F(3, 23) = 7.00$, $p < .001$, $\eta_p^2 = 0.23$, indicating stronger facilitation for the real words.

Although the pseudohomophones do not match the exact orthographic format of the semantic representations of the target word, they successfully induced a facilitation effect. The magnitude of the effect was, however, greatly reduced compared to the real word condition, and this pattern mimics the relationship between Romaji and other Japanese scripts. The current experiment implies that the semantic representations of Japanese words do not match the Romaji form as well as the other scripts do. Similarly, the pseudohomophones do not exactly match the semantic representations of real words. More research is necessary to investigate how words, which can be written in multiple scripts, are represented semantically, but one possible explanation for the slower processing of Romaji is script familiarity effect, which is described in more detail in the General Discussion.

General discussion

The picture-word task has proved to be a very useful means of investigating the processes responsible for lexical retrieval for speech production. Overall, the results from the present picture naming experiments robustly demonstrated that

Table 6 Summary over orthographic effects (in milliseconds) across all experiments (facilitation effects for exp. 1, 2 and 5; interference effects for exp. 3 and 4). Statistically significant effects in bold. (PH = Pseudohomophones.)

	Naming in L1				Naming in L2				
	Kanji	Hiragana	Katakana	Romaji	English	Kanji	Hiragana	Romaji	English
Exp. 1: Identity facilitation monolinguals	98	120	111	57					
Exp. 2: Identity facilitation bilinguals	103	174		77	5	3	32	-3	83
	Kanji	Hiragana	Picture	Romaji	English	Kanji	Hiragana	Romaji	English
Exp. 3: Semantic interference monolinguals									
Close	1	30	9	2					
Distant	12	31	12	-5					
Exp. 4: Semantic interference bilinguals									
Close	13	24		1	12	5	2	-14	5
Distant	7	20		-5	5	5	15	-10	-3
	Real	PH							
Exp. 5: Pseudohomophones English	123	81							

each Japanese script is processed somewhat differently (see Table 6 for an overview of the orthographic effects observed in each experiment). In line with previous research using the picture-word paradigm (see Hall, 2011, for a review), we generally observed both identity facilitation and semantic interference effects, albeit to various extent for the different Japanese scripts. The Hiragana script showed the strongest effects throughout the experiments, emphasizing the importance of phonological processing of written words in the picture naming task. Importantly, however, many theoretical accounts of the picture-word task have failed to incorporate reading processes, and most have focused on semantic activation. However, reading processes must clearly be engaged if the distractor words are to have any effect on the naming latencies of the target pictures. In fact, the ease of reading Hiragana (or the difficulty of reading Kanji) explains a great deal of the results found in the present study. Dylman and Barry's model (2018) incorporates the reading routes as proposed by the "three-route" model of reading (e.g., Coltheart et al., 2001), in explaining effects found using the picture-word task.

Experiments 1 to 3 consistently found that Hiragana produced the strongest facilitation and interference effects. Especially Experiment 1 and 2 highlighted that the difference in the magnitude of facilitation between Hiragana and Kanji trials are derived from slower naming time for Hiragana in the unrelated condition (no observable difference between the two scripts in the identity condition). This indicates that there is some degree of difficulty in excluding easy-to-pronounce Hiragana distractors from the articulation buffer when they are not to be produced. On the contrary, Kanji words seem to have been excluded from the articulation buffer relatively swiftly when they were not to be produced (and were in that sense treated similarly to distractor pictures). The overall pattern of these three experiments is, therefore, explained better by the account of response exclusion than by competition (or absence of competition) between semantic nodes. However, we lack sufficient statistical evidence to support the response exclusion hypothesis strongly, so what we can say at this stage is that the performance difference for Hiragana and Kanji must be derived from the difference in the "ease of reading" between the two scripts. The differences found concerning the presence or magnitude of effects observed for the various Japanese orthographies across the experiments support the conclusions drawn by Dylman and Kikutani (2018) in that the Japanese scripts involve different processing mechanisms.

For bilinguals, we found within-language lexical facilitation from identity distractors for all Japanese scripts, as well as English when naming in English (Experiment 2). Importantly, cross-language facilitation was also found and the pattern of results was generally consistent with Dylman and Barry's (2018) findings of asymmetric connections (the L1 to L2 connections being stronger than the L2 to L1 connection). When naming in L1 Japanese, English translations produced no reliable effect; but when naming in L2 English, Japanese translations facilitated naming times. This supports Dylman and Barry's theoretical interpretation concerning the asymmetric flow of facilitation between interconnected lexical representations of translations. This inter-lexical translation connection hypothesis proposes that the lexical representations of translations (and indeed for all alternative names, such as synonyms

and dialect-equivalents) are interconnected asymmetrically, such that more facilitatory activation is passed along the L1-to-L2 connection than along the weaker L2-to-L1 connection.

We discuss why only distractor words in Hiragana produced facilitation (in the Japanese-English bilinguals) by using Dylman and Barry's model. In this model, a written word activates semantics and phonological lexical representations. For bilinguals, the lexical store contains representations of different languages together, at which above mentioned inter-lexical translation connection occurs, but representations of semantic concepts at the semantic store are language independent. Structurally, the semantic store is not directly connected to phoneme retrieval, and connect only via the phonological lexical store. In order for cross-language facilitation to occur, the relevant representations in one language in the lexical store need to be activated strongly enough to send sufficient activation to the representations in the other language. Hiragana, being a phonological script, is able to send strong and direct activation for phonological lexical representations. Kanji, however, is read primarily by a visual-semantic process in a forward fashion from visual recognition to semantics, and so does not send strong activation to phonological lexical representations directly, and instead relies mostly on indirect activation from the semantic store. This gives Hiragana a higher probability of cross-language inter-lexical activations at the lexical store compared to Kanji. Furthermore, the semantic activation through Hiragana is not necessarily that much weaker than through Kanji. Although Hiragana was long thought to be processed more strictly phonologically, it was recently found that Hiragana reading relies more heavily on the lexical route than initially suspected, a characteristic that is similar to Kanji reading (Dylman & Kikutani, 2018). Thus, Hiragana may have strong enough semantic activation as well as having the benefit of more direct activation of phonology, which together will ensure rapid processing resulting in L1-to-L2 activation. Another possibility is that the phonological activation of Kanji happens much more slowly than Hiragana, due to a greater ambiguity of the phonological activation of Kanji words (ON- vs. KUN-readings).⁴

One consideration that needs to be made regards script familiarity, which refers to the notion that words written in their "typically used script" are processed faster than words transcribed in a different Japanese script. For example, foreign loanwords written in Katakana (the script which is typically used precisely for foreign loanwords) are processed faster than when the same words are transcribed in Hiragana (Tamaoka, 2014). Likewise, there may be an effect of script frequency in the current study, especially with regard to Romaji, which can be used to phonologically spell out any Japanese word, but is not the typically used script for Japanese words.

⁴ Similar to the argument for the higher semantic ambiguity (relatively speaking) of Kana processing, the argument here is that Kanji reading must involve a process whereby the reader (generally speaking) needs to make a decision as to the correct pronunciation in any given context. Now, this decision is governed by learnt rules regarding context and compound status (i.e., whether any given Kanji character is combined with other, Kanji or Hiragana, characters). Nonetheless, the decision must be made before the character/word can be read aloud using correct pronunciation. Kana reading, in contrast, does not involve this particular extra step as each character only maps onto a single phonological node.

The within-language facilitation from the identity distractors typed in Romaji may therefore be at a purely phonological level, via the phonological assembly route. However, this is an unlikely explanation for the difference in facilitation between Hiragana and Kanji, as care was taken to ensure that the stimuli all had common and frequently used Kanji forms. Although the present research did not determine the most common form for each stimulus word, Kanji form is most likely to be the preferred form for the majority of them and they appear frequently in reading materials for adults. In fact, Dylman and Kikutani (2018) specifically controlled for script familiarity and found a similar effect where naming latencies were significantly faster for words written in Hiragana compared to Kanji, even when the preferred form of the word was Kanji and the Kanji form was highly frequent.

One important note to make regards the differences between the samples across the five experiments. While the participants from all experiments were university students, and had fairly similar age ranges, there were some differences to note. For example, as mentioned in the discussion section for Experiment 4, the bilinguals who participated in Experiment 2 were all living in the UK during data collection, whereas the bilingual participants in Experiment 4 lived in Japan. The two groups, therefore, differed in terms of immersion, and, thereby, likely second language exposure, use and proficiency. While added heterogeneity can lead to increased generalisability, it is, nonetheless, an important aspect to note, especially for future studies. Another issue concerns the definition of the term bilingualism itself, which is a complex and multifaceted term. Bilinguals are a highly heterogeneous group (e.g., de Bruin, 2019) with individually and uniquely created language compositions. Indeed, recent research suggests that the mere operationalisation of bilingualism as a measurable variable can lead to different outcomes even when analysing the same data set (Champoux-Larsson & Dylman, 2020).

Most experiments using the picture-word task with visual distractor words have tested participants whose language is written in an alphabetical script. Our study has been concerned with possible differences between scripts of distractor words in the facilitation effect from within-language identical words, and the facilitation effect from translations in bilinguals. The reported experiments highlight the role of orthography in the picture-word task, which can be more directly investigated using the multiscriptal language Japanese, thereby eliminating potential confounds such as linguistic differences. It is evident from these findings that orthography, independent of phonology plays a significant role in effects found within the picture-word paradigm. Further research is necessary to determine whether the existing models of reading and object naming, which are established based on data using alphabetical languages, can be applied to other languages using different orthographic systems. Especially languages like Japanese, using multiple orthographies, are likely to have some unique characteristics involved in processing. Models for bilinguals may also need to incorporate the aspect of orthography more clearly. The effect of orthography, at the same time, can also work to assess the credibility of existing models as the present research attempted to compare the lexical selection by competition account and the response exclusion account, not only for multiscriptal languages such as Japanese, but as general models of speech production. Crucially, the current study has examined fundamental issues concerning the role of orthography in the

picture-word task whilst removing methodological constraints involving linguistic differences or specific bilingual populations, which has typically limited previous studies investigating these issues.

Appendix 1

See Table 7.

Table 7 The stimuli used in Experiment 1

Target picture	Kanji	Kanji unrel	Hiragana	Hiragana unrel	Katakana	Katakana unrel	Romaji	Romaji unrel
Ashtray	灰皿	白鳥	はいざら	はくちょう	ハイザラ	ハクチョウ	Haizara	Hakucho
Balloon	風船	牛	ふうせん	うし	フウセン	ウシ	Fusen	Ushi
Bear	熊	眼鏡	くま	めがね	クマ	メガネ	Kuma	Megane
Bird	鳥	車	とり	くるま	トリ	クルマ	Tori	Kuruma
Book	本	馬	ほん	うま	ホン	ウマ	Hon	Uma
Box	箱	本	はこ	ほん	ハコ	ホン	Hako	Hon
Car	車	雲	くるま	くも	クルマ	クモ	Kuruma	Kumo
Cat	猫	手	ねこ	て	ネコ	テ	Neko	Te
Clock	時計	猫	とけい	ねこ	トケイ	ネコ	Tokei	Neko
Cow	牛	風船	うし	ふうせん	ウシ	フウセン	Ushi	Fuusen
Desk	机	電球	つくえ	でんきゅう	ツクエ	デンキユウ	Tsukue	Denkyu
Dog	犬	傘	いぬ	かさ	イヌ	カサ	Inu	Kasa
Ear	耳	星	みみ	ほし	ミミ	ホシ	Mimi	Hoshi
Eye	目	旗	め	はた	メ	ハタ	Me	Hata
Flag	旗	魚	はた	さかな	ハタ	サカナ	Hata	Sakana
Flower	花	信号	はな	しんごう	ハナ	シンゴウ	Hana	Shingo
Foot	足	木	あし	き	アシ	キ	Ashi	Ki
Glasses	眼鏡	熊	めがね	くま	メガネ	クマ	Megane	Kuma
Hand	手	椅子	て	いす	テ	イス	Te	Isu
Horse	馬	箱	うま	はこ	ウマ	ハコ	Uma	Hako
Key	鍵	豚	かぎ	ぶた	カギ	ブタ	Kagi	Buta
Light bulb	電球	犬	でんき ゆう	いぬ	デンキ ユウ	イヌ	Denkyu	Inu
Mountain	山	電話	やま	でんわ	ヤマ	デンワ	Yama	Denwa
Pig	豚	鍵	ぶた	かぎ	ブタ	カギ	Buta	Kagi
Star	星	目	ほし	め	ホシ	メ	Hoshi	Me
Swan	白鳥	灰皿	はくちょう	はいざら	ハクチョ ウ	ハイザラ	Hakucho	Haizara
Telephone	電話	山	でんわ	やま	デンワ	ヤマ	Denwa	Yama
Traffic lights	信号	花	しんごう	はな	シンゴウ	ハナ	Shingo	Hana
Tree	木	足	き	あし	キ	アシ	Ki	Ashi
Umbrella	傘	耳	かさ	みみ	カサ	ミミ	Kasa	Mimi

Appendix 2

See Table 8.

Table 8 The stimuli used in Experiment 2

Target picture	Kanji	Kanji unrel	Hiragana	Hiragana unrel	Romaji	Romaji unrel	English	English unrel
Apple	林 鴨 檣*		りんご	かも	Ringo	Kamo	Apple	Duck
Book	本	蟹	ほん	かに	Hon	Kani	Book	Crab
Chair	椅子	鼠	いす	ねずみ	Isu	Nezumi	Chair	Mouse
Clock	時計	豚	とけい	ぶた	Tokei	Buta	Clock	Pig
Cloud	雲	熊	くも	くま	Kumo	Kuma	Cloud	Bear
Cow	牛	山羊	うし	やぎ	Ushi	Yagi	Cow	Goat
Crocodile	鱷*	豆	わに	まめ	Wani	Mame	Crocodile	Peas
Deer	鹿	葡萄	しか	ぶどう	Shika	Budou	Deer	Grapes
Dog	犬	封筒	いぬ	ふうとう	Inu	Fuutou	Dog	Envelope
Eagle	鷲*	玉葱	わし	たまねぎ	Washi	Tamanegi	Eagle	Onion
Fish	魚	南瓜	さかな	かぼちゃ	Sakana	Kabocha	Fish	Pumpkin
Flower	花	人参	はな	にんじん	Hana	Ninjin	Flower	Carrot
Hat	帽子	虎	ぼうし	とら	Boushi	Tora	Hat	Tiger
Key	鍵	象	かぎ	ぞう	Kagi	Zou	Key	Elephant
Mountain	山	馬	やま	うま	Yama	Uma	Mountain	Horse
Mouse	鼠*	猫	ねずみ	ねこ	Nezumi	Neko	Mouse	Cat
Peach	桃	蛇	もも	へび	Momo	Hebi	Peach	Snake
Shoe	靴	兎	くつ	うさぎ	Kutsu	Usagi	Shoe	Rabbit
Train	電車	狐	でんしゃ	きつね	Densha	Kitsune	Train	Fox
Turtle	亀	狼	かめ	おおかみ	Kame	Ookami	Turtle	Wolf
Umbrella	傘	耳	かさ	みみ	Kasa	Mimi	Umbrella	Ear
Water-melon	西瓜*	莓	すいか	いちご	Suika	Ichigo	Water-melon	Strawberry
Whale	鯨*	鰻	くじら	うなぎ	Kujira	Unagi	Whale	Eel
Wolf	狼	旗	おおかみ	はた	Ookami	Hata	Wolf	Flag

A total of 6 characters marked with asterisks were not listed in Joyo list (second characters for Apple and Watermelon)

Appendix 3

See Table 9.

Table 9 The stimuli used in Experiments 3 and 4

Target object	Close				Distant				Unrelated			
	Close		Distant		Distant		Unrelated		Unrelated		Unrelated	
	Kanji	Hiragana	Romaji	English	Kanji	Hiragana	Romaji	English	Kanji	Hiragana	Romaji	English
Airplane	気球	ききゅう	Kikyuu	Air balloon	船	ふね	Fune	Boat	帽子	ぼうし	Boushi	Hat
Apple	梨	なし	Nashi	Pear	豆	まめ	Mame	Peas	時計	とけい	Tokei	Clock
Axe	金槌	かなづち	Kanazuchi	Hammer	縄	なわ	Nawa	Rope	耳	みみ	Mimi	Ear
Barrel	桶	おけ	Oke	Bucket	箱	はこ	Hako	Box	雲	くも	Kumo	Cloud
Cat	虎	とら	Tora	Tiger	蟹	かに	Kani	Crab	目	め	Me	Eye
Chair	腰掛	こしかけ	Koshikake	Stool	戸棚	とだな	Todana	Cupboard	封筒	ふうとう	Futou	Envelope
Chicken	鳩	はと	Hato	Pigeon	蝶	ちよう	Chou	Butterfly	鍵	かぎ	Kagi	Key
Crocodile	蛙	かえる	Kaeru	Frog	鼠	ねずみ	Nezumi	Mouse	本	ほん	Hon	Book
Dog	狼	おおかみ	Ookami	Wolf	雀	すずめ	Suzume	Sparrow	旗	はた	Hata	Flag
Donkey	馬	うま	Uma	Horse	鴨	かも	Kamo	Duck	太陽	たいよう	Taiyou	Sun
Fish	鰻	うなぎ	Unagi	Eel	鹿	しか	Shika	Deer	山	やま	Yama	Mountain
Goat	牛	うし	Ushi	Cow	鯨	くじら	Kujira	Whale	花	はな	Hana	Flower
Jar	壺	つぼ	Tsubo	Vase	籠	かご	Kago	Basket	電車	でんしゃ	Densha	Train
Onion	人参	にんじん	Ninjin	Carrot	西瓜	すいか	Suika	Watermelon	靴	くつ	Kutsu	Shoe
Pig	芋	ひつじ	Hitsuji	Sheep	猿	さる	Saru	Monkey	電話	でんわ	Denwa	Telephone
Pumpkin	芋	いも	Imo	Potato	葡萄	ぶどう	Budou	Grapes	羽根	はね	Hane	Feather
Rabbit	狸	たぬき	Tanuki	Raccoon	象	ぞう	Zou	Elephant	傘	かさ	Kasa	Umbrella
Strawberry	桃	もも	Momo	Peach	葱	ねぎ	Negi	Leek	車	くるま	Kuruma	Car
Swan	鶯	わし	Washi	Eagle	熊	くま	Kuma	Bear	井戸	いど	Ido	Well
Turtle	蛇	へび	Hebi	Snake	狐	きつね	Kitsune	Fox	木	き	Ki	Tree

Appendix 4

See Table 10.

Table 10 The stimuli used in Experiment 5

Target Object	Identity	Identity Unrelated	Pseudo-Homophone	Pseudo-Homophone Unrelated
Bowl	Bowl	Horn	Bole	Hawn
Cake	Cake	Bird	Caik	Burd
Cloud	Cloud	Sheet	Clowd	Sheat
Coat	Coat	Page	Cote	Paige
Eagle	Eagle	Bacon	Eegle	Beykon
Ear	Ear	Oil	Ere	Oyl
Flute	Flute	Spade	Floot	Spaid
Goat	Goat	Root	Gote	Rute
Key	Key	Ice	Kee	Ise
Leaf	Leaf	Sole	Leef	Soal
Nail	Nail	Wrap	Nale	Rapp
Needle	Needle	Pillow	Neadel	Pilloe
Sheep	Sheep	Tooth	Sheap	Tufe
Shirt	Shirt	Knife	Shert	Nife
Shoe	Shoe	Gate	Shoo	Gaite
Snail	Snail	Whirl	Snale	Whurl
Snake	Snake	Sweat	Snaik	Swett
Spoon	Spoon	Float	Spune	Flote
Tie	Tie	Cup	Tye	Kup
Train	Train	Floor	Trane	Flore
Tree	Tree	Tape	Trea	Teip
Wheel	Wheel	Coast	Wheal	Coste

Funding Open access funding provided by Stockholm University.

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

References

- Barry, C., Morrison, C.M., & Ellis, A.W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology*, 50, 560–585. <https://doi.org/10.1080/783663595>.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41, 977–990. <https://doi.org/10.3758/BRM.41.4.977>
- Champoux-Larsson, M.-F., & Dylman, A. S. (2020). Different measurements of bilingualism and their effect on performance on a Simon task. *Applied Psycholinguistics*, 42, 505–526. <https://doi.org/10.1017/S0142716420000661>
- 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. <https://doi.org/10.1006/jmla.2001.2825>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256. <https://doi.org/10.1037/0033-295X.108.1.204>
- Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Bilingualism: Language and Cognition*, 2, 231–244. <https://doi.org/10.1017/S1366728999000334>
- Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365–397. <https://doi.org/10.1006/jmla.1999.2651>
- Damian, M.F., & Bowers, J.S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review*, 10, 111–117. <https://doi.org/10.3758/BF03196474>
- de Bruin, A. (2019). Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences*, 9, 33–33. <https://doi.org/10.3390/bs9030033>
- Dhooge, E., & Hartsuiker, R. J. (2010). The distractor frequency effect in picture-word interference: Evidence for response exclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 878–891. <https://doi.org/10.1037/a0019128>
- Dylman, A. S., & Barry, C. (2018). When having two names facilitates lexical selection: Similar results in the picture-word task from translations in bilinguals and synonyms in monolinguals. *Cognition*, 171, 151–171. <https://doi.org/10.1016/j.cognition.2017.09.014>
- Dylman, A. S., & Kikutani, M. (2018). The role of semantic processing in reading Japanese orthographies: An investigation using a script-switch paradigm. *Reading and Writing: An Interdisciplinary Journal*, 31, 503–531. <https://doi.org/10.1007/s11145-017-9796-3>
- Feldman, L. B., & Turvey, M. T. (1980). Words written in Kana are named faster than the same words written in Kanji. *Language and Speech*, 23, 141–147. <https://doi.org/10.1177/002383098002300201>
- Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers? What's the (hard) problem? *Bilingualism: Language and Cognition*, 9, 153–166.
- Flaherty, M. (1993). Are Japanese kanji processed like pictures? *Psychologia: An International Journal of Psychology in the Orient*, 36, 144–150.
- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118, 13–42. <https://doi.org/10.1037/0096-3445.118.1.13>
- Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81. <https://doi.org/10.1017/S1366728998000133>
- Hall, M. L. (2011). Bilingual picture-word studies constrain theories of lexical selection. *Frontiers in Psychology*, 2, 381. <https://doi.org/10.3389/fpsyg.2011.00381>
- Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1, 213–230. <https://doi.org/10.1017/S1366728998000364>

- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, 5, 67–103. <https://doi.org/10.1080/02643298808252927>
- Janssen, N., Schirm, W., Mahon, B. Z., & Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 249–256. <https://doi.org/10.1037/0278-7393.34.1.249>
- Klein, G.S. (1964). Semantic power measured through the interference of words with color-naming. *The American Journal of Psychology*, 77, 576–588. <https://doi.org/10.2307/1420768>.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.
- 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. <https://doi.org/10.1017/S0140525X99001776>
- Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory and Cognition*, 7, 485–495. <https://doi.org/10.3758/BF03198265>
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 503–535. <https://doi.org/10.1037/0278-7393.33.3.503>
- Marsh, J. E., Crawford, J. C., Pilgrim, L. K., Sörqvist, P., & Hughes, R. W. (2017). Trouble articulating the right words: Evidence for a response-exclusion account of distraction during semantic fluency. *Scandinavian Journal of Psychology*, 58, 367–372. <https://doi.org/10.1111/sjop.12386>
- Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 1146–1160. <https://doi.org/10.1037/0278-7393.17.6.1146>
- Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 555–563. <https://doi.org/10.1037/0278-7393.28.3.555>.
- Morton, J., & Sasanuma, S. (1984). Lexical access in Japanese. In L. Hendersson (Ed.), *Orthographies and reading: Perspectives from Cognitive Psychology, Neuropsychology and Linguistics* (pp. 25–56). Erlbaum.
- Park, S., & Arbuckle, T.Y. (1977). Ideograms versus alphabets: Effects of script on memory in "bisciprual" Koreansubjects. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 631–642. <https://doi.org/10.1037/0278-7393.3.6.631>
- R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org>.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114, 510–532. <https://doi.org/10.1037/0033-2909.114.3.510>
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249–284. [https://doi.org/10.1016/S0010-0277\(97\)00027-9](https://doi.org/10.1016/S0010-0277(97)00027-9)
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110, 88–125. <https://doi.org/10.1037/0033-295X.110.1.88>
- Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin and Review*, 13, 33–37. <https://doi.org/10.3758/BF03193809>
- Rusted, J. (1988). Orthographic effects for Chinese-English bilinguals in a picture-word interference task. *Current Psychology: Research and Reviews*, 7, 207–220.
- Sasanuma, S. (1975). Kana and Kanji processing in Japanese aphasics. *Brain and Language*, 2, 360–383.
- Sasanuma, S., & Fujimura, O. (1971). Selective impairment of phonetic and nonphonetic transcriptions of words in Japanese aphasic patients: Kana versus Kanji on visual recognition and writing. *Cortex*, 1, 1–18. [https://doi.org/10.1016/S0010-9452\(71\)80019-9](https://doi.org/10.1016/S0010-9452(71)80019-9)
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86–102. [https://doi.org/10.1016/0749-596X\(90\)90011-N](https://doi.org/10.1016/0749-596X(90)90011-N)
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215. <https://doi.org/10.1037/0278-7393.6.2.174>

- Spalek, K., Damian, M., & Bölte, J. (2013). Is lexical selection in spoken word production competitive? Introduction to the Special Issue on lexical competition in language production. *Language & Cognitive Processes* 28, 597–614 <https://doi.org/10.1080/01690965.2012.718088>
- Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 896–918. <https://doi.org/10.1037/0278-7393.22.4.896>
- Tamaoka, K. (2014). The Japanese writing system and lexical understanding. *Japanese Language and Literature (the American Association of Teachers of Japanese, AATJ)*, 48, 431–471.
- Verdonschot, R. G., Kiyama, S., Tamaoka, K., Kinoshita, S., La Heij, W., & Schiller, N. O. (2011a). The functional unit of Japanese word naming: Evidence from masked priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1458–1473. <https://doi.org/10.1037/a0024491>
- Verdonschot, R. G., La Heij, W., Paolieri, D., Zhang, Q., & Schiller, N. O. (2011b). Homophonic context effects when naming Japanese kanji: Evidence for processing costs? *The Quarterly Journal of Experimental Psychology*, 64, 1836–1849. <https://doi.org/10.1080/17470218.2011.585241>
- 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. *The Quarterly Journal of Experimental Psychology*, 66, 2023–2038. <https://doi.org/10.1080/17470218.2013.773050>
- Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, 85, B61–B69. [https://doi.org/10.1016/S0010-0277\(02\)00107-5](https://doi.org/10.1016/S0010-0277(02)00107-5)
- Wydell, T. N., & Kondo, T. (2015). Behavioral and neuroimaging research of reading: A case of Japanese. *Current Developmental Disorders Reports*, 2, 339–345. <https://doi.org/10.1007/s40474-015-0066-2>
- Zhang, Q., & Weekes, B. S. (2009). Orthographic facilitation effects on spoken word production: Evidence from Chinese. *Language and Cognitive Processes*, 24, 1082–1092. <https://doi.org/10.1080/01690960802042133>
- Zhu, X., Damian, M. F., & Zhang, Q. (2015). Seriality of semantic and phonological processes during overt speech in Mandarin as revealed by event-related brain potentials. *Brain and Language*, 144, 16–25. <https://doi.org/10.1016/j.bandl.2015.03.007>
- Zhu, X., Zhang, Q., & Damian, M. F. (2016). Additivity of semantic and phonological effects: Evidence from speech production in Mandarin. *The Quarterly Journal of Experimental Psychology*, 69, 2285–2304. <https://doi.org/10.1080/17470218.2015.112942>