Canadian Journal of Experimental Psychology
2010, Vol. 64, No. 3, 180–196

1196-1961/10/$12.00 DOI: 10.1037/a0020475

The Relatedness-of-Meaning Effect for Ambiguous Words in Lexical-Decision Tasks: When Does Relatedness Matter?

Yasushi Hino and Yuu Kusunose
Waseda University

Stephen J. Lupker
The University of Western Ontario

Effects of the number of meanings (NOM) and the relatedness of those meanings (ROM) were examined for Japanese Katakana words using a lexical-decision task. In Experiment 1, only a NOM advantage was observed. In Experiment 2, the same Katakana words produced a ROM advantage when Kanji words and nonwords were added. Because the Kanji nonwords consisted of unrelated characters whereas the Kanji words consisted of related characters, participants may have used the relatedness of activated meanings as a cue in making lexical decisions in this experiment, artificially creating a ROM advantage for Katakana words. Consistent with this explanation, no ROM effect for Katakana words was observed in Experiment 3 when the Kanji nonwords consisted of characters with similar (i.e., related) meanings. These results pose a further challenge to the position that the speed of semantic coding is modulated by ROM for ambiguous words.

Keywords: a number-of-meaning effect, a relatedness-of-meaning effect, a lexical decision task

A key question in reading research is how meanings are retrieved from the printed forms. While this process is a complicated one even for unambiguous (i.e., single meaning) words, it is substantially more complicated for words that have multiple meanings (e.g., BANK). One type of model that potentially seems equipped to handle the nuances of processing ambiguous words and that, in general, has had some success explaining the nature of the orthographic-semantic interaction is parallel distributed processing (PDP) models (e.g., Borowsky & Masson, 1996; Kawamoto, 1993; Kawamoto, Farrar, & Kello, 1994; Plaut, 1997; Plaut & McClelland, 1993; Plaut, McClelland, Seidenberg, & Patterson, 1996; Rodd, Gaskell, & Marslen-Wilson, 2004; Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990). The present research is an attempt to investigate some of the assumptions models of this sort have made about the meaning retrieval process for ambiguous words.

According to PDP models, orthographic, phonological, and semantic information are represented as patterns of activation over sets of units representing features of the respective domains. These units are interconnected with each other and the weights on these connections are adjusted through a learning process in order to produce correct associations between the input representations (i.e., orthography) and the output representations (i.e., pronunciations and meanings). In addition, the strength of association between these representations is determined by the consistency of the input-output relationships possessed by the various words and, hence, the speed of the settling process at the output level is determined by these associative strengths. As such, the models predict that the speeds of phonological coding and semantic coding should be modulated by the nature of orthographic-phonological and orthographic-semantic relationships possessed by the words being read.

Research on the processing of homographs provides a nice demonstration of exactly this point. Homographs are words with two pronunciations (e.g., WIND), and the typical finding is that these words take longer to name than frequency matched nonhomographs (e.g., Kawamoto & Zemblidge, 1992; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). Using a PDP model, Seidenberg and McClelland (1989) were able to successfully simulate this homograph disadvantage. When their model was trained to read a nonhomograph, a reasonably strong association was established between orthographic and phonological representations because a single orthographic pattern of activation was always mapped onto a single phonological pattern of activation. In contrast, when the model was trained to read a homograph, the associations established between the single orthographic pattern and each of the two phonological representations were weaker because the weights most appropriate for one of the orthographic-phonological associations would be altered during the process of learning the other association. Because of the weaker relationships between orthography and phonology for the homographs, the settling process in the model was slower for the homographs than for the nonhomographs, providing an explanation of the homograph disadvantage in the latency data.

In the semantic realm, these models would postulate that there are more complicated (i.e., one-to-many) relationships between orthography and semantics for ambiguous words (i.e., words with multiple meanings) than for unambiguous words. Therefore, as Joordens and Besner (1994) have noted, this type of model should also predict a processing time disadvantage for ambiguous words (compared to unambiguous words) because the settling process at the semantic level should be more difficult for ambiguous words.

Yasushi Hino, Faculty of Arts and Sciences, Waseda University; Yuu Kusunose, Graduate School of Letters, Arts and Sciences, Waseda University; Stephen J. Lupker, Department of Psychology, The University of Western Ontario.

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

This prediction, however, has not been borne out in the data from lexical decision experiments, most of which appear to show a number of meanings (NOM; i.e., ambiguity) advantage (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino, Lupker, & Pexman, 2002; Hino, Lupker, Sears, & Ogawa, 1998; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989; Pexman & Lupker, 1999; Rubenstein, Garfield, & Millikan, 1970; Rubenstein, Lewis, & Rubenstein, 1971, although see Forster & Bednall, 1976, and Gernsbacher, 1984). As a result, these types of models have had to develop additional assumptions about the nature of representations for ambiguous words as well as the nature of processing in a lexical-decision task (e.g., Borowsky & Masson, 1996; Hino et al., 2002; Kawamoto, Farrar, & Kello, 1994; Klepousniotou & Baum, 2007; Piercey & Joordens, 2000; Rodd, Gaskell, & Marslen-Wilson, 2002, 2004).

One way to try to solve this problem has been to argue that lexical decision performance is not necessarily sensitive to the nature of semantic coding (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Kawamoto et al., 1994; Rueckl, 1995) and, instead, that lexical decisions are based mainly on the nature of orthographic processing (e.g., Balota & Chumbley, 1984; Besner & McCann, 1987). For example, Kawamoto et al. (1994) simulated the ambiguity advantage in lexical decision tasks using a fully recurrent PDP network with a least mean square learning algorithm. Working on the assumption that lexical decisions are based mainly on orthographic properties of stimuli, Kawamoto et al. used the number of cycles required for fully settling within the orthographic module in order to simulate lexical decision performance. In their simulations, when a single orthographic pattern was mapped onto multiple semantic patterns (i.e., the situation for ambiguous words), the model developed stronger associations amongst orthographic units. In essence, the weaker associations between orthographic and semantic units for ambiguous words were compensated for by developing stronger associations amongst orthographic units. As a result, the speed of settling within the orthographic module was faster for ambiguous words than for unambiguous words, successfully simulating the ambiguity advantage.

An alternative approach to the problem was provided by Borowsky and Masson (1996) who argued that the sum of energy within the orthographic and semantic modules, which reflects the network's activity toward a learnt pattern of activation (a basin of attraction), is the metric that drives performance in a lexical-decision task. According to Borowsky and Masson, the energy within the orthographic and semantic modules is assumed to reflect the familiarity of stimuli, so that the energy will decrease more quickly for familiar stimuli than for unfamiliar stimuli. As such, lexical decision latencies were simulated using the number of cycles required for the energy to reach a certain criterion in their Hopfield network model trained with a Hebbian learning algorithm. With this procedure, Borowsky and Masson successfully simulated an ambiguity advantage in lexical decision tasks.

Homonymous – Polysemous Distinction

One advantage of models that adopt the assumption of distributed representations at the semantic level is that these models have a straightforward way of representing a potentially important distinction between types of ambiguous words; polysemous words and homonyms. Polysemous words are words for which the various meanings of the word are related, while homonyms are words in which the meanings are unrelated. An example of a homonym is BANK, which has two completely unrelated meanings, "a financial institution" and "the edge of a river". Homonyms are assumed to arise because unrelated word meanings are arbitrarily attached to the same form. In contrast, an ambiguous word like "PAPER" is a typical example of a polysemous word because the meanings it has are all related to one another (e.g., "a writing material" and "the content written on that material"). The term "multiple senses" is often used, rather than "multiple meanings", when describing polysemous words. Polysemous words are assumed to arise by extending the use of a word into new contexts.

Although the homonymous-polysemous distinction is etymological, within a PDP framework, this distinction can be reflected quite naturally in terms of how these two types of words are represented in semantic memory, which, as a result, should give rise to processing differences when those words are read. In particular, Azuma and Van Orden (1997) and Rodd et al. (2002, 2004) proposed that the homonymous-polysemous distinction could be reflected within a PDP framework by assuming that semantic representations for related senses share a set of semantic features while the semantic representations for unrelated meanings do not. That is, whereas homonymous meanings would be assumed to be represented by separate sets of semantic units (features), polysemous senses would be assumed to share at least a subset of semantic units. As such, the orthographic-to-semantic mappings would be more consistent for polysemous words than for homonyms. As a result, PDP models would predict that the settling process at the semantic level should be slower for homonyms than for polysemous words.

Azuma and Van Orden (1997) did indeed report a result consistent with this prediction, specifically, an advantage for words with related meanings (i.e., a relatedness-of-meaning (ROM) advantage) in a lexical decision experiment when pseudohomophones were used as nonwords (although such was only the case for words having few meanings overall). More specifically, in their experiments, the number of meanings/senses and the relatedness amongst meanings/senses were orthogonally manipulated. When legal nonwords were used, no significant effects were observed. When pseudohomophones were used as nonwords, however, a ROM advantage was observed for the few-meaning words (i.e., ambiguous words with related meanings were responded to faster than word with unrelated meanings). In addition, there was a main effect of NOM (i.e., the basic ambiguity advantage).

In order to simulate NOM and ROM effects in lexical decision, Rodd et al. (2004) developed a PDP model based on the Hopfield network used by Joordens and Besner (1994) and Borowsky and Masson (1996). Instead of using a Hebbian learning algorithm, Rodd et al. employed the least mean square learning algorithm used by Kawamoto et al. (1994). In contrast to Kawamoto et al.'s model, however, Rodd et al. assumed that the settling time in the semantic module indexed lexical decision time.

In their model, each meaning/sense is represented by an attractor basin in a semantic module and the state of the semantic system moves into one of the attractor basins when a meaning/sense is computed for a word. For an unambiguous word, a single attractor basin is assumed, whereas multiple attractor basins are assumed for homonyms and polysemous words. In addition, because polysemous senses are related, their attractor basins are clustered together in the semantic system. In essence, the polysemous senses are assumed to essentially create a broad attractor basin containing the attractor basins of all the multiple senses. Thus, although it would take longer for the network to fully compute a meaning/sense for a polysemous word than for an unambiguous word, the network should find its way into the broad basin fairly quickly. Under the assumption that the lexical status of a stimulus could be evaluated once the state of the semantic system enters an attractor basin, a processing advantage for polysemous words over unambiguous words in lexical decision can be explained. In contrast, there should be no such advantage for homonyms because their multiple meanings are not similar and, hence, their representations do not create a broad attractor basin. In fact, when homonyms are processed, the inconsistent relationship between orthography and semantics should produce a processing cost in comparison to both polysemous and unambiguous words.

Consistent with these predictions, Rodd et al. (2002) reported a significant polysemy advantage (multiple-sense words were responded to faster than few-sense words) and a small trend toward a homonymy disadvantage (multiple-meaning words were responded to slightly slower than few-meaning words) in their (visual) lexical-decision task, essentially consistent with the predictions those authors made based on their PDP model. That is, a processing advantage was observed when words possessed more senses (when the number of meanings was kept constant); however, words having more meanings, but having the same number of senses, produced a small processing time disadvantage (See also Beretta, Fiorentino, & Poeppel, 2005, for the identical results using the same stimuli. In these experiments, the number of unrelated meanings and the number of related senses were manipulated based on the online *Wordsmyth English Dictionary – Thesaurus*: Parks, Ray, & Bland, 1998). As in Azuma and Van Orden (1997), obtaining these results in a visual lexical-decision task required using pseudohomophones as the nonwords.

Rodd et al. (2002) also analysed the stimuli used in three previous studies that reported ambiguity advantages in lexical decision tasks (Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Millis & Button, 1989) using the same dictionary (Parks et al., 1998). Their analysis suggested that, in the two of three studies, there was no difference in the number of meanings between many-meaning words (the "ambiguous" words) and few-meaning words (the "unambiguous" words). In contrast, the number of senses was larger for many-meaning words than for few-meaning words in all the three studies. As such, it appears that the ambiguity manipulation in these experiments may have involved manipulating senses rather than meanings. Based on these observations, Rodd et al. suggested that, consistent with their own findings, the ambiguity advantages reported in previous studies were due to the ambiguous words having multiple senses rather than having multiple meanings (although such is probably not the case in all of the prior literature, e.g., see Hino & Lupker, 1996). Hence, Rodd et al. argued that the ambiguity advantage observed previously may be reconcilable with their claims about the nature of semantic representations for homonyms and polysemous words.

Empirical Challenges to Rodd et al.'s (2002; 2004) Theoretical Position

Although Rodd et al.'s (2002) assumption about the nature of semantic representations for homonyms and polysemous words is consistent with their own data, it garnered little support in a series of studies by Klein and Murphy (2001, 2002) using a sensicality judgement task. In this task, participants are asked to decide whether a given phrase makes sense, responding by pressing one of two buttons. In their experiments, noun phrases were created by presenting polysemous words preceded by a modifier (e.g., SHREDDED PAPER). The same polysemous word was accompanied by different modifiers in different phrases, so that the polysemous word in pairs of phrases denoted either the same sense (e.g., WRAPPING PAPER – SHREDDED PAPER) or different senses (e.g., DAILY PAPER – SHREDDED PAPER). If a polysemous word (e.g., PAPER) is represented essentially as an abstract core meaning that encompasses all the senses (i.e., the broad attractor basin described above), similar sets of semantic features should be activated by the same-sense and different-sense words. Hence, the two prime types should produce equivalent latencies. In contrast, if polysemous senses are separately represented (as the multiple meanings for homonyms are presumed to be), the same-sense primes should activate many more features common to the target meaning, producing faster latencies. Klein and Murphy's results were consistent with the latter prediction. The sensicality judgement responses were faster for the same-sense pairs than for the different-sense pairs.

Also using their sensicality judgement task, Klein and Murphy (2001) attempted to examine another implication of these representational assumptions. Because homonymous meanings are assumed to be represented separately, same-meaning pairs (e.g., COMMERCIAL BANK – SAVINGS BANK) should lead to much faster latencies than different-meaning pairs (e.g., CREEK BANK – SAVINGS BANK). It is more important to note that if polysemous senses are represented by a core meaning, the size of the difference between same-sense pairs and different-sense pairs for polysemous words should be much smaller than the size of the difference between same-meaning and different-meaning pairs when using homonyms. Contrary to this prediction, Klein and Murphy's results showed that the differences for the two types of ambiguous words were identical.[1]

Based on the results of the sensicality judgement task, as well as some additional experiments, Klein and Murphy (2001, 2002) suggested that relatedness of meanings/senses is essentially independent of the semantic overlap. That is, although polysemous senses are related, most polysemous senses, like homonymous

[1] Klepousniotou, Titone, and Romero (2008) recently attempted to replicate Klein and Murphy's (2001) results. For ambiguous words with low-overlap and moderate-overlap among meanings, the results were similar to those of Klein and Murphy. For ambiguous words with more highly overlapping meanings, however, Klepousniotou et al. observed comparable response latencies for the same-meaning and different-meaning pairs when the target phrase denoted a dominant meaning of the ambiguous word, indicating the possibility that multiple meanings are represented by a single core representation for this type of ambiguous words. When a word's senses/meanings are highly overlapping, however, the word may actually be better thought of as being unambiguous.

meanings, are represented separately in a reader's semantic system. For example, although the two senses, "a writing material" and "a content written on that material" for the polysemous word PAPER are related (the latter being an extension of the former), semantic overlap between these two senses is actually fairly minimal because the writing material is dissimilar to the content written on that material. As a result, most related senses would only minimally share semantic features (i.e., there is not a broad attractor basin for multiple-sense words), meaning that the degree of consistency of the orthographic-to-semantic mappings would be essentially similar for homonyms and polysemous words. If so, in contrast to the position of Rodd et al. (2002, 2004) amongst others, there would be no reason to expect that the speed of the settling process would be different for homonyms than for polysemous words (within a PDP framework) nor would there be any obvious reason, within this framework, why polysemy would facilitate lexical-decision making while homonymy would not.

Data consistent with Klein and Murphy's (2001, 2002) argument have been reported by Hino, Pexman, and Lupker (2006). Hino et al. examined the effect of ambiguity and relatedness of meanings (ROM) using lexical decision and semantic categorization tasks. Following Azuma and Van Orden (1997), Hino et al. manipulated the relatedness of the various meanings of ambiguous words in a quantitative fashion using subjective ratings rather than using dictionary definitions to find words that could be classified as either homonymous or polysemous.[2] Ambiguity itself was also manipulated based on subjective ratings, using a procedure similar to those used by Hino and Lupker (1996) and Kellas et al. (1988). Their experiments involved three word groups: ambiguous words with related meanings, ambiguous words with unrelated meanings, and unambiguous words. In their lexical-decision task, an ambiguity advantage was observed: lexical decision latencies were faster for the two types of ambiguous words than for the unambiguous words. Lexical decision performance was, however, virtually identical for the two types of ambiguous words (i.e., there was no ROM effect).

Hino et al. (2006) interpreted the processing advantages for the two types of ambiguous words as implying that performance in the lexical-decision task has little to do with the process of settling at the semantic level. Rather, this effect was due to semantic feedback (i.e., feedback from the semantic level to the orthographic/lexical level which is the key level for making lexical decisions). The amount of semantic activation generated by reading a word would be greater for ambiguous words than for unambiguous words and, in line with the conclusions of Klein and Murphy (2001, 2002), the amount of semantic activation would be comparable for the two types of ambiguous words since all meanings, even related ones, are essentially represented separately. Thus, the processing advantage produced by the semantic feedback into the orthographic/lexical system should also be similar for the two types of ambiguous words.

The lack of a ROM effect in Hino et al.'s (2006) lexical-decision task is, of course, inconsistent with the data reported by Rodd et al. (2002). Note, however, that the nonwords used in Rodd et al.'s and Hino et al.'s tasks were different. Rodd et al. used pseudohomophones, nonwords that are pronounced the same as real words. In contrast, Hino et al. used standard nonwords. Because Hino et al.'s stimuli were all written in Japanese Katakana script, pseudohomophones could not be used in those experiments because any character string written in Katakana that sounds like a word is considered a word by Japanese readers.

According to Azuma and Van Orden (1997), the type of nonwords used in a lexical-decision task is important because different types of nonwords cause participants to strategically alter the nature of the lexical decision making process. As previously noted, Azuma and Van Orden only observed a ROM advantage when the nonword foils were pseudohomophones in their lexical decision experiments. They argued that when nonword foils are standard nonwords participants typically engage in orthographically-based processing, which is typically sufficient to allow them to accurately make lexical decisions (e.g., Balota & Chumbley, 1984; Besner & McCann, 1987). In contrast, because pseudohomophones are more word-like, when pseudohomophones are used as nonword foils, participants are encouraged to engage a deeper, semantic processing strategy, and hence, lexical decision performance would be more sensitive to the process of semantic activation. As a result, semantic effects, particularly a ROM advantage, would emerge.

In an effort to respond to these claims and, more specifically, in a effort to examine the potential nature of semantic representations and processing more directly, Hino et al. (2006) used a series of semantic categorization tasks in which the word stimuli from their lexical-decision task were used in the negative trials (as nonexemplars). If the speed of semantic activation is modulated by the nature of the orthographic-to-semantic mappings for these words, a ROM advantage (i.e., faster responding to ambiguous words with related meanings than to ambiguous words with unrelated meanings) and a NOM disadvantage (slower responding to ambiguous words with unrelated meanings in comparison to unambiguous words) would be expected in semantic categorization tasks.

In fact, Hino et al. (2006) did observe such a pattern when a broad semantic category (e.g., Living Things) was used in the task (see also Hino et al., 2002). At the same time, however, there were no effects at all when narrower categories (e.g., Vegetables and Animals) were used, suggesting to Hino et al. (2006) that the different results were due to the differences in the decision-making strategies used in these tasks. In the task with a narrower category, participants are able to make a decision by merely checking a small set of features, and hence, the number of meanings would be irrelevant in making decisions for words on the negative trials. In contrast, because a broader category is characterized by more complicated combinations of features (e.g., family resemblance), the decisions would be more difficult and all the activated features would have to be checked before making a decision. Thus, the number of meanings would affect the decision times for words in the negative trials not because of differences in settling time but due to a more complicated decision process. That is, Hino et al. (2006) concluded that the NOM disadvantage and the ROM advantage observed in their Living Thing decision task are best explained in terms of the decision-making process in which spe-

[2] One of Gernsbacher's (1984) main points was that word properties are best measured by determining how those properties are represented in readers' minds (e.g., through norms) rather than relying on what one finds in dictionaries. That was the approach taken by Kellas et al. (1988), Azuma and Van Orden (1997), and Hino et al. (2006) among others.

cific operations are required depending on the category used in the task.

Note also that Hino et al.'s (2006) semantic categorization results stand in contrast with those of Gottlob, Goldinger, Stone, and Van Orden (1999) and Piercey and Joordens (2000), both of whom reported an ambiguity disadvantage in the positive trials of their relatedness judgement tasks (hence, supporting the assumption that the speed of semantic activation is modulated by the nature of orthographic-to-semantic mappings). However, as Pexman, Hino, and Lupker (2004) have noted, a NOM disadvantage in the positive trials of relatedness judgement tasks (e.g., are they related?: VAMPIRE - BAT) can also be produced by the decision-making process due to the negative bias created by the alternative meaning of the ambiguous word (i.e., a baseball bat). Pexman et al. went on to show that NOM (and ROM) effects do not exist when the same stimuli are used on the negative trials of their relatedness judgement tasks (e.g., are they related? TABLE - BAT), a pattern that should not have emerged if those effects were due to the nature/speed of semantic activation.

In sum, although the general notions underlying semantic representations of ambiguous words in a PDP framework (as described by Rodd et al., 2002) seem logical, there is, at present, very little behavioural data supporting these notions and considerable data that are inconsistent with them. Further, most of the supporting data seem to come from lexical decision tasks (Azuma & Van Orden, 1997; Beretta et al., 2005; Klepousniotou & Baum, 2007; Rodd et al., 2002), rather than tasks in which participants are explicitly required to access word meanings, like categorization and/or relatedness judgement tasks (e.g., Hino et al., 2006; Klein & Murphy, 2001, 2002; Pexman et al., 2004).

The Present Research

Given that it is the lexical-decision task that provides the most support for the PDP-based account of semantics under discussion, the present research was an attempt to reexamine the ROM effect using lexical decision tasks. In particular, we conducted three lexical decision experiments using Japanese Katakana-written stimuli. In contrast to the manipulation of ROM in Hino et al. (2006), in these experiments, ROM was manipulated by explicitly selecting homonyms, polysemous words, and unambiguous words as done in Rodd et al. (2002). In Experiment 1, we used pronounceable Katakana-written nonwords as in Hino et al.'s experiment in order to determine whether this change in the operational definition of ROM will produce a different pattern of results. To anticipate our results, the data were no different than those in Hino et al., that is, there was a significant NOM effect but no ROM effect, in spite of the fact that we defined NOM and ROM just as Rodd et al. did. Even though the existence of a NOM effect indicates that semantics must have played some role in the lexical decision making process in both of these experiments, because pseudohomophones were not used as the nonwords one can still question whether responding in either experiment was based on a strategy that sufficiently involved semantic processing. In an attempt to induce even more attention to semantic information during lexical decision making, in Experiments 2 and 3, we included Japanese Kanji words and nonwords in addition to the Katakana words and nonwords.

Although Katakana is a shallow orthography in which the character-to-sound correspondences are highly regular, Kanji is a deep orthography in which the character-to-sound correspondences are more complicated. Typically, a Kanji character has Kun-reading and On-reading pronunciations. The Kun-readings are Japanese original pronunciations, whereas the On-readings are pronunciations imported from China. More relevant to the present study, multicharacter Kana and Kanji words also differ in their morphological structure. That is, while each Katakana character is a phonetic character representing a mora, a combination of a consonant and a vowel, each Kanji is a morpheme and, hence, carries meanings.

Our logic here follows that of Taft (2003, 2004), who suggested that deeper processing would be required in making lexical decisions for polymorphemic words when the nonwords consisted of wrong combinations of real morphemes than when the nonwords were merely orthographically legal letter strings. That is, when nonwords are letter strings, participants would be able to make a correct "word" decision as soon as a morpheme is detected in the stimulus. In contrast, when nonwords are wrong combinations of real morphemes, a "word" decision requires participants to examine whether the given combination of the morphemes creates a meaningful word. Thus, by including multicharacter Kanji words and nonwords in our stimulus set, it should be possible to induce participants to engage in a strategy that explicitly involves semantic processing. If so, task performance should become more sensitive to the process of semantic activation even for Katakana words. Therefore, if the homonymous-polysemous distinction has the representational implications incorporated within PDP models, we should have more chance of observing a ROM effect for the Katakana words when the multicharacter Kanji words and nonwords are involved in the stimulus set in Experiments 2 and 3.

Experiment 1

Method

Participants. Thirty undergraduate students from Chukyo University participated in this experiment for course credit. All were native Japanese speakers who had normal or corrected-to-normal vision.

Stimuli. We selected 14 homonyms, 14 polysemous words, and 14 unambiguous words from a Japanese dictionary (Umesao, Kindaichi, Sakakura, & Hinohara, 1995).[3] These words were three to five Katakana characters in length. Following Rodd et al. (2002), the number of dictionary entries and the number of senses listed within each entry were counted for these words as indices of the number of meanings and the number of senses. All the homonyms had more than one dictionary entry, whereas both the polysemous words and unambiguous words possessed only a single entry. Thus, the number of dictionary entries was significantly larger for the homonyms (2.79) than for the other two types of words (1.00), $F(2, 39) = 23.83$, $MSE = 0.63$, $p < .001$. The polysemous words and unambiguous words differed in terms of

[3] Although we attempted to select as many items as possible in each of the three conditions, we could only use 14 items per condition because we had to match the words on a number of dimensions across the three conditions as described in Table 1.

the number of senses listed within the dictionary entry (4.79 vs. 1.71), $F(1, 26) = 32.26$, $MSE = 2.05$, $p < .001$. In addition, we attempted to equate the number of senses between the polysemous words (4.79) and the homonyms (4.50), $F(1, 26) = 0.13$, $MSE = 4.46$. As such, the number of senses were significantly smaller for the unambiguous words than for the two types of ambiguous words, $F(2, 39) = 12.81$, $MSE = 3.15$, $p < .001$.[4]

According to National Language Research Institute (1970), frequency counts of these words were all less than 50 per 940,533. As shown in Table 1, word length, $F(2, 39) = 0.00$, $MSE = 0.32$, the number of syllables (morae), $F(2, 39) = 0.08$, $MSE = 0.30$, word frequency, $F(2, 39) = 0.49$, $MSE = 94.17$, and orthographic neighbourhood size based on National Language Research Institute (1993), $F(2, 39) = 0.79$, $MSE = 14.83$, were all equated across the three word groups.[5]

In addition, according to the norming data collected by Hino et al. (2006), the familiarity ratings were equivalent across the three word groups, $F(2, 39) = 0.00$, $MSE = 0.43$. NOM ratings for the homonyms and polysemous words were all more than 1.30, whereas the NOM ratings were all less than 1.15 for the unambiguous words.[6]

The degrees of relatedness amongst meanings/senses were also evaluated for the homonyms and the polysemous words using the two types of rating data collected by Hino et al. (2006). Both types of ratings indicated that the degree of relatedness were higher for the polysemous words than for the homonyms.[7]

In addition to the 42 Katakana words, 42 Katakana nonwords were selected from a list of Katakana nonwords that Hino et al. (2006) used in their NOM ratings. The numbers of characters and syllables of these Katakana nonwords were matched to those of the 42 Katakana words. The NOM ratings collected by Hino et al. were all less than 0.05, with an average of 0.01 for the 42 nonwords.

Table 1
Mean Word Frequency (Freq), Word Length (Len), Syllabic Length (Syl), Orthographic Neighbourhood Size (N), Experiential Familiarity Rating (Fam), Number-of-Meaning Rating (NOM), Number of Dictionary Meanings (M), Number of Dictionary Senses (S), Relatedness-of-Meaning Rating Based on Items (Rel-A), and Relatedness-of-Meaning Rating Based on Meaning Pairs (Rel-B) for Each Word Type Used in Experiments 1, 2, and 3

Word type	Freq	Len	Syl	N	Fam
Polysemous words	9.71	3.29	3.29	4.14	4.74
Homonyms	6.07	3.29	3.21	5.36	4.73
Unambiguous words	7.93	3.29	3.21	3.57	4.72
Word type	**NOM**	**M**	**S**	**Rel-A**	**Rel-B**
Polysemous words	1.49	1.00	4.79	4.21	4.04
Homonyms	1.56	2.79	4.50	2.93	2.09
Unambiguous words	1.05	1.00	1.71	—	—

Note. Word frequencies are from National Language Research Institute (1970). Orthographic neighbourhood sizes are counted using National Language Research Institute (1993). The numbers of dictionary meanings and senses are counted using Umesao et al. (1995). Familiarity ratings, number-of-meaning ratings, and the two types of relatedness-of-meaning ratings are from Hino et al. (2006).

Procedure. Participants were tested individually in a normally lit room. Participants were asked to make a word-nonword discrimination for each stimulus appearing on a video monitor (NEC, PC-TV455) by pressing either a “Word” or a “Nonword” key on a keyboard. The two keys that flank the space-key were used as the word and nonword keys (“XFER” and “NFER” keys in a NEC Japanese keyboard). Participants were also told that their responses should be made as quickly and as accurately as possible. Eighteen practise trials were given prior to the 84 experimental trials. Nine Katakana words and nine Katakana nonwords were presented in a random order in the practise trials. None of these stimuli were the same as those used in the experimental trials. During the practise trials, participants were informed about their lexical decision latency and whether their response was correct after each trial. No feedback was given during the experimental trials. The order of stimulus presentation for the experimental trials was randomized for each participant.

Each trial was initiated with a 50 ms 400 Hz beep signal. Following the beep, a fixation point appeared at the centre of the video monitor. One second after the onset of the fixation point, a

[4] The number of dictionary senses for some of our unambiguous words was more than one because it was difficult to collect words with only a single sense. Thus, in our manipulation, the unambiguous words were words with few senses. As shown in Table 1, however, the NOM ratings were all less than 1.15 for these words. Therefore, we termed these words unambiguous words.

[5] Orthographic neighbourhood sizes were calculated using “sakuin.dat” in National Language Research Institute (1993), which is a computer-based dictionary with 36,780 word entries.

[6] According to the NOM ratings collected by Hino et al. (2006), the mean ratings were significantly lower for the unambiguous words (1.05) than for the two types of ambiguous words, $F(2, 39) = 78.56$, $MSE = 0.01$, $p < .001$. However, there was no significant difference in the NOM ratings between the homonyms (1.56) and the polysemous words (1.49), $F(1, 26) = 1.73$, $MSE = 0.02$.

[7] Given a list of ambiguous Katakana words, Hino et al. asked 41 participants to think of all the meanings each word has and to rate the degree of relatedness among these meanings using a 7-point scale from *Unrelated* (1) to *Related* (7). The relatedness ratings were significantly higher for the polysemous words (4.21) than for the homonyms (2.93), $F(1, 26) = 18.18$, $MSE = 0.63$, $p < .001$. In addition, Hino et al. also collected another type of relatedness ratings for the same ambiguous words using a procedure similar to that used by Azuma and Van Orden (1997). In order to first identify the meanings of these words, Hino et al. asked 50 participants to write down all the meanings they could think of given each of these ambiguous words. These responses were classified by three judges using the definitions listed in the dictionary (Umesao et al., 1995). A response was classified as a meaning of a word only if the response was made by more than five participants (10%). After identifying the meanings of each ambiguous word in this manner, all the meanings of the same word were paired with each other and listed in a questionnaire along with a 7-point scale ranging from *Unrelated* (1) to *Related* (7). Thirty-two participants were, then, asked to rate the degree of relatedness for each pair of meanings by circling the appropriate number on the scale. Mean ratings were, then, calculated across all the meaning pairs for each word. The mean relatedness ratings were also significantly higher for the polysemous words (4.04) than for the homonyms (2.09), $F(1, 26) = 45.40$, $MSE = 0.59$, $p < .001$. In addition, the number of identified meanings with this technique was quite comparable between the polysemous words (3.07) and the homonyms (3.07), $F(1, 26) = 0.00$, $MSE = 1.07$.

stimulus was presented above the fixation point. The fixation point and the stimulus were presented in white on a black background. Participants were seated in front of the video monitor at a distance of about 50 cm and asked to respond to the stimulus by pressing either the word or nonword key on the keyboard. The "word" response was made using the participant's dominant hand. The participant's response terminated the presentation of the stimulus and the fixation point. The lexical decision latencies from the onset of the stimulus to the participant's key press and whether the response was correct were automatically recorded by a computer (NEC, PC-9801FA). The intertrial interval was two seconds.

Results

Lexical decision latencies were classified as outliers if they were out of the range of 2.5 standard deviations (*SD*) from the mean lexical decision latency in each condition for each participant. Fifty-one data points (2.02%) were classified as outliers and were excluded from the statistical analyses. In addition, there were 132 error responses (5.24%). These trials were also excluded from the latency analysis. Mean lexical decision latencies for correct responses and mean error rates were calculated across both subjects and items. The mean lexical decision latencies and error rates were submitted to one-way analyses of variance (ANOVAs). In the subjects (F_1) analysis, Stimulus Type was a within-subject factor, while it was a between-item factor in the items (F_2) analysis. The mean lexical decision latencies and error rates from the subjects analysis are presented in Table 2.

In the analyses of lexical decision latencies, the main effect of Stimulus Type was significant in both analyses, $F_1(2, 58) = 23.29$, $MSE = 825.31$, $p < .001$; $F_2(2, 39) = 5.86$, $MSE = 1575.15$, $p < .01$. Lexical decision latencies were significantly shorter for the polysemous words than for the unambiguous words, $F_1(1, 29) = 25.29$, $MSE = 1214.10$, $p < .001$; $F_2(1, 26) = 8.12$, $MSE = 1800.41$, $p < .01$, as well as being significantly shorter for the homonyms than for the unambiguous words, $F_1(1, 29) = 32.66$, $MSE = 821.29$, $p < .001$; $F_2(1, 26) = 7.73$, $MSE = 1685.64$, $p < .025$. In contrast, lexical decision latencies were comparable for the polysemous words and homonyms, $F_1(1, 29) = 0.30$, $MSE = 440.55$; $F_2(1, 26) = 0.04$, $MSE = 1239.39$.

In the analyses of error rates, the main effect of Stimulus Type was not significant in either analysis, $F_1(2, 58) = 0.62$, $MSE = 42.96$; $F_2(2, 39) = 0.29$, $MSE = 43.19$.

Discussion

In Experiment 1, lexical decision latencies were faster for the two types of ambiguous words than for the unambiguous words (i.e., there was a NOM effect). In addition, lexical decision performance was comparable for the two types of ambiguous words (i.e., there was no ROM effect). These results are extremely similar to those of Experiment 1 in Hino et al. (2006), in which a processing advantage was observed for ambiguous words, but no ROM effect was detected in lexical decision performance. They contrast, however, with the results of Azuma and Van Orden (1997) and Rodd et al. (2002). In particular, Rodd et al. reported evidence of a homonymy disadvantage in addition to a polysemy advantage (and, hence, a ROM effect).

As previously noted, the nonwords in Azuma and Van Orden's (1997) and Rodd et al.'s (2002) studies differed from those in Experiment 1 (and from the experiments in Hino et al., 2006). Whereas Azuma and Van Orden and Rodd et al. used pseudohomophones in their tasks, we used Katakana-written pronounceable nonwords. When those authors used standard pronounceable nonwords, they observed no effects. As such, the argument could be made that the discrepancy between our results and those of Azuma and Van Orden and Rodd et al. is due to our participants engaging in a shallower, more orthographically-based processing strategy when making their lexical decisions, whereas Azuma and Van Orden and Rodd et al.'s participants engaged in a more semantically-based processing strategy. As a result, lexical decision performance in the present experiment may have not reflected the speed of semantic activation, leading to the null ROM effect.

If such an argument is correct, of course, then a further implication is that the NOM advantage for the two types of ambiguous words also must have had little to do with activation at the semantic level. That is, if the lack of a ROM effect was due to the fact that responding in Experiment 1 was not based on processing at the semantic level but, rather, was based only on processing at the orthographic level, then it follows that the observed NOM effects also couldn't have been due to processing at the semantic level. Therefore, an account like Rodd et al.'s (2002; 2004) would be unable to provide a complete explanation of the results in Experiment 1 even if one were to accept the assumption that the nature of processing was different (i.e., shallower) in Experiment 1 than it was in their experiments. Note, however, that the present results would be readily explained in terms of an account based on

Table 2
Mean Lexical Decision Latencies in Millisecond (Ms) and Error Rates in Percent (%) for Each Stimulus Group in Experiment 1

			Ambiguity effect		ROM effect	
Word type	RT	ER	RT	ER	RT	ER
Polysemous words	537 (9.37)	4.13 (1.49)	+46*	+1.84	+3	+0.57
Homonyms	540 (9.70)	4.70 (0.96)	+43*	+1.27		
Unambiguous words	583 (14.30)	5.97 (1.18)				
Katakana nonwords	613 (14.26)	4.57 (0.94)				

Note. RT and ER stand for mean reaction time and error rate, respectively. Standard error of the mean is in parenthesis (). The asterisk * denotes a significant effect, $p < .05$, in the subjects' analysis.

the idea that there was more semantic feedback to the orthographic level for the two types of ambiguous words than for the unambiguous words, facilitating orthographic processing (Hino et al., 2006; Pexman et al., 2004).

Experiment 2

The goal of Experiment 2 was to try to alter our participants' processing strategy, particularly to encourage them to engage in deeper, semantically-based processing in an attempt to determine whether the data pattern would change. To accomplish this, we added a set of two-character Kanji words and nonwords to our 42 Katakana words and 42 Katakana nonwords from Experiment 1. Because Kanji characters are morphemes, deeper processing would have to be carried out when a stimulus set involves multi-character Kanji words and nonwords and, as a result, the expectation was that participants would be biased to engage in more semantically-based processing in general (e.g., Taft, 2003, 2004). If so, there would, presumably, be more chance of observing a ROM effect for our Katakana words if the speed of semantic activation actually is modulated by ROM.

Method

Participants. Thirty undergraduate students from Chukyo University participated in this experiment for course credit. All were native Japanese speakers who had normal or corrected-to-normal vision. None had participated in Experiment 1.

Stimuli. In addition to the 42 Katakana words and 42 Katakana nonwords used in Experiment 1, 42 Kanji words and 42 Kanji nonwords were used in Experiment 2. Kanji words and nonwords consisted of two Kanji characters. The 42 Kanji words consisted of 21 high frequency and 21 low frequency Kanji words. According to National Language Research Institute (1970), frequency counts for the high frequency Kanji words were all more than 150, with an average of 346.38, while frequency counts for the low frequency Kanji words were all five. These Kanji words were two to four syllables (morae) in length, with an average of 3.24.

The 42 Kanji nonwords were created by combining two unrelated Kanji characters that had only a single pronunciation, and thus, the syllabic lengths of the nonwords were matched with those of the Kanji words. As such, the complete stimulus set in Experiment 2 consisted of 42 Katakana words, 42 Katakana nonwords, 42 Kanji words, and 42 Kanji nonwords.

Procedure. The procedure and apparatus in Experiment 2 were the same as those in Experiment 1 except in the following respects. Twenty-four practise trials were followed by 168 experimental trials. Six Katakana words, six Katakana nonwords, six Kanji words, and six Kanji nonwords were presented in a random order in the practise trials. None of these stimuli were used in the experimental trials.

Results

Lexical decision latencies were classified as outliers if they were out of the range of 2.5 *SD* from the mean lexical decision latency in each condition for each participant. Across all trials, 115 data points (2.28%) were classified as outliers and were excluded from the statistical analyses. In addition, 309 error trials (6.13%) were excluded from the latency analysis. Mean lexical decision latencies for correct responses and mean error rates were calculated across both subjects and items. The mean lexical decision latencies and error rates from the subjects analysis are presented in Table 3.

Lexical decision latencies and error rates for the Katakana word trials were separately submitted to one-way ANOVAs. In the analyses of lexical decision latencies, the main effect of Stimulus Type was significant in the subjects analysis, $F_1(2, 58) = 8.13$, $MSE = 732.45$, $p < .01$, although not in the items analysis, $F_2(2, 39) = 1.90$, $MSE = 1367.56$. In addition, the 22 ms difference in lexical decision latencies between the polysemous words and homonyms was significant in the subjects analysis, $F_1(1, 29) = 12.41$, $MSE = 578.41$, $p < .01$, although not in the items analysis, $F_2(1, 26) = 2.40$, $MSE = 1502.28$. The 26 ms difference in lexical decision latencies between the polysemous words and the unambiguous words was significant in the subjects analysis, $F_1(1, 29) = 20.17$, $MSE = 515.34$, $p < .001$, and marginally significant in the items analysis, $F_2(1, 26) = 3.69$, $MSE = 1125.31$, $p < .07$. In contrast, lexical decision latencies were comparable between the homonyms and the unambiguous words, $F_1(1, 29) = 0.27$, $MSE = 1103.60$; $F_2(1, 26) = 0.01$, $MSE = 1475.09$.

In the analyses of error rates, the main effect of Stimulus Type was not significant in either analysis, $F_1(2, 58) = 1.73$, $MSE = 23.71$; $F_2(2, 39) = 0.93$, $MSE = 18.59$.

Table 3
Mean Lexical Decision Latencies in Millisecond (Ms) and Error Rates in Percent (%) for Each Stimulus Group in Experiment 2

			Ambiguity effect		ROM effect	
Word type	RT	ER	RT	ER	RT	ER
Polysemous words	544 (10.47)	3.60 (0.96)	+26*	+1.63	+22*	−0.63
Homonyms	566 (12.09)	2.97 (0.77)	+4	+2.26		
Unambiguous words	570 (13.27)	5.23 (1.00)				
HF Kanji word	508 (9.86)	1.67 (0.60)				
LF Kanji word	673 (14.75)	17.77 (2.20)				
Katakana nonword	634 (12.72)	2.83 (0.54)				
Kanji nonword	664 (15.22)	6.17 (1.22)				

Note. RT and ER stand for mean reaction time and error rate, respectively. Standard error of the mean is in parenthesis (). The asterisk * denotes a significant effect, $p < .05$, in the subjects' analysis.

In the analyses of lexical decision latencies for the Kanji word trials, the 165 ms frequency effect was significant in both analyses, $F_1(1, 29) = 380.21$, $MSE = 1064.22$, $p < .001$; $F_2(1, 40) = 151.04$, $MSE = 1964.68$, $p < .001$. In addition, the 16.1% frequency effect on error rates was also significant in both analyses, $F_1(1, 29) = 56.85$, $MSE = 68.39$, $p < .001$; $F_2(1, 40) = 31.76$, $MSE = 89.23$, $p < .001$.

Combined Analyses of the Data in the Katakana Word Trials in Experiments 1 and 2

In order to contrast the data for the Katakana words in Experiments 1 and 2, we conducted Stimulus Type by Experiment ANOVAs for lexical decision latencies and error rates in the two experiments. In the subjects analysis, Stimulus Type was a within-subject factor, whereas Experiment was a between-subject factor. In the items analysis, Stimulus Type was a between-item factor, while Experiment was a within-item factor.

In the analyses of lexical decision latencies, the main effect of Stimulus Type was significant both in the subjects analysis, $F_1(2, 116) = 25.43$, $MSE = 778.88$, $p < .001$, and in the items analysis, $F_2(2, 39) = 3.45$, $MSE = 2556.68$, $p < .05$. The main effect of Experiment was not significant in either analysis, $F_1(1, 58) = 0.17$, $MSE = 10715.52$; $F_2(1, 39) = 1.74$, $MSE = 386.02$. It is most important to note that the interaction between Stimulus Type and Experiment was significant in both analyses, $F_1(2, 116) = 6.89$, $MSE = 778.88$, $p < .01$; $F_2(2, 39) = 7.81$, $MSE = 386.02$, $p < .01$. The significant interaction reflects the fact that different patterns of results were observed for the same Katakana words in Experiments 1 and 2. No significant effect was detected in the analyses of error rates, all Fs < 1.90.

Discussion

The significant Stimulus Type by Experiment interaction in the combined analyses of lexical decision latencies indicates that the pattern of results in Experiment 2 was different from that in Experiment 1. Specifically, the results in Experiment 2 were closer to those reported by Rodd et al. (2002) in that a processing time advantage was observed for the polysemous words over both the homonyms and the unambiguous words. Therefore, it is clear that the manipulations used in Experiment 2 did cause our participants to invoke a different processing strategy than the one used in Experiment 1. The question, however, is exactly what was the nature of that strategy?

To begin with, note that, in contrast to the ROM advantage observed in the present experiment, we failed to observe any evidence for the expectation, based on Rodd et al.'s (2004) model, that homonyms would be more difficult to process than unambiguous words. That is, in spite of the fact that homonyms presumably are represented at the semantic level by at least two different attractor basins, which should cause competition in comparison to unambiguous words, which have only one, there was no significant difference between latencies in those two conditions. In fact, the small difference that was observed went in the opposite direction.

One could potentially argue that the lack of the homonymy disadvantage could have been due to the fact that the homonyms possessed more senses than the unambiguous words. As shown in Table 1, the unambiguous words possessed 1.71 senses on average, whereas the average number of senses was 4.50 for the homonyms. As such, the processing disadvantage for the homonyms could have been cancelled by the processing advantage created by having more senses. Note, however, that the number of senses for the homonyms was the total number of senses listed in multiple dictionary entries (2.79 entries on average). Thus, it reflects the fact that the homonyms possessed, on average, 1.61 senses for each of the 2.79 dictionary entries, indicating that the number of senses involved for each dictionary entry was quite similar for the homonyms and the unambiguous words (1.61 vs. 1.71). As such, the homonyms would not have had attractor basins that were any larger than those for the unambiguous words. Therefore, there seems to be no reason not to expect a processing disadvantage for the homonyms over the unambiguous words if Rodd et al.'s (2002; 2004) analysis is correct.

What should also be noted was that the lexical decision latencies for the experimental Katakana words were comparable across the two experiments (i.e., the main effect of Experiment was not significant for the Katakana words). If participants actually were engaging in a deeper, semantic processing strategy when the Kanji words and nonwords were involved in the stimulus set in Experiment 2, lexical decision latencies would be expected to have been longer in Experiment 2 than in Experiment 1.

There are a number of aspects of the results, therefore, that are inconsistent with the idea that the ROM effect in Experiment 2 was caused by participants engaging in deeper, semantic-level processing. Nonetheless, as is clear, there must have been was a change in the processing strategy from Experiment 1 to Experiment 2, presumably caused by the insertion of the Kanji words and nonwords. One possible way in which the processing strategy might have changed is that the relatedness of meanings possessed by Kanji characters may have been used as a cue to make lexical decisions for the two-character Kanji words and nonwords. Given the fact that meanings are activated based on a subset of a word's spellings (e.g., Bowers, Davis, & Hanley, 2005), it is quite likely that semantic activation did arise for the constituent characters of Kanji words and nonwords. The meanings activated by these characters would have been related for the Kanji words but unrelated for the Kanji nonwords because the Kanji nonwords were created by combining two "unrelated" Kanji characters. Thus, the relatedness of activated meanings would indicate that the presented stimulus was a word. Past research, indicating that prime-target relatedness can be used as a decision cue in a lexical-decision task (e.g., Balota & Lorch, 1986; McNamara & Altarriba, 1988; Neely, 1991), would appear to provide support for idea that participants can use this type of strategy in these types of tasks.

If participants were using such a decision-making strategy, it would, presumably, also have implications for performance on Katakana trials. In particular, the relatedness of activated meanings may have been used in making lexical decisions for Katakana stimuli, with the activated senses of a polysemous word providing a clue to make a positive decision and the activated meanings of a homonym biasing participants to make a negative decision, producing the observed ROM effect.

Experiment 3

Experiment 3 was an attempt to examine this alternative possibility. If the ROM effect in Experiment 2 was due to a decision-making bias deriving from the relatedness of activated meanings, this effect should disappear if the Kanji nonwords were created by combining two related Kanji characters. Specifically, if Kanji nonwords are created by taking Kanji words with related characters and transposing those characters, the activated meanings would be related both for words and nonwords. In such a situation, the relatedness amongst activated meanings would not provide any clue to making lexical decisions, and hence, any ROM effect produced by that strategy should disappear.

In Experiment 3, the Katakana words and nonwords and the Kanji words were the same as those used in Experiment 2; however, a new set of two-character Kanji nonwords was created in which the constituent characters have similar meanings. In order to create the Kanji nonwords, we first collected Kanji words with two characters similar in meaning. For example, "道路" was selected because both characters denote "road". The two characters were, then, transposed to create a nonword, so that the nonword, "路道" was created from "道路". Because the relatedness amongst activated meanings no longer provides a clue for making lexical decisions for our Kanji stimuli, if a strategy based on relatedness of meanings is what produced the ROM effect in Experiment 2, then no ROM effect would be expected in Experiment 3.

It should also be noted that Taft, Zhu, and Peng (1999) reported that the nonwords created by transposing characters from real words produce longer lexical decision latencies for Chinese readers. According to Taft et al., the transposed character stimuli activate lexical representations of the (nontransposed) base words, which delays responding. If the transposed Kanji nonwords do activate the lexical representations of their base words, using this type of nonword should also produce longer lexical decision latencies in Experiment 3 than in Experiment 2, at least for the Kanji stimuli. In addition, because this type of nonword would be more word-like in the sense that these nonwords would activate lexical and semantic representations of their base words, using these nonwords should actually be more likely to bias participants to engage in a semantic processing strategy than the manipulation in Experiment 2. Thus, if the ROM effect in Experiment 2 was due to a reliance on activation at the semantic level, a ROM effect should also emerge in Experiment 3.

Method

Participants. Thirty-eight undergraduate and graduate students from Waseda University participated in this experiment. All were native Japanese speakers who had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Stimuli. The 42 Katakana words and 42 Katakana nonwords used in Experiments 1 and 2 and the 42 Kanji words used in Experiment 2 were once again used in this experiment. In addition, 42 new Kanji nonwords were created in the following manner.

We first selected 42 Kanji words that consisted of two Kanji characters with similar meanings from National Language Research Institute (1970). Frequency counts of these words were all more than 15, with an average of 48.33. These words were two to four syllables in length, with an average of 3.21. From these words, 42 Kanji nonwords were created by transposing the two characters.

In order to assure that these Kanji nonwords consisted of characters with similar meanings, we asked a new set of 26 students from Waseda University to rate the similarity of meanings for the two constituent characters of each of the Kanji nonwords used in Experiments 2 and 3. None of the students participated in any of the lexical decision experiments. The 84 Kanji nonwords from Experiments 2 and 3 were randomly ordered and printed along with the 7-point scale ranging from *Dissimilar* (1) to *Similar* (7). Participants were asked to rate the similarity in meanings of the two Kanji characters by circling the appropriate number on the scale.

The similarity ratings were all less than 3.00 for the two constituent Kanji characters of the nonwords used in Experiment 2, with an average of 1.82. In contrast, for the constituent characters of the nonwords created for Experiment 3, the similarity ratings were all more than 4.00, with an average of 5.32. As such, the similarity ratings for the Kanji pairs were significantly higher for the Kanji nonwords used in Experiment 3 than for those used in Experiment 2, $F(1, 82) = 1229.85$, $MSE = 0.21$, $p < .001$.[8]

Procedure. Stimuli were presented on a video monitor (Iiyama HM204DA) driven by an IBM-AT compatible computer. The participants' responses were collected using two buttons on a button-box connected to the computer. In all other respects, the procedure was identical to that in Experiment 2.

Results

Lexical decision latencies were classified as outliers if they were out of the range of 2.5 *SD* from the mean lexical decision latency in each condition for each participant. Across all trials, 152 data points (2.38%) were classified as outliers and were excluded from the statistical analyses. In addition, 540 error trials (8.46%) were excluded from the latency analysis. Mean lexical decision latencies for correct responses and mean error rates were calculated across both subjects and items. The mean lexical decision latencies and error rates from the subjects analysis are presented in Table 4.

In the analyses of lexical decision latencies, the main effect of Stimulus Type was significant in the subjects analysis, $F_1(2, 74) = 15.23$, $MSE = 732.50$, $p < .001$, and marginally significant in the items analysis, $F_2(2, 39) = 2.88$, $MSE = 1216.94$, $p < .07$. In addition, the 31 ms ambiguity advantage for the polysemous words was significant in the subjects analysis, $F_1(1, 37) = 20.02$, $MSE = 881.49$, $p < .001$, and marginally significant in the items analysis, $F_2(1, 26) = 3.89$, $MSE = 1404.45$, $p < .06$. The 29 ms ambiguity advantage for the homonyms was significant in both analyses, $F_1(1, 37) = 25.49$, $MSE = 618.39$, $p < .001$; $F_2(1, 26) = 4.66$, $MSE = 1084.02$, $p < .05$. In contrast, lexical decision latencies

[8] Because the Kanji nonwords in Experiment 3 were transposed-character nonwords, they should be more word-like than those used in Experiment 2 according to Taft et al. (1999). In addition to the higher similarity ratings of the constituent characters, the summed frequencies of the characters computed based on National Language Research Institute (1993) were also higher for the Kanji nonwords in Experiment 3 than for those in Experiment 2, 446.19 vs. 231.86, $F(1, 82) = 9.46$, $MSE = 102032.04$, $p < .01$. The similar trend was also found in the orthographic neighbourhood sizes, 22.69 vs. 16.36, $F(1, 82) = 3.62$, $MSE = 232.45$, $p = .06$. As such, these characteristics also indicate that the Kanji nonwords in Experiment 3 were more word-like than those in Experiment 2.

Table 4
Mean Lexical Decision Latencies in Millisecond (Ms) and Error Rates in Percent (%) for Each Stimulus Group in Experiment 3

Word type	RT	ER	Ambiguity effect RT	Ambiguity effect ER	ROM effect RT	ROM effect ER
Polysemous words	545 (10.18)	2.66 (0.68)	+31*	+2.81*	+2	+0.79
Homonyms	547 (11.37)	3.45 (0.80)	+29*	+2.02		
Unambiguous words	576 (11.39)	5.47 (0.99)				
HF Kanji word	552 (13.20)	1.84 (0.51)				
LF Kanji word	780 (23.55)	16.45 (2.29)				
Katakana nonword	643 (18.35)	3.45 (0.65)				
Kanji nonword	844 (24.69)	16.11 (1.90)				

Note. RT and ER stand for mean reaction time and error rate, respectively. Standard error of the mean is in parenthesis (). The asterisk * denotes a significant effect, $p < .05$, in the subjects' analysis.

were comparable for the polysemous words and homonyms, $F_1(1, 37) = 0.08$, $MSE = 697.61$; $F_2(1, 26) = 0.01$, $MSE = 1162.37$.

In the analyses of error rates, the main effect of Stimulus Type was significant only in the subjects analysis, $F_1(2, 74) = 4.57$, $MSE = 17.54$, $p < .025$; $F_2(2, 39) = 1.63$, $MSE = 19.29$. The 2.81% difference between polysemous words and unambiguous words was significant in the subjects analysis, $F_1(1, 37) = 7.86$, $MSE = 19.16$, $p < .01$, and marginally significant in the items analysis, $F_2(1, 26) = 3.27$, $MSE = 18.39$, $p < .09$. The 2.02% difference between homonyms and unambiguous words was marginally significant in the subjects analysis, $F_1(1, 37) = 3.58$, $MSE = 21.77$, $p < .07$, although not in the items analysis, $F_2(1, 26) = 1.33$, $MSE = 21.03$. In contrast, error rates were comparable for the polysemous words and homonyms, $F_1(1, 37) = 1.01$, $MSE = 11.68$; $F_2(1, 26) = 0.33$, $MSE = 18.46$.

In the analyses of lexical decision latencies for the Kanji word trials, the 228 ms frequency effect was significant in both analyses, $F_1(1, 37) = 154.71$, $MSE = 6369.98$, $p < .001$; $F_2(1, 40) = 90.90$, $MSE = 6503.65$, $p < .001$. The 14.61% frequency effect on error rates was also significant in both analyses, $F_1(1, 37) = 40.93$, $MSE = 99.02$, $p < .001$; $F_2(1, 40) = 48.20$, $MSE = 47.78$, $p < .001$.

Combined Analyses of the Katakana Word Data in Experiments 1 and 3

In order to compare the data from the Katakana word trials in Experiment 3 with those in Experiment 1, we conducted Stimulus Type by Experiment ANOVAs for lexical decision latencies and error rates of the two experiments. In the analyses of lexical decision latencies, the main effect of Stimulus Type was significant in both analyses, $F_1(2, 132) = 39.06$, $MSE = 773.28$, $p < .001$; $F_2(2, 39) = 4.87$, $MSE = 2477.83$, $p < .025$. The main effect of Experiment was not significant in either analysis, $F_1(1, 66) = 0.33$, $MSE = 11275.72$; $F_2(1, 39) = 0.25$, $MSE = 314.26$. It is most important to note that the interaction between Stimulus Type and Experiment was not significant in either analysis, $F_1(2, 132) = 1.45$, $MSE = 773.28$; $F_2(2, 39) = 2.17$, $MSE = 314.26$. These results reflect the fact that the pattern of results was quite similar for the same Katakana words in Experiments 1 and 3.

In the analyses of error rates, the main effect of Stimulus Type was significant only in the subjects analysis, $F_1(2, 132) = 3.34$, $MSE = 28.71$, $p < .05$; $F_2(2, 39) = 0.96$, $MSE = 43.84$. No other effect was significant in either analysis, all Fs < 1.3.

Combined Analyses of the Katakana Word Data in Experiments 2 and 3

In addition, we also conducted Stimulus Type by Experiment ANOVAs for lexical decision latencies and error rates of the Katakana word trials in Experiments 2 and 3. In the analyses of lexical decision latencies, the main effect of Stimulus Type was significant in the subjects analysis, $F_1(2, 132) = 18.64$, $MSE = 732.48$, $p < .001$, although not in the items analysis, $F_2(2, 39) = 2.12$, $MSE = 2260.85$. The main effect of Experiment was not significant in either analysis, $F_1(1, 66) = 0.06$, $MSE = 11956.84$; $F_2(1, 39) = 0.90$, $MSE = 323.65$. It is most important to note that the interaction between Stimulus Type and Experiment was significant in both analyses, $F_1(2, 132) = 3.88$, $MSE = 732.48$, $p < .05$; $F_2(2, 39) = 4.03$, $MSE = 323.65$, $p < .05$. These results reflect the fact that the pattern of results was different for the same Katakana words in Experiments 2 and 3.

In the analyses of error rates, the main effect of Stimulus Type was significant only in the subjects analysis, $F_1(2, 132) = 5.28$, $MSE = 20.25$, $p < .01$; $F_2(2, 39) = 1.39$, $MSE = 29.72$. No other effect was significant in either analysis, all Fs < 1.

Combined Analyses of the Kanji Word Data From Experiments 2 and 3

In order to compare the data from the Kanji word trials in Experiments 2 and 3, Frequency by Experiment ANOVAs were conducted for lexical decision latencies and error rates. In the subjects analysis, Frequency was a within-subject factor, whereas Experiment was a between-subject factor. In the items analysis, Frequency was a between-item factor, whereas Experiment was a within-item factor.

In the analyses of lexical decision latencies, the main effect of Frequency was significant in both analyses, $F_1(1, 66) = 318.91$, $MSE = 4038.66$, $p < .001$; $F_2(1, 40) = 132.69$, $MSE = 6502.38$, $p < .001$, reflecting the fact that lexical decision latencies were faster for high frequency words than for the low frequency words. The main effect of Experiment was also significant in both analyses, $F_1(1, 66) = 12.30$, $MSE = 15640.38$, $p < .001$; $F_2(1, 40) =$

67.28, $MSE = 1965.95$, $p < .001$, reflecting the fact that lexical decision latencies were longer in Experiment 3 than in Experiment 2. In addition, the interaction between Frequency and Experiment was also significant in both analyses, $F_1(1, 66) = 8.37$, $MSE = 4038.66$, $p < .01$; $F_2(1, 40) = 12.78$, $MSE = 1965.95$, $p < .01$, reflecting the fact that the frequency effect size was larger in Experiment 3 than in Experiment 2.

In the analyses of error rates, the main effect of Frequency was the only significant effect, $F_1(1, 66) = 92.37$, $MSE = 85.56$, $p < .001$; $F_2(1, 40) = 51.49$, $MSE = 99.50$, $p < .001$. No other effect was significant in either analysis, all Fs < 1.

Discussion

As in Experiment 2, the stimulus set in Experiment 3 contained Kanji words and nonwords. Further, the Kanji nonwords used in Experiment 3 were even more word-like than those in Experiment 2 because they were created by transposing Kanji characters of real words and, hence, would be expected to activate lexical and semantic representations of their base words as in Taft et al. (1999). Thus, if the pattern of results in Experiment 2 reflected the fact that lexical decisions were based on activation at the semantic level and that there are different speeds of semantic activation for homonyms, polysemous words, and unambiguous words written in Katakana, the results for the Katakana words in Experiment 3, specifically the occurrence of a ROM effect, should mimic those in Experiment 2.

In contrast to the results of Experiment 2, however, lexical decision performance was comparable for the polysemous words and homonyms in Experiment 3 with both word types showing an advantage over unambiguous words. The significant interaction between Stimulus Type and Experiment in the combined analyses of Experiments 2 and 3, along with the lack of the same interaction in the combined analyses of Experiments 1 and 3, clearly reflects the fact that the pattern of results observed in Experiment 3 was quite different from that in Experiment 2 and similar to that in Experiment 1. Specifically, processing time advantages were observed for the two types of ambiguous words with no ROM effect being observed in Experiments 1 and 3.

The implication of these results is that the ROM effect in Experiment 2 does not appear to have been due to a reliance on semantic-level activation in making lexical decisions with different speeds of semantic activation for homonyms and polysemous words producing different lexical decision latencies. Note also that, although the lexical decision latencies for the same Kanji words were longer in Experiment 3 than in Experiment 2, lexical decision latencies were comparable across all the experiments for the Katakana words. These results suggest that there was virtually no change in the processing depth for the Katakana words across the three experiments, and hence, there is no reason to believe that the task performance reflected the speed of semantic-level activation only in Experiment 2.

Instead, the results appear to indicate that the ROM effect in Experiment 2 was due to the use of a decision strategy based on the relatedness amongst activated meanings. As the similarity rating data clearly illustrate, the meanings were dissimilar for the two constituent characters of the Kanji nonwords used in Experiment 2. In contrast, the meanings of the two constituent characters were related in a meaningful manner for the Kanji words. Thus, the relatedness amongst activated meanings provides a clue to discriminate Kanji words from nonwords. If the relatedness amongst activated meanings was also used as a clue to make lexical decisions for Katakana stimuli, a positive bias would be created for the polysemous words but not for homonyms. As a result, a ROM effect would emerge due to the decision-making process in Experiment 2, as was observed.

In Experiment 3, on the other hand, the relatedness amongst activated meanings was not a valid cue for making lexical decisions because Kanji nonwords consisted of two Kanji characters with similar meanings. As a result, the relatedness amongst activated meanings could not be effectively used in making lexical decisions, and hence, the decision-making process would not be affected by ROM for the ambiguous Katakana words. The lack of a ROM effect in Experiment 3, therefore, suggests that the ROM effect in Experiment 2 was most likely a decision-making effect based on a strategy that was useful in that specific task situation.

General Discussion

Ambiguous words present a clear challenge for those wishing to understand the nature of semantic representations and how those representations are activated from print. PDP models, models in which representational units can be viewed as semantic features and semantic activation can be viewed as the process of activating the correct set of features, seem to provide a promising way of addressing this challenge. Unfortunately, the most straightforward way of implementing these ideas within PDP frameworks seems to be generally inconsistent with the existing data.

Within PDP frameworks, the speed of semantic activation is assumed to be modulated by the nature of a word's orthographic-to-semantic mappings. Thus, as argued by Rodd et al. (2002, 2004) the settling speed at the semantic level would be expected to be modulated by both the number of meanings a word has and the relatedness of those meanings. In particular, the basic expectation is that the settling speed should be slower for words with more meanings (ambiguous words) than for words with fewer meanings (unambiguous words). Under the straightforward assumption that polysemous senses share semantic units while homonymous meanings are separately represented, the orthographic-to-semantic mappings would be less consistent for homonyms than for polysemous words. Therefore, PDP models predict that the speed of semantic activation should be slower for homonyms than for both polysemous words and unambiguous words. Although there is at least some data (Azuma & Van Orden, 1994; Rodd et al., 2002) consistent with this prediction, most of the data reported in the literature show that ambiguous words are responded to more rapidly than unambiguous words regardless of what factor creates that ambiguity (Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino et al., 2002; Hino et al., 1998; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas et al., 1988; Millis & Button, 1989; Pexman & Lupker, 1999; Rubenstein et al., 1970; Rubenstein et al., 1971).

There are also data (Klein & Murphy, 2001, 2002) suggesting that both polysemous senses and homonymous meanings are separately represented in the semantic system. If Klein and Murphy's position were incorporated into a PDP framework, such a model would not predict that the speed of semantic activation would be modulated by the homonymous-polysemous distinction. Consistent with this position, Hino et al. (2006) and Pexman et al. (2004) failed to find evidence indicating that the speed of semantic activation is modulated by ROM and NOM in their semantic-decision tasks. However, even if this assumption were implemented within a PDP model, such a model would still not be able to explain the advantage polysemous words have over unambiguous words in lexical decision tasks if one also assumes that lexical decision making is driven by semantic activation.

Given this situation, and given the fact that a lexical-decision task does involve task-specific processing components that, in theory, could produce a ROM effect, the question is whether Azuma and Van Orden's (1997), Rodd et al.'s (2002) and Klepousniotou and Baum's (2007) findings provide good support for the claim that there are different speeds of semantic activation for homonyms and polysemous words. In order to reexamine the locus of the ROM effect in lexical decision, therefore, we conducted three lexical decision experiments using Katakana-written homonyms, polysemous words, and unambiguous words. Following Rodd et al., we manipulated the number of homonymous meanings and the number of polysemous senses for these words by counting the number of dictionary entries and the number of senses within these entries. The homonyms had more than one dictionary entry whereas the polysemous words and unambiguous words possessed only a single entry. The number of senses listed in the dictionary was, of course, smaller for the unambiguous words than for the two other word types, but the number of senses were comparable for the polysemous words and the homonyms.

In Experiment 1, these Katakana words were used with pronounceable Katakana-written nonwords. In this experiment, only a NOM effect was observed (both the polysemous words and the homonyms were faster than unambiguous words, and there was no difference between polysemous words and homonyms). These results seem most consistent with Klein and Murphy's (2001, 2002) position concerning representation, coupled with the assumption that any NOM effect is not due to the process of settling at the semantic level but rather is due to something like greater semantic feedback to the orthographic level for the homonyms and polysemous words (e.g., Hino & Lupker, 1996; Hino et al., 2002; Hino et al., 2006; Lupker, 2007; Pexman et al., 2004; Pexman & Lupker, 1999).

At the same time, however, one could argue that the reason the results in Experiment 1 were different from those in the lexical decision studies showing a ROM effect (e.g., Azuma & Van Orden, 1997; Beretta et al., 2005; Rodd et al., 2002) is that semantic activation played, at most, a small role in the process because our nonwords were not pseudohomophones. Thus, in Experiment 2, we attempted to bias participants to engage in a deeper, semantically-based processing by including two-character Kanji words and nonwords in the stimulus set (Taft, 2003, 2004). With the addition of Kanji words and nonwords in Experiment 2, the results for the Katakana words were more similar to those reported by Rodd et al. (2002). That is, a ROM advantage was observed in that the polysemous words were responded to faster than the homonyms in spite of the fact that the number of senses was equated across the two conditions.

In contrast to the other predictions drawn from the Rodd et al.'s position (2002; 2004), however, a) there was no evidence of a homonymy disadvantage compared to the unambiguous words, a disadvantage that should accrue due to the competition between the unrelated meanings, and b) lexical decision latencies for the experimental Katakana words were quite comparable across the two experiments. That is, because homonyms possess multiple unrelated meanings, they are assumed to have multiple basins of attraction at the semantic level, whereas a single attractor basin is assumed for an unambiguous word. As a result, the settling process should be harder for homonyms than for unambiguous words. In addition, if a ROM advantage was observed in Experiment 2 because the decisions were made based on deeper semantic processing, the lexical decision latencies should have been longer in Experiment 2 than when lexical decision making was presumably based on shallower processing (i.e., in Experiments 1 and 3).

As such, we proposed an alternative account for these data in terms of a change in the decision-making strategy when Kanji stimuli were involved in the stimulus set. Because Kanji characters are morphemes, it's likely that their meanings are activated by the constituent Kanji characters for the Kanji words and nonwords. These activated meanings would be related for the Kanji words but unrelated for the Kanji nonwords because the Kanji nonwords were created by pairing unrelated Kanji characters. As such, the relatedness of activated meanings would provide a clue to making lexical decisions. That is, the relatedness amongst activated meanings would produce a bias toward a positive decision. If such a decision strategy were employed not only for Kanji stimuli but also for Katakana stimuli, a positive bias would be created for the polysemous words but not for the homonyms and unambiguous words, producing a processing time advantage for polysemous words over homonyms and unambiguous words.

We evaluated this idea in Experiment 3 by using Kanji nonwords with constituent Kanji characters with similar meanings. To accomplish this, we collected two-character Kanji words that consisted of characters with similar meanings and then created Kanji nonwords by transposing the characters. Because it is likely that transposed Kanji nonwords activate the lexical and semantic representations of their base words (e.g., Taft et al., 1999), these nonwords would have been more word-like than those used in Experiment 2. Thus, these nonwords should have biased participants to engage a semantic processing strategy to at least as great an extent as the nonwords used in Experiment 2. Therefore, according to an account of the sort provided by Rodd et al. (2002, 2004), the pattern of results in Experiment 3 should have been essentially the same as that in Experiment 2. In contrast, our alternative, decision making account of the ROM effect in Experiment 2 leads to the prediction that the results in Experiment 3 should be similar to those in Experiment 1 because the relatedness of activated meanings would no longer provide a valid clue in making lexical decisions.

The results were consistent with the latter prediction. That is, as in Experiment 1, a NOM effect but no ROM effect was observed in Experiment 3. Therefore, our conclusion is that the results of Experiment 2 were most likely not due to differences in the semantic activation process for polysemous words and homonyms.

Where Is the Locus of the ROM Effect?

According to the three combined analyses of lexical decision latencies of the Katakana word trials, the main effect of Experiment was never significant, reflecting the fact that lexical decision latencies for the Katakana words were comparable across the three experiments. Although we used different sets of participants across the three experiments, given the comparable latencies for the Katakana words, one may still wish to argue that we were not successful in biasing participants to engage in semantic processing in Experiments 2 and 3, and hence, the results of these experiments as well as those in Experiment 1 were not sensitive to the nature of semantic activation. Such an argument, however, would have great difficulty explaining the ROM effect in Experiment 2, an effect that is clearly a semantically-based effect (as well as having no obvious way to explain the NOM effects in all three experiments).

One alternative argument could be that, given that there often are specific decision-making strategies in lexical decision tasks, that there was some strategy in play that masked the existence of a ROM effect in Experiments 1 and 3. In particular, because we had twice as many ambiguous words as unambiguous words in the Katakana word trials, there may have been a bias to make positive decisions whenever multiple meanings/senses are activated. If so, the different speeds of semantic coding depending on word types could have been masked in Experiments 1 and 3. However, we think that such a decision strategy is unlikely because a) it is unclear why such a strategy wouldn't have also been employed in Experiment 2, and b) an ambiguity advantage has been observed in a number of studies when equal numbers of ambiguous and unambiguous words have been used in lexical decision tasks (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino et al., 2002; Hino et al., 1998; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas et al., 1988; Pexman & Lupker, 1999; Rubenstein et al., 1970; Rubenstein et al., 1971).

What Is the Impact of Using Pseudohomophones?

Given the present results, as well as those of Hino et al. (2006), Klein and Murphy (2001, 2002), and Pexman et al. (2004), it appears to be somewhat difficult to take the ROM effects reported by Azuma and Van Orden (1997), Beretta et al. (2005), Klepousniotou and Baum (2007), and Rodd et al. (2002) as good evidence that there are different speeds of semantic activation for homonyms, polysemous words and unambiguous words. In addition, the question must be asked as to whether the seeming dependence of those effects upon using pseudohomophones implies that these types of effects actually are due to participants engaging in deeper semantically-based processing in a lexical-decision task.

In fact, an obvious question is, why would pseudohomophones have that specific impact in the first place? Although pseudohomophones do not have lexical representations, they do have the ability to activate semantics to a much greater extent than standard nonwords do. Seeing the nonword KAT inevitably brings to mind (feline-based) semantic information whereas seeing KAG does not. Thus, if participants were increasing their reliance on semantic information in the face of pseudohomophones, what they would be doing is making the task of discriminating words from nonwords more difficult for themselves (see also Yap, Tse, & Balota, 2009).

Empirically, what using pseudohomophones sometimes does is to increase effect sizes. For example, as shown by Stone and Van Orden (1993), pseudohomophones increase the size of frequency effects in lexical decision tasks. Pexman and Lupker (1999) argued that pseudohomophones also simultaneously increase both the size the ambiguity advantage and the size of the homophony disadvantage in a lexical-decision task. If the impact of using pseudohomophones is to bias participants to engage in a deeper semantic processing, why would frequency effects (presumably a lexical rather than a semantic factor) and homophone effects (presumably a phonological rather than a semantic factor) increase in size? Given these facts, what Pexman and Lupker proposed is that the impact of using pseudohomophones is to require more extensive processing at the orthographic/lexical level, the level that is most involved in making lexical decisions in general and the level at which pseudohomophones clearly do differ from real words. Hence, any effects at that level or any effects due to semantic or phonological feedback to that level would be increased, resulting in an increased frequency effect, an increased ambiguity advantage and an increased homophony disadvantage.

If it is the case that pseudohomophones do bias participants to engage in more extensive orthographic/lexical processing (e.g., Pexman & Lupker, 1999), there would seem to be no obvious reason why their use would create ROM effects in lexical decision tasks using English stimuli (e.g., Azuma & Van Orden, 1997; Beretta et al., 2005; Rodd et al., 2002). Would it, then, be possible to explain those ROM effects in terms of some sort of strategy like that suggested here to explain the results in Experiment 2? Unfortunately, it's not clear that it would be. While using relatedness amongst senses/meanings might be a useful cue in the present Experiment 2, leading to the ROM advantage, it's not clear why using pseudohomophones in an English lexical-decision task would accomplish the same purpose. At this point, therefore, it doesn't appear that one can conclusively determine why Azuma and Van Orden (1997) and Rodd et al. (2002) got evidence for a ROM effect only when they used pseudohomophones. However, when thinking about this question, the following fact should also be kept in mind. In both Rodd et al.'s and Klepousniotou and Baum's (2007) studies, the ROM effect was weaker in their visual lexical-decision task (even with pseudohomophones in Rodd et al., 2002) than in their auditory lexical-decision task, a task in which pseudohomophones do not exist (i.e., if it sounds like a word, it is a word). Thus, it is reasonably clear that using pseudohomophones is not a prerequisite for getting a ROM effect.

Also worth noting is that Beretta et al. (2005) conducted MEG recordings in their lexical decision experiment, which led them to the conclusion that both the polysemy advantage and homonymy disadvantage in the lexical-decision task were due to different speeds of semantic activation for homonyms, polysemous words, and unambiguous words. Using Rodd et al.'s (2002) stimuli, Beretta et al. replicated the polysemy advantage and the homonymy disadvantage on lexical decision latencies in their task. At the same time, Beretta et al. found that the results of M350 latencies (the neuromagnetic evoked component in the left temporal cortex peaking at 300–420 ms after the visual presentation of stimuli) were similar to those of the lexical decision latencies. That is, although the effects were marginal, Beretta et al. observed a polysemy advantage and a homonymy disadvantage in their analysis of M350 latencies. Following Pylkkanen, Stringfellow, and Marantz (2002), Beretta et al. assumed that M350 latencies in MEG recordings are sensitive only to early processing and, hence, concluded that the speed of semantic activation is modulated by both NOM and ROM in a way consistent with the proposals of PDP models, providing some support for those models.

Overall, however, given the preponderance of the evidence, including the evidence from Hino et al. (2006), Pexman et al. (2004), and Siakaluk, Pexman, Sears, and Owen (2007) concerning the impact of decision-making processes in terms of producing ambiguity effects, there appears to be little support for a PDP-type account of ambiguity effects like that proposed by Rodd et al. (2002). Following Klein and Murphy (2001, 2002), it seems more likely that different meanings and senses are both represented by distinctive sets of features or units at the semantic level; thus, the speed of semantic activation is not actually affected by ROM, which is why Hino et al. (2006) and Pexman et al. (2004) failed to find a ROM effect in their semantic tasks. Hino et al.'s and Pexman et al.'s failure to find NOM effects in semantic tasks also suggests that the basic assumption of competition during the semantic activation process is also probably incorrect. In lexical decision tasks, of course, NOM effects (i.e., advantages) do arise. While the locus of those effects is not yet pinned down, a reasonable possibility is that the NOM advantage is due to those words that have more meanings and/or more senses producing more semantic activation and, hence, stronger semantic feedback, which facilitates processing at the orthographic/lexical level (e.g., Hino et al., 2002; Hino et al., 2006; Lupker, 2007).

Everything considered, one must acknowledge that the locus of any ROM effect is, unfortunately, still far from being determined. An important message offered by the present research, however, is that ROM effects could arise at the decision-making stage due to a specific strategy employed by participants in order to accomplish a task. As such, all the ROM effects reported in the literature as well as those observed in future studies must be carefully interpreted in order to reach a reasonable conclusion regarding the locus of those effects.

Résumé

Les effets du nombre de significations (NS) et du lien entre ces significations (LS) ont été examinés pour les mots japonais katakana à l'aide d'une tâche de décision lexicale. Dans l'Expérience 1, seul un avantage NS a été observé. Dans l'Expérience 2, les mêmes mots katakana ont conduit à un avantage LS lorsque des mots kanjis et des non-mots ont été ajoutés. Puisque les non-mots kanjis étaient composés de caractères indépendants alors que les mots kanjis étaient formés de caractères reliés, les participants pourraient avoir utilisé le lien des significations activées comme indice pour prendre des décisions lexicales dans cette expérience, créant artificiellement un avantage LS pour les mots katakana. Conformément à cette explication, aucun effet LS n'a été observé pour les mots katakana dans l'Expérience 3 lorsque les non-mots kanjis étaient composés de caractères ayant des significations similaires (c'est-à-dire, reliées). Ces résultats représentent un obstacle additionnel pour la position selon laquelle la vitesse de codage sémantique est modulée par les LS pour les mots ambigus.

Mots-clés : effet du nombre de significations, effet du lien entre significations, tâche de décision lexicale

References

Azuma, T., & Van Orden, G. C. (1997). Why SAFE is better than FAST: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language, 36,* 484–504.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 340–357.

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 336–345.

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research, 24,* 57–65.

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 201–219). Hillsdale, NJ: Erlbaum.

Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experiment Psychology: Learning, Memory, and Cognition, 22,* 63–85.

Bowers, J. S., Davis, C. J., & Hanley, D. (2005). Automatic semantic activation of embedded words: Is there a "hat" in "that"? *Journal of Memory and Language, 52,* 131–143.

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition, 4,* 53–61.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113,* 256–281.

Gottlob, L. R., Goldinger, S. D., Stone, G. O., & Van Orden, G. C. (1999). Reading homographs: Orthographic, phonologic, and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 561–574.

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

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

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

Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language, 55,* 247–273.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13,* 278–305.

Jastrzembski, J. E., & Stanners, R. F. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior, 14,* 534–537.

Joordens, S., & Besner, D. (1994). When banking on meaning is not (yet) money in the bank: Explorations in connectionist modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1051–1062.

Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language, 32,* 474–516.

Kawamoto, A. H., Farrar, W. T., & Kello, C. (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1233–1247.

Kawamoto, A. H., & Zemblidge, J. (1992). Pronunciation of homographs. *Journal of Memory and Language, 31,* 349–374.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 601–609.

Klein, D. E., & Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language, 45,* 259–282.

Klein, D. E., & Murphy, G. L. (2002). Paper has been my ruin: Conceptual relations of polysemous senses. *Journal of Memory and Language, 47,* 548–570.

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics, 20,* 1–24.

Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses: The comprehension of polysemy depends of sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1534–1543.

Lupker, S. J. (2007). Representation and processing of lexically ambiguous words. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 159–174). New York, NY: Oxford University Press.

McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27,* 545–559.

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition, 17,* 141–147.

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

National Language Research Institute. (1993). *Bunrui Goi Hyou: Floppy Disk version* [Thesaurus: Floppy Disk version]. Tokyo, Japan: Shuei Shuppan.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Parks, R., Ray, J., & Bland, S. (1998). *Wordsmyth English dictionary – thesaurus*. Retrieved from http://www.wordsmyth.net/

Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic ambiguity and the process of generating meaning from print. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1252–1270.

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

Piercey, C. D., & Joordens, S. (2000). Turning an advantage into a disadvantage: Ambiguity effects in lexical decision versus reading tasks. *Memory & Cognition, 28,* 657–666.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes, 12,* 765–805.

Plaut, D. C., & McClelland, J. L. (1993). Generalization with componential attractors: Word and nonword reading in an attractor network. *Proceedings of the 15th Annual Conference of the Cognitive Science Society* (pp. 824–829). Hillsdale, NJ: Erlbaum.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Pylkkanen, L., Stringfellow, A., & Marantz, A. (2002). Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density, *Brain and Language, 81,* 666–678.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46,* 245–266.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science, 28,* 89–104.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9,* 487–494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645–657.

Rueckl, J. G. (1995). Ambiguity and connectionist networks: Still settling into a solution. Commentary on Joordens and Besner (1994). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 501–508.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 521–568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404.

Siakaluk, P. D., Pexman, P. M., Sears, C. R., & Owen, W. J. (2007). Multiple meanings are not necessarily a disadvantage in semantic processing: Evidence from homophone effects in semantic categorization. *Language and Cognitive Processes, 22,* 453–467.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 744–774.

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. M. H. Assink & D. Sandra (Eds.), *Reading complex words: Cross-language studies* (pp. 113–137). Amsterdam, The Netherlands: Kluwer.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology, 57*A, 745–765.

Taft, M., Zhu, X., & Peng, D. (1999). Positional specificity of radicals in Chinese character recognition. *Journal of Memory and Language, 40,* 498–519.

Umesao, T., Kindaichi, H., Sakakura, A., & Hinohara, S. (1995). *Nihongo dai-jiten: Dai2-ban* [The great Japanese dictionary: 2nd ed.]. Tokyo, Japan: Kodansha.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488–522.

Yap, M. J., Tse, C.-S., & Balota, D. A. (2009). Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language, 61,* 305–325.

(*Appendix follows*)

Appendix

Polysemous, Homonymous, and Unambiguous Katakana Words Used in Experiments 1, 2, and 3, Along With Their English Translations

Polysemous Words:

カップ cup, mug/prize cup/ice cream cup/cup for brassiere

カバー cover/compensation/singing other person's song/a book cover

キー key, lock/key, an important point/hint/key in music

サイン signature/signal, motion

スーパー supermarket/super, excellent

セット set, a suite/setting up/set, a match/having one's hair set

ソフト soft, gentle, tender/softball/soft serve ice cream/software

ヒット hit in baseball/hit, a success

ブリッジ bridge (to cross a river)/bridge in wrestling

プレス press, place (one's hand, finger) with force/the press, new, journalism/make smooth with a hot iron

フロント front desk (in a hotel)/front, the foremost part

ポイント the point, the gist/point, a score/point, a place

ボックス box, container/a certain type of dance steps

モード mode in fashion, popularity, trend/type, manner/mood, temper

Homonyms:

コート wearing coat/court in sports/coating with paint

サンド sandwich/place tightly between two other things/sand, tiny grains of crushed rocks

スチール steel/still (picture)/base steal (in baseball)

セーブ saving, economization/save, record data/save, help/save in baseball/Seibu, a name of a company

ソース sauce/source

チェック cheque, investigate/cheque, a pattern made of squares

パンク a flat tyre/punk rock

ベース base, basis/base in baseball/bass (guitar)

ボーリング bowling/boring (to dig)

リーチ reach in mahjong/reach, length of the arm/reach, a name of a toothbrush

リード in the lead, have an advantage/lead, take the head/lead, a name of a paper used for cooking/read/lead in baseball

リスト list/wrist

リング a finger ring/ring, a title of a movie/things placed in a circle/a ring for boxing or wrestling

ルート Route, a way to go/square root/connections, contacts

Unambiguous Words:

アイデア idea

ギフト gift

コスト cost

シーズン season

シューズ shoes

タワー tower

ピンク pink

プラン plan

フロア floor

ベンチ bench

ポット pot

モラル moral

ランク rank

リサーチ research

Received December 10, 2008
Accepted March 25, 2010 ■