Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 55 (2006) 247–273
Journal of Memory and Language
www.elsevier.com/locate/jml

Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? ☆

Yasushi Hino [a,*], Penny M. Pexman [b], Stephen J. Lupker [c]

[a] Faculty of Arts and Sciences, Waseda University, 1-24-1 Toyama, Shinjuku-ku, Tokyo 162-8644, Japan
[b] Department of Psychology, University of Calgary, 2500 University Drive N.W. Calgary, Alberta T2N 1N4 Canada
[c] Department of Psychology, The University of Western Ontario, London, Ontario N6A 5C2 Canada

Received 21 December 2005; revision received 13 April 2006
Available online 23 June 2006

Abstract

According to parallel distributed processing (PDP) models of visual word recognition, the speed of semantic coding is modulated by the nature of the orthographic-to-semantic mappings. Consistent with this idea, an ambiguity disadvantage and a relatedness-of-meaning (ROM) advantage have been reported in some word recognition tasks in which semantic processing is presumed to be required for responding. To further evaluate this idea, we examined ambiguity and ROM effects in lexical decision and semantic categorization tasks. In the lexical decision task, there was an ambiguity advantage but no ROM effect. In the semantic categorization tasks, we used various semantic categories and found a processing time disadvantage for ambiguous words with less related meanings when the decisions were relatively difficult, but observed no effect when the decisions were easier. These results suggest that both the ambiguity disadvantage and the ROM advantage in the semantic categorization tasks are due to decision-making, rather than semantic-coding, processes. The implications of these results for issues concerning the nature of semantic coding and semantic representations are discussed.

Keywords: Ambiguity effect; Relatedness-of-meanings effect; Semantic coding; Words' orthographic-to-semantic relationships

Introduction

In reading research, one of the central questions has been: how are meanings retrieved from print, that is, how do readers use orthographic information to activate semantic information? While there are now a number of models of this process, one clear fact is that how the orthographic-to-semantic activation process is characterized is highly dependent on the assumptions made about the nature of the mental representations.

In the present investigation, our focus was on the distributed representation assumption inherent in 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,

☆ This research was supported in part by Waseda University Grant for Special Research Projects 2005B-060. We thank Akiko Uchida, Yoko Atoda, Kayo Matsumoto, Yoko Ikemura, Akane Suganuma, Ayami Ebihara and Yukari Ando for their assistance in the data collection. We would also like to thank David Plaut, Marcus Taft and Steve Joordens for their comments on earlier versions of the manuscript.
* Corresponding author. Fax: +11 81 3 5286 3759.
E-mail address: hino@waseda.jp (Y. Hino).

0749-596X/$ - see front matter
doi:10.1016/j.jml.2006.04.001

1990). According to these types of models, orthographic and semantic information are assumed to be represented by patterns of activation over sets of units representing orthographic and semantic features. These units are interconnected and, through a learning process, these connections come to be weighted in a way that reflects the appropriate relationships among units. When a word is presented, a set of orthographic units is first activated and this activation spreads to semantic units through the weighted connections. Because the weights on connections depend on the nature of the relationships between orthographic and semantic units (i.e., their consistency), PDP models predict that the speed and accuracy of semantic coding essentially depend on the nature of the feedforward relationships between orthography and semantics.

As Joordens and Besner (1994; see also Besner and Joordens, 1995) have noted, ambiguous words (i.e., words with multiple meanings) appear to provide a good opportunity to examine these predictions from the PDP models. According to the models, ambiguous words involve the mapping of a single orthographic code onto a number of different semantic codes and, as such, the one-to-many feedforward mappings should produce a cost in terms of the time needed to settle on a particular semantic code. That is, when a particular orthographic pattern is associated with multiple semantic patterns within the same set of semantic units, the model must learn two different associations based on the same orthographic pattern. Further, these two associations must both be represented within the same set of weights, that is, the weights that connect the orthographic and semantic levels. In such a circumstance, the weights on connections that have been adjusted to produce one of the orthographic–semantic associations would be disrupted by learning another association. Hence, it would be difficult to establish the strong orthographic-semantic associations appropriate for any of the meanings of ambiguous words. Consequently, PDP models predict that the speed of semantic coding would be slower for words having multiple meanings than for words having a single meaning.

In fact, in their simulations, Borowsky and Masson (1996) and Kawamoto et al. (1994) reported that the time (i.e., the number of processing cycles) taken to settle on a semantic code was slower and the settling process was more error prone for ambiguous words than for unambiguous words in the PDP models they examined. In addition, Joordens and Besner (1994) reported that there was a "blend state" problem in their simulations with ambiguous words. That is, although they observed faster settling time for ambiguous words than for unambiguous words when the model settled accurately, most of the time, their model failed to settle into a correct semantic pattern for ambiguous words. Instead, semantic units settled into a blend state in which multiple meanings were each partially activated. The obvious question, therefore, is what does the relevant literature tell us about how humans process ambiguous words?

Ambiguity effects

One result that appears consistently in the literature is that ambiguous words have a processing advantage in lexical decision tasks. In these tasks, response latencies are typically faster for ambiguous words (e.g., BANK, LEAN) than for unambiguous words (e.g., FOOD, TENT) (e.g., Borowsky & Masson, 1996; Ferraro & Hansen, 2002; Hino & Lupker, 1996; Hino, Lupker, & Pexman, 2002; Hino, Lupker, & Besner, 1998; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989; Pexman & Lupker, 1999; Rubenstein, Garfield, & Millikan, 1970, 1971, although see Forster & Bednall, 1976; Gernsbacher, 1984). A second result, which will be more central to the present discussion, is the report of an ambiguity disadvantage in semantic tasks using the same set of items that produce an ambiguity advantage in lexical decision. These semantic tasks include the semantic categorization task with the Living Thing category in Hino et al. (2002) and the relatedness judgment task in Piercey and Joordens (2000).

As a number of researchers have suggested, the central process in the lexical decision task is the decision-making process, a process that is assumed to be typically based on the nature of the orthographic code (e.g., Balota, 1990; Balota & Chumbley, 1984; Besner, 1983; Besner & McCann, 1987; Hino & Lupker, 1996, 1998, 2000; Hino et al., 2002; McCann, Besner, & Davelaar, 1988; Pexman & Lupker, 1999; Seidenberg & McClelland, 1989). Thus, whether or not ambiguous words take longer to settle at the semantic level may be irrelevant in the lexical decision task (see Borowsky & Masson, 1996; Kawamoto, 1993; Kawamoto et al., 1994; Masson & Borowsky, 1995; Rueckl, 1995, for similar discussions). In contrast, performance in semantic tasks, such as semantic categorization or relatedness judgment tasks, is clearly dependent on semantic processing and, in particular, on how easy it is to complete semantic coding. Thus, the ambiguity disadvantage observed in semantic tasks appears to provide important support for PDP-based accounts.

Consider, for example, Hino et al.'s (2002) PDP-based account of the impact of semantics in these tasks. As in Balota, Ferraro, and Connor's (1991) interactive-activation proposal, the impact of semantics in lexical-decision making is assumed to come about through feedback activation from the semantic level to the orthographic level. Because ambiguous words would activate multiple meanings, the amount of feedback activation from the semantic level to the orthographic level would be greater for ambiguous words than for unambiguous words. As a result, the orthographic processing required in making lexical decisions would receive more support from semantic feedback for ambiguous words, which would produce the typical ambiguity advantage in lexical decision. Thus, looked at from this perspective, an ambiguity advantage in lexical decision tasks would not be at all inconsistent with PDP-based accounts.

Semantic categorization tasks, on the other hand, require participants to determine whether a meaning of a word falls into a specific semantic category. Thus, semantic categorization responses must be based on the results of semantic processing and, hence, should be sensitive to the speed of semantic coding. As such, Hino et al. (2002) suggested that their ambiguity disadvantage in semantic categorization was perfectly compatible with the predictions of PDP models. That is, because the speed of semantic coding is modulated by the nature of orthographic-to-semantic mappings, it would be expected that ambiguous words would suffer a disadvantage due to their one-to-many feedforward mappings.

Based on their observation of an ambiguity advantage (on error rates) in lexical decision and an ambiguity disadvantage in a relatedness judgment task, Piercey and Joordens (2000) also offered a PDP-based account. Those authors suggested that lexical decisions can be made before semantic coding is completed, whereas semantic tasks such as the relatedness judgment task require identification of a meaning for each word (see also Gottlob, Goldinger, Stone, & Van Orden, 1999). Specifically, Piercey and Joordens suggested that lexical decisions are made based mainly on the familiarity of the target word. Ambiguous words generate a high level of familiarity more quickly than unambiguous words due to the fact that ambiguous words have multiple basins of attraction (correct patterns of activation). Because of these multiple basins, early in processing, the activation pattern of an ambiguous word will typically be closer to one of those basins than the activation pattern of an unambiguous word would be to its single basin. Thus, there would be an ambiguity advantage. To make a relatedness judgment, on the other hand, semantic coding would have to be completed because it is necessary to identify the meaning of a word to decide whether it is related to the word it is paired with. Thus, task performance would reflect the speed of completing semantic coding. The semantic activation would easily settle into a learned pattern for unambiguous words, but for ambiguous words, the multiple meanings would create a competition. As a result, an ambiguity disadvantage would emerge.

Further evidence supporting these ideas comes from Duffy, Morris, and Rayner (1988) and Rayner and Duffy (1986). Those authors reported that, in on-line reading tasks, gaze durations were longer for unbiased ambiguous words than for unambiguous words when the words were presented in a neutral sentential context, results that, as noted by Borowsky and Masson (1996), Gottlob et al. (1999) and Piercey and Joordens (2000), support the idea that semantic coding is slower for ambiguous words. Thus, in all of these experiments, experiments in which semantic coding is presumably required, ambiguous words did suffer in contrast to unambiguous words, consistent with the prediction of PDP-based accounts.

Unfortunately, these results stand in contrast to two other recently reported results. First, in contrast to the ambiguity disadvantage in Hino et al.'s (2002) semantic categorization task using the Living Thing category, Forster (1999) reported no effect of ambiguity in his semantic categorization task using the Animal category. Second, using a relatedness judgment task, Pexman, Hino, and Lupker (2004) reported that while there was a clear ambiguity disadvantage for positive pairs (e.g., MONEY–BANK), replicating Gottlob et al. (1999) and Piercey and Joordens (2000), there was no effect of ambiguity for negative pairs (e.g., HORSE–BANK).

As a result, Pexman et al. (2004) suggested an alternative account of the ambiguity disadvantage observed in the relatedness judgment task, an account based on decision-making, rather than semantic-coding, processes. On the positive trials in relatedness judgment experiments, there is inevitably a response conflict due to there being two meanings of the ambiguous word. That is, one meaning is consistent with a positive response (i.e., the meaning that is related to the meaning of its paired word) and the other(s) is(are) consistent with a negative response (i.e., unrelated to the meaning of its paired word). Thus, an ambiguity disadvantage in this circumstance could be due either to more difficult semantic coding for ambiguous words or to a decision-making conflict for ambiguous words. In contrast, on negative trials in the relatedness judgment task, there would be no decision-making conflict. All meanings of the ambiguous word would be consistent with a negative response. Thus, an ambiguity disadvantage in this circumstance could be unequivocally attributed to slower semantic coding for ambiguous words. The lack of an ambiguity disadvantage for these negative pairs, therefore, suggests that the disadvantage for positive pairs is due to the decision-making process. (A similar argument can be made to explain the ambiguity disadvantage in on-line reading tasks, tasks in which the reader must select the intended meaning of the ambiguous word in order to understand the passage, e.g., Duffy et al., 1988; Rayner & Duffy, 1986.) As such, in contrast to the explanation provided by PDP-based accounts, these results suggest that the speed of semantic coding does not depend on the nature of words' orthographic-to-semantic relationships.

The only results that are not consistent with this analysis are those of Hino et al. (2002). In their semantic categorization task, participants were asked to decide whether or not a presented word falls into the Living Thing category. The important data come from their negative trials, trials in which both their ambiguous and unambiguous words possessed no Living Thing meanings. To establish that none of their ambiguous words had Living Thing meanings even among their subordinate meanings, Hino et al., consulted an unabridged Japanese dictionary (Umesao, Kindaichi, Sakakura, & Hinohara, 1995). In addition, they asked a group of participants to rate whether each of these ambiguous words consisted of all Nonliving Thing meanings (0), both Living and Nonliving Thing meanings (1), or all Living Thing meanings (2). Both sources suggested that their ambiguous words most likely did not involve Living Thing meanings. Thus, there would be no reason to expect a decision-making conflict to arise for the ambiguous words in their semantic categorization task. As such, the ambiguity disadvantage in Hino et al.'s task does imply that the speed of semantic coding is slower for ambiguous words than for unambiguous words.

As noted above, however, Hino et al.'s (2002) conclusion is problematic because Pexman et al. (2004) failed to observe an ambiguity effect in the negative trials of their relatedness judgment task. Further, Hino et al.'s results itself may be problematic due to the fact that Forster (1999) failed to observe an ambiguity disadvantage using a semantic categorization task similar to that of Hino et al. Both Forster's null effect and Hino et al.'s ambiguity disadvantage were obtained on negative trials. However, as will be discussed below, there were a couple of differences between the two experiments that could, potentially, account for the different results. One goal of the present investigation was to examine these differences in an attempt to determine when ambiguity disadvantages do emerge in semantic categorization tasks. In addition, we also examined the impact of another factor that may be important when considering the processing of ambiguous words, the relatedness of the various meanings.

Relatedness-of-meanings (ROM) effects

The question of whether the relatedness of the various meanings of ambiguous words has processing implications has only recently started to attract some attention in the literature (e.g., Azuma & Van Orden, 1997; Klein & Murphy, 2001, 2002; Locker, Simpson, & Yates, 2003; Rodd, Gaskell, & Marslen-Wilson, 2002; although see Jastrzembski, 1981). Words like "BANK" which have at least two completely unrelated meanings (e.g., a place where you keep your money, the edge of the river) are examples of ambiguous words with unrelated meanings while words like "PAPER" which have many related meanings (e.g., an object to write on, a manuscript printed on that object) are examples of ambiguous words with related meanings.[1]

Within a PDP framework, a reasonable assumption would be that the semantic representations for related meanings share semantic features. If so, the orthographic-to-semantic mappings would be more consistent (i.e., those mappings may produce less competition during the settling process) for ambiguous words with related meanings than for ambiguous words with unrelated meanings. As a result, if the speed of semantic coding is modulated by the nature of orthographic-to-semantic mappings, semantic coding should be faster for ambiguous words with related meanings than for ambiguous words with unrelated meanings, producing a relatedness advantage in semantically-based tasks. The further implication, of course, is that any ambiguity disadvantage (in a comparison to unambiguous words) should be larger for ambiguous words with unrelated meanings than for ambiguous words with related meanings.

The implications of ROM for the lexical decision task, however, are somewhat less clear. If one had to rely on semantic coding to make lexical decisions, a relatedness advantage would be expected (see Azuma & Van Orden, 1997). If one assumed that lexical decisions are made based on orthographic codes, however, there is no obvious prediction to be made. Assuming that lexical decision performance is affected by semantic feedback to the orthographic level, Locker et al. (2003) have suggested that the amount of semantic feedback could be modulated by the processing competition at the semantic level. Because stronger competition is expected at the semantic level for ambiguous words with unrelated meanings than for ambiguous words with related meanings, less semantic feedback might be expected for ambiguous words with unrelated meanings. If so, a relatedness advantage would be expected.

On the other hand, very similar accounts (e.g., Hino et al., 2002; Pexman et al., 2004) could also make the opposite prediction. That is, if related meanings do share semantic features, the amount of semantic activation may actually be greater for ambiguous words with unrelated meanings than for ambiguous words with related meanings because more semantic units (representing semantic features) would be activated by unrelated

[1] Some researchers have suggested that the relatedness of meanings reflects the etymological distinction between homonyms (ambiguous words with unrelated meanings) and polysemous words (ambiguous words with related senses) (e.g., Caramazza & Grober, 1976; Jastrzembski, 1981; Klein & Murphy, 2001, 2002). However, because our ROM manipulation involved a relative classification scheme based on the subjective ratings, instead of using the terms homonyms and polysemous words, we simply used the terms ambiguous words with more related and less related meanings to describe that manipulation in this paper.

meanings. If so, there would be more support from semantic feedback for ambiguous words with unrelated meanings than for ambiguous words with related meanings. As such, if there were any relatedness effect at all in the lexical decision task, one would expect a relatedness disadvantage, rather than a relatedness advantage.

The present research

The main purpose of the present research was to determine both when ambiguity disadvantages emerge in semantic categorization tasks and whether those disadvantages are due to the semantic-coding process. To address these issues, the first step was to examine the discrepancy between the results of Hino et al. (2002), who obtained an ambiguity disadvantage in a semantic categorization task, and Forster (1999), who did not. As noted, both results were obtained from words used on negative trials (i.e., trials on which the targets were not exemplars of the designated category). There were, however, some differences between the two experiments. To begin with, although there is no obvious reason why it would matter, Forster used English words while Hino et al. used Japanese Katakana words. Indeed, it would be surprising if this difference were important because Hino and colleagues have had just as much success producing ambiguity effects with Japanese Katakana words (e.g., Hino et al., 1998, 2002; Hino, Lupker, Sears, & Ogawa, 1998; Pexman et al., 2004) as they and others have had with English words (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas et al., 1988; Millis & Button, 1989; Pexman et al., 2004; Pexman & Lupker, 1999; Rubenstein et al., 1970, Rubenstein, Lewis, & Rubenstein, 1971) in lexical decision experiments. Second, as previously noted, the category used in Forster's experiment (i.e., the Animal category) was much smaller and more well-defined than that used in Hino et al.'s experiment (i.e., the Living Thing category). Thus, there would likely be very different decision demands in the two experiments. If this difference does turn out to be important, it would provide additional support for Pexman et al.'s (2004) decision-making account because the different results would, presumably, be attributable to the different processing demands during the decision-making process.

The stimuli selected for all these experiments were Japanese Katakana words. It is not possible to manipulate relatedness of meanings for unambiguous words, so three word groups were created: ambiguous words with more related meanings, ambiguous words with less related meanings, and unambiguous words. The first experiment was a lexical decision experiment. This experiment had essentially two purposes. The first purpose was to provide a manipulation check. That is, it was felt that it would be useful to demonstrate that our ambiguity manipulation was strong enough to produce the standard ambiguity advantage in lexical decision. The second purpose was to determine whether relatedness of meanings matters for these words in this task. The second and third experiments were attempts to replicate Hino et al.'s (2002) and Forster's (1999) results using semantic categorization tasks similar to the ones that they used. That is, we conducted two semantic categorization experiments using the same words in the negative trials, one using a broader semantic category (i.e., Living Things) and the other using a narrower category (i.e., Vegetables). Subsequent experiments also involved the same words in various semantic categorization tasks, tasks that allowed us to vary the decision-making requirements of the task in order to get a better idea of when ambiguity disadvantages emerge in semantic categorization tasks.

Experiment 1

Method

Participants

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

Stimuli

One hundred and twenty ambiguous words with Nonliving Thing meanings, 120 unambiguous words with Nonliving Thing meanings, and 90 unambiguous words with Living Thing meanings were selected based on the first author's intuition from word frequency norms of Katakana-written words ("table of loan words listed in order of their frequencies") in National Language Research Institute (1971). These were all Katakana words between 2 and 6 characters in length. Forty-two undergraduate students from Chukyo University were then asked to rate the experiential familiarity of these words. The 330 words were randomly ordered and listed in a questionnaire. Each word was accompanied by a 7-point scale ranging from *Very Unfamiliar* (1) to *Very Familiar* (7). The participants were asked to rate the experiential familiarity by circling the appropriate number on the scale.

In addition, a new group of 41 students from Chukyo University was asked to rate how typical the meanings of these words are as exemplars of Living Thing or Nonliving Thing categories. These 330 words were, once again, randomly ordered and listed in a questionnaire. In the questionnaire, each word was accompanied by a seven-point scale ranging from *Nonliving* (1) to *Living* (7). The participants were asked to give their typicality ratings by circling the appropriate number on the scale.

Further, for the 240 Katakana words with Nonliving Thing meanings, the number of meanings was evaluated based on subjective ratings. In the number-of-meanings (NOM) ratings, the 240 words were mixed with 90 Katakana-written nonwords, randomly ordered and listed in a questionnaire. A new group of 41 students from Chukyo University was asked to count the number of meanings for each Katakana character string and write down that number.

For comparison purposes, we also collected another NOM measure for the ambiguous words using Azuma and Van Orden's (1997) technique. To accomplish this, a new group of 50 students from Chukyo University was asked to write down all the meanings that they could think of given each of these 120 ambiguous words. Next, those meanings were classified as the same or different meanings based on meanings listed in an unabridged Japanese dictionary (Umesao et al., 1995). Three judges (including the first author) classified these responses. As in Azuma and Van Orden's procedure, when a participant generated multiple responses that are classified as the same meaning in the dictionary, these responses were considered as a single response of that meaning. In addition, when participants' responses were not matched with any of the meanings listed in the dictionary, they were classified as new meanings. After classifying the responses, meanings that were given by more than 5 participants (10%) were counted as meanings of each word. These meaning counts correlated significantly with the NOM ratings collected using the first procedure, $r = .34$, $p < .001$.

To determine whether the 120 ambiguous words (with Nonliving Thing meanings) involve any Living Thing meanings, meaning-type ratings were also collected. A new group of 38 students from Chukyo University was asked to decide whether these ambiguous words consisted of *all Nonliving Thing meanings* (0), *Nonliving and Living Thing meanings* (1), or *all Living Thing meanings* (2). The 120 ambiguous words were randomly ordered and listed in a questionnaire accompanied by a 3-point scale ranging from 0 to 2. The participants were asked to circle the appropriate number on the scale.

To estimate the degree of relatedness of meanings for the 120 ambiguous words, two types of ROM ratings were also collected. First, these 120 ambiguous words were randomly ordered and listed in a questionnaire. In the questionnaire, each word was accompanied by a 7-point scale ranging from *Unrelated* (1) to *Related* (7). A new group of 41 students from Chukyo University was asked to think of all the meanings of each ambiguous word and to rate the relatedness of these meanings by circling the appropriate number on the scale.

Second, we collected relatedness scores in a similar way to that used by Azuma and Van Orden (1997). From the 120 ambiguous words, 90 ambiguous words were selected based on the NOM ratings: the NOM ratings were all more than 1.3 for these 90 ambiguous words. For each of the 90 ambiguous words, all the meanings that were given by at least 5 participants were paired and listed in a questionnaire, accompanied by a 7-point scale from *Unrelated* (1) to *Related* (7). A new group of 32 students from Chukyo University was, then, asked to rate the degree of relatedness for each pair of meanings by circling the appropriate number on the scale. In Azuma and Van Orden's relatedness ratings, relatedness was not evaluated among subordinate meanings. In contrast, we collected relatedness ratings based on all the meanings (that had been given by at least 5 participants). For the 90 ambiguous words, the two types of relatedness ratings were highly correlated with one another, $r = .68$, $p < .001$.

Based on these ratings, 20 ambiguous words with less related meanings, 20 ambiguous words with more related meanings, and 20 unambiguous words were selected. The statistical characteristics of these words are given in Table 1. These words were all between 2 and 5 Katakana characters in length. Word length, $F(2,\ 57) = 0.09$, $MSE = 0.54$, and the number of syllables (morae), $F(2,\ 57) = 0.00$, $MSE = 0.53$, were matched as closely as possible across the three word groups. The frequency counts of these words were all less than or equal to 50 per 940,533. Word frequency counts were also matched as closely as possible across the three word groups, $F(2,\ 57) = 0.44$, $MSE = 121.30$. In addition, experiential familiarity ratings, $F(2,\ 57) = 0.04$, $MSE = 0.60$, and typicality ratings, $F(2,\ 57) = 0.04$, $MSE = 0.23$ were also equated as much as possible. In addition, orthographic neighborhood sizes were calculated for these words (e.g., Coltheart, Davelaar, Jonasson, & Besner, 1977). A word's orthographic neighborhood size is defined as the number of words that can be created by replacing one character in the word. These counts for each word were based on a computer-based dictionary with 36,780 word entries (National Language Research Institute, 1993). The mean orthographic neighborhood sizes were comparable across the three word groups, $F(2,\ 57) = 0.80$, $MSE = 14.11$. In addition, the meaning-type ratings for the two groups of ambiguous words were all less than 0.30 with means of 0.13 for the ambiguous words with less related meanings and 0.12 for the ambiguous words with more related meanings, $t(38) = 0.64$. Thus, it is unlikely that these words involve (subordinate) Living Thing meanings.

The NOM ratings for these ambiguous words were all more than 1.30, whereas the ratings for the unambiguous words were all less than 1.30: the mean ratings were 1.53 for the ambiguous words with more related meanings, 1.57 for the ambiguous words with less related meanings, and 1.09 for the unambiguous words. As such, the NOM ratings were comparable for the two groups of ambiguous words, $t(38) = 0.96$, whereas these

Table 1
Mean word frequency (freq), word length (len), syllabic length (Syl), orthographic neighborhood size (N), experiential familiarity rating (FAM), number-of-meaning rating (NOM), typicality (living thing) rating (LIV), meaning-type rating (M), relatedness rating based on Items (Rel-A), relatedness rating based on meaning pairs (Rel-B), and meaning count based on 50 subjects' responses (MC) for each word group used in all the experiments

Word type	Freq	Len	Syl	N	FAM	NOM	LIV
Ambig./More Rel.	10.85	3.40	3.40	3.80	4.84	1.53	2.16
Ambig./Less Rel.	7.80	3.50	3.40	4.50	4.78	1.57	2.19
Unambig.	8.35	3.45	3.40	3.00	4.84	1.09	2.15

Word type	M	Rel-A	Rel-B	MC
Ambig./More Rel.	0.12	4.23	4.27	3.15
Ambig./Less Rel.	0.13	2.47	2.30	2.85
Unambig.	—	—	—	—

ratings for the unambiguous words were significantly smaller than those for the ambiguous words with more related meanings, $t(38) = 11.52$, $p < .001$, and for the ambiguous words with less related meanings, $t(38) = 10.82$, $p < .001$. The meaning counts based on the 50 participants' responses were all more than 2 for these ambiguous words, ranging from 2 to 6. The mean meaning counts were 2.85 for the ambiguous words with less related meanings and 3.15 for the ambiguous words with more related meanings, $t(38) = 0.96$.[2]

The ROM ratings based on items were all more than 3.20 with a mean of 4.23 for the ambiguous words with more related meanings, whereas these ratings were all less than 3.20 with a mean of 2.47 for the ambiguous words with less related meanings, $t(38) = 11.31$, $p < .001$. In addition, the ROM ratings based on meaning pairs were all more than 3.20 with a mean of 4.27 for the ambiguous words with more related meanings, whereas these ratings were all less than 3.20 with a mean of 2.30 for the ambiguous words with less related meanings, $t(38) = 9.08$, $p < .001$. The words in the three experimental word groups and their English translations are listed in AppendixA.[3]

In addition to these word stimuli, 60 Katakana nonwords, which were created by replacing one character from actual Katakana words, were added to the stimulus set used in the lexical decision task of Experiment 1. The mean character length of these nonwords was 3.40, ranging from 2 to 5. The mean syllabic length of these nonwords was 3.33, ranging from 2 to 5. The mean NOM rating for these nonwords was 0.00. It should be noted that none of these nonwords were pseudohomophones (i.e., nonwords that are pronounced like words). Due to the nature of Katakana, pseudohomophones can't actually exist in that script because any Katakana character string that is pronounced like a word would be regarded as a word by Japanese readers.

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 (FlexScan 54T, EIZO NANAO Corporation) by pressing either a "Word" or a "Nonword" key on a response box connected to a computer. The "Word" response was made using the participant's dominant hand.

Participants were also instructed that their responses should be made as quickly and as accurately as possible. Eighteen practice trials were given prior to the 90 experimental trials. The practice items were 8 Katakana words and 8 Katakana nonwords. None of these items were used in the experimental trials. During the practice

[2] Although we collected meaning counts by asking 50 participants to write down all the meanings that they could think of for each word, our ambiguity manipulation was based on the NOM ratings. We used NOM ratings because a significant ambiguity effect has been repeatedly reported in lexical decision tasks using this ambiguity measure (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino et al., 1998, 2002; Kellas et al., 1988; Millis & Button, 1989; Pexman & Lupker, 1999). In addition, as has been argued by Millis and Button, this measure appears to correctly reflect the number of meanings that each participant can access in memory. If, for example, we had, instead, counted the total number of meanings generated by 50 participants, this measure may not have reflected the number of meanings that each participant could access in memory because all meanings are not generated by all the participants. In contrast, the NOM ratings do reflect the average number of meanings that each participant can access in memory.

[3] We classified the ambiguous words into the more related versus less related meaning categories based on the two types of ROM ratings. Because the relatedness of meanings should be a matter of degree, our ROM manipulation was based on a relative classification depending on the cutoff criteria that we set for these ratings (i.e., 3.20 for the ROM ratings based on items and 3.20 for the ROM ratings based on meaning pairs).

trials, participants were informed about their response 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 center of the video monitor. One second after the onset of the fixation point, a 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 responded to the stimulus by pressing one of two keys on the response box. The participant's response terminated the presentation of the stimulus and the fixation point. The response latency from the onset of the stimulus to the participant's key press and whether the response was correct were automatically recorded by the computer on each trial. 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 cell mean in each condition. Seventy-six data points (2.44%) were classified as outliers and were excluded from the statistical analyses. In addition, there were 87 error responses (2.79%). These trials 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 2.

In the analyses of lexical decision latencies, the main effect of Word Type was significant both in the subjects' analysis, $F_s(2, 50) = 13.61$, $MSE = 308.70$, $p < .001$, and in the items' analysis, $F_i(2, 57) = 3.24$, $MSE = 1169.88$, $p < .05$ ($minF'(2, 84) = 2.62$, $p < .10$; Clark, 1973).

Planned comparisons further revealed that lexical decision latencies were faster for the two types of ambiguous words than for unambiguous words. That is, the ambiguity advantage for ambiguous words with more related meanings was significant in both analyses, $t_s(25) = 5.11$, $p < .001$; $t_i(38) = 2.25$, $p < .05$. The ambiguity advantage for ambiguous words with less related meanings was significant in the subjects' analysis, $t_s(25) = 3.89$, $p < .001$, although not in the items' analysis, $t_i(38) = 1.52$. Lexical decision latencies were comparable for the two types of ambiguous words, $t_s(25) = 1.39$; $t_i(38) = 1.09$.

In the analyses of error rates, the main effect of Word Type was significant in the subjects' analysis, $F_s(2, 50) = 3.18$, $MSE = 12.05$, $p < .05$, although not in the items' analysis, $F_i(2, 57) = 1.34$, $MSE = 21.11$ ($minF'(2, 103) = .94$, ns; Clark, 1973) . The only significant difference detected in the planned comparisons for error rates was the ambiguity advantage for ambiguous words with less related meanings (relative to unambiguous words) in the subjects' analysis, $t_s(25) = 2.28$, $p < .05$, although this effect was not significant in the items' analysis, $t_i(38) = 1.40$.

Discussion

The results of Experiment 1 are quite clear. Equivalent ambiguity advantages were observed for both ambiguous words with less related meanings and for ambiguous words with more related meanings. The existence of an ambiguity advantage shows that our selected Katakana words produce the same effects typically observed for English words in a lexical decision task.

Azuma and Van Orden (1997) and Rodd et al. (2002) have suggested that lexical decision performance is affected by the relatedness of the meanings of ambiguous words when semantics are activated and used in making the decision (which, according to those authors, occurs when the nonwords are pseudohomophones). Although the lack of a relatedness advantage in Experiment 1 provides no support for that suggestion, we should note again that our nonword stimuli were not pseudohomophones because pseudohomophones don't exist when stimuli are written in Japanese Kana scripts. Hence, our lexical decision task may not have been optimal for producing relatedness effects. In all of the following experiments, we used semantic categorization tasks, tasks that clearly do require activation of semantic information, in our examination of NOM and ROM effects.

Table 2
Mean response latencies in millisecond and error rate in percent for each stimulus groups in Experiment 1(lexical decision task)

Word type	RT	ER	Ambiguity effect		Relatedness effect	
			RT	ER	RT	ER
Ambig./More Rel.	538 (12.55)	2.56 (0.72)	+25*	+1.74	+8	−0.60
Ambig./Less Rel.	546 (12.10)	1.96 (0.57)	+17*	+2.34*		
Unambig.	563 (12.51)	4.30 (0.96)				
Nonword	614 (17.05)	3.36 (0.52)				

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.

These tasks will allow us to more directly address the issue of whether the speed of semantic coding is modulated by the nature of the orthographic-to-semantic mappings, including the role that relatedness of meanings may play in that process.

The ambiguity advantage observed in this lexical decision experiment can be explained as being due to semantic feedback. The null relatedness effect, considered in the context of the feedback account, suggests that the two sets of ambiguous words provide equivalent feedback activation to their respective orthographic units. That is, although the semantic organization of these two word types might be somewhat different (e.g., it may be more tightly organized for the words with more related meanings), the semantic richness (e.g., the number of activated semantic units) of and, hence, the amount of feedback from the semantic representations of the two word types appear to have been essentially equivalent.

An additional point to note is that if one assumes that the amount of semantic feedback is modulated by the amount of competition at the semantic level (e.g., Locker et al., 2003), the null relatedness effect would suggest that the amount of competition was equivalent for the two types of ambiguous words. As such, these results would imply that initial semantic activation is not affected by the relatedness of meanings for ambiguous words. We will return to this issue in the General discussion.

Experiments 2 and 3

Experiment 2 was an attempt to replicate Hino et al.'s (2002) results using the identical semantic categorization task that they used (i.e., is it a living thing?). Experiment 3 was a conceptual replication of Forster's (1999) experiment. Because there are many Katakana names of Vegetables in Japanese, we used the Vegetable category instead of the Animal category. Most importantly, although we changed the category, we used the same experimental Katakana words (in the "No" trials) in both experiments to examine whether semantic categorization performance is affected by the type of semantic category chosen for the task (in particular, whether there is a processing time disadvantage for ambiguous words).

Method

Participants

Eighty-seven undergraduate students from Chukyo University participated in these experiments for course credit. All were native Japanese speakers who had normal or corrected-to-normal vision. Forty-seven students participated in Experiment 2 and 40 participated in Experiment 3. None had participated in Experiment 1.

Stimuli

In both experiments, the stimuli for the negative trials were the same 60 Katakana words with Nonliving Thing meanings used in Experiment 1. To create a stimulus set for the positive trials in Experiment 2, 60 (of the original 90) Katakana words with Living Thing meanings were further selected and added to the stimulus set to create the positive trials. The mean word length and syllabic length of these fillers were 3.47 and 3.40, respectively. The "Living Thing" ratings for these fillers (i.e., how typical they are as exemplars of the "Living Thing" category) were all more than 4.00 with a mean of 5.87.

To further evaluate the typicality of the experimental Katakana words as exemplars of the Vegetable category, the 120 ambiguous words with Nonliving Thing meanings and 120 unambiguous words with Nonliving Thing meanings used in the previous typicality ratings (to create the Nonliving Thing stimulus set for Experiments 1 and 2) were mixed with 90 Katakana names of Vegetables and randomly listed in a questionnaire. In the questionnaire, each word was accompanied by a seven-point scale ranging from *Unlikely to be a Vegetable* (1) to *Likely to be a Vegetable* (7). Forty-two additional students from Chukyo University were asked to rate the typicality as an exemplar of the Vegetable category by circling the appropriate number on the scale. The mean "Vegetable" ratings were 1.72 for the 20 ambiguous words with more related meanings, 1.81 for the 20 ambiguous words with less related meanings, and 1.70 for the 20 unambiguous words. These ratings were comparable across the three word groups, $F(2, 57) = .77$, $MSE = .09$.

In addition to the 60 experimental Katakana words, 60 filler Katakana Vegetable names were selected (from the original 90) and added to the stimulus set to create positive trials in Experiment 3. The mean word length and syllabic length of these fillers were 3.45 and 3.40, respectively. The mean "Vegetable" rating for these fillers was 5.84.

Procedure

In Experiment 2, participants were asked to decide whether or not a word appearing on a video monitor (PC-TV455, NEC Corporation) is a name of a Living Thing by pressing either a "Yes" or a "No" key. In Experiment 3, participants were asked to decide whether or not a word appearing on the video monitor is a name of a Vegetable by pressing either a "Yes" or a "No" key. In both experiments, two keys flanking the space key (XFER and NFER keys on a NEC Japanese keyboard) were used as the "Yes" and "No" keys, respectively. The "Yes" response was made using the participant's dominant hand. The response latency from the onset of the stimulus to the participant's key press and whether the response was correct were automatically recorded by a computer (PC-9801FA, NEC Corporation) on each trial. Sixteen practice trials were given prior to the 90 experimental trials in both experiments. The practice trials consisted of 8 positive and 8 negative trials. None of the items in the practice trials were used in the experimental trials. In all other respects, the procedure was identical to that in Experiment 1.

Results

Participants' data were discarded if they had more than 15% error rates. Three participants' data were discarded in Experiment 2. As a result, 44 participants' data were submitted to the statistical analyses in Experiment 2. As in Experiment 1, response latencies were classified as outliers if they were out of the 2.5 *SD* range from the cell mean in each condition. The outliers were excluded from the statistical analyses.

Experiment 2 (Living Thing category)

There were 321 error responses (6.08%). These trials were excluded from the latency analysis. For correct responses, there were 125 outliers (2.37%). They were also excluded from the statistical analyses. Mean response latencies for correct responses and mean error rates were calculated across both subjects and items. The mean response latencies and error rates from the subjects' analysis are presented in Table 3.

In the analyses of response latencies, the main effect of Word Type was significant in both the subjects' and the items' analyses, $F_s(2, 86) = 16.48$, $MSE = 1042.07$, $p < .001$; $F_i(2, 57) = 3.72$, $MSE = 2980.04$, $p < .05$ ($minF'(2, 96) = 3.03$, $p < .10$; Clark, 1973). Planned comparisons revealed that response latencies were slower for ambiguous words with less related meanings than any other word types: the ambiguity disadvantage was significant for ambiguous words with less related meanings, $t_s(43) = 4.98$, $p < .001$; $t_i(38) = 2.18$, $p < .05$. In contrast, response latencies were comparable for the ambiguous words with more related meanings and the unambiguous words, $t_s(43) = 0.55$; $t_i(38) = 0.13$. For the two types of ambiguous words, the relatedness advantage was also significant, $t_s(43) = 4.58$, $p < .001$; $t_i(38) = 2.40$, $p < .025$.

Similarly, in the analyses of error rates, the main effect of Word Type was significant in the subjects' and the items' analyses, $F_s(2, 86) = 37.88$, $MSE = 20.34$, $p < .001$; $F_i(2, 57) = 5.38$, $MSE = 68.95$, $p < .01$ ($minF'(2, 90) = 4.71$, $p < .025$; Clark, 1973). Planned comparisons revealed that more errors were observed for ambiguous words with less related meanings than any other word types: the ambiguity disadvantage was significant for ambiguous words with less related meanings, $t_s(43) = 7.63$, $p < .001$; $t_i(38) = 2.38$, $p < .025$. In contrast, error rates were comparable for the ambiguous words with more related meanings and the unambiguous words, $t_s(43) = 0.01$; $t_i(38) = 0.01$. For the two types of ambiguous words, the relatedness advantage was also significant, $t_s(43) = 6.74$, $p < .001$; $t_i(38) = 2.36$, $p < .025$.

Experiment 3 (Vegetable category)

There were 246 error responses (5.13%). These trials were excluded from the latency analysis. For correct responses, there were 135 outliers (2.81%), which were also excluded from the statistical analyses. Mean response latencies for correct responses and mean error rates were calculated across both subjects and items. The mean response latencies and error rates from the subjects' analysis in this experiment are also presented in Table 3.

Table 3
Mean response latencies in millisecond and error rate in percent for each stimulus group in Experiments 2 (semantic categorization task with Living Thing category) and 3 (semantic categorization task with Vegetable category)

Word type	RT	ER	Ambiguity effect		Relatedness effect	
			RT	ER	RT	ER
Experiment 2 (with Living Thing Category)						
Ambig./More Rel.	732 (12.99)	2.77 (0.62)	+3	−0.01	+36*	+7.25*
Ambig./Less Rel.	768 (15.03)	10.02 (0.90)	−33*	−7.26*		
Unambig.	735 (12.24)	2.76 (0.70)				
Filler	665 (10.04)	7.24 (0.79)				
Experiment 3 (with Vegetable Category)						
Ambig./More Rel.	564 (7.07)	3.20 (0.71)	+2	−1.28	0	−0.03
Ambig./Less Rel.	564 (8.44)	3.17 (0.83)	+2	−1.25		
Unambig.	566 (6.27)	1.92 (0.54)				
Filler	583 (11.48)	7.75 (0.69)				

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 main effect of Word Type was not significant in the analyses of response latencies, $F_s(2, 78) = 0.16$, $MSE = 437.59$; $F_i(2, 57) = 0.01$, $MSE = 1214.28$; ($minF'(2, 96) = .01$, ns; Clark, 1973). The main effect of Word Type was also nonsignificant in the analyses of error rates, $F_s(2, 78) = 1.58$, $MSE = 13.57$; $F_i(2, 57) = 0.42$, $MSE = 33.26$ ($minF'(2, 101) = .33$, ns; Clark, 1973).

Combined analyses for the experimental items in Experiments 2 and 3

To examine the differences between the two experiments, we further conducted combined 3 (Word Type) × 2 (Experiment) ANOVAs on response latencies and error rates for the experimental items.

In the analyses of response latencies, the main effect of Experiment was significant in the subjects' and the items' analyses, $F_s(1, 82) = 144.73$, $MSE = 7652.92$, $p < .001$; $F_i(1, 57) = 655.73$, $MSE = 1501.03$, $p < .001$ ($minF'(1, 111) = 118.56$, $p < .01$; Clark, 1973), reflecting the fact that response latencies were faster in Experiment 3 than in Experiment 2. The main effect of Word Type was significant in the subjects' analysis, $F_s(2, 164) = 10.35$, $MSE = 754.57$, $p < .001$, although not in the items' analysis, $F_i(2, 57) = 1.94$, $MSE = 2693.29$ ($minF'(2, 91) = 1.63$, $p > .10$; Clark, 1973). The interaction between Word Type and Experiment was significant in the subjects' and the items' analyses, $F_s(2, 164) = 11.42$, $MSE = 754.57$, $p < .001$; $F_i(2, 57) = 3.91$, $MSE = 1501.03$, $p < .05$ ($minF'(2, 118) = 2.91$, $p < .10$; Clark, 1973). This interaction reflects the fact that the ambiguity disadvantage was observed for the ambiguous words with less related meanings in Experiment 2 but not in Experiment 3.

In the analyses of error rates, the main effect of Experiment, $F_s(1, 82) = 11.18$, $MSE = 32.91$, $p < .001$; $F_i(1, 57) = 4.99$, $MSE = 33.61$, $p < .05$ ($minF'(1, 134) = 3.45$, $p < .10$; Clark, 1973), the main effect of Word Type, $F_s(2, 164) = 25.70$, $MSE = 17.12$, $p < .001$; $F_i(2, 57) = 3.21$, $MSE = 68.60$, $p < .05$ ($minF'(2, 87) = 2.85$, $p < .10$; Clark, 1973) and the interaction between Experiment and Word Type were significant in both analyses, $F_s(2, 164) = 18.48$, $MSE = 17.12$, $p < .001$; $F_i(2, 57) = 4.90$, $MSE = 33.61$, $p < .025$ ($minF'(2, 119) = 3.87$, $p < .025$; Clark, 1973). The results of the error analysis essentially parallel the results from the latency analysis.

Discussion

When participants were asked to decide whether a word falls into the Living Thing category in Experiment 2, the results were noticeably different from the lexical decision results in Experiment 1. In particular, consistent with Hino et al. (2002), ambiguity created a disadvantage rather than an advantage in Experiment 2. In addition, this disadvantage only existed for the ambiguous words with less related meanings. Ambiguous words with more related meanings produced no ambiguity disadvantage and, hence, a relatedness advantage was observed for the ambiguous words. These results clearly support the idea that not only are the processes involved in making a semantic decision quite different from those involved in making a lexical decision but the role of semantics is as well. This conclusion is, of course, quite consistent with Hino et al.'s (2002) feedback account in which the impact of semantics in lexical decision is due to the feedback-activation process and not the semantic-coding process.

When participants were asked to decide whether a word names a Vegetable in Experiment 3, on the other hand, response latencies and accuracy were quite comparable across the three word groups. Thus, consistent with Forster's (1999) results, it appears that when a narrow semantic category is used, there is no processing time disadvantage for ambiguous words (with either related or unrelated meanings). These results appear to indicate that the nature of the ambiguity effect is modulated by the type of semantic category used in a semantic categorization task.

Before proceeding, however, it is necessary to consider the viability of two potential artifactual explanations of the results of Experiment 2. First, when we were selecting our stimuli for Experiments 1 and 2, we attempted to select ambiguous words for which all their meanings fall outside the Living Thing category. That is, in addition to the typicality ratings, we collected meaning-type ratings to examine whether these ambiguous words involve Living Thing meanings. It is not impossible, however, that our stimulus set involves some ambiguous words for which some aspect of their subordinate meanings might have a Living Thing connotation. If so, activating these meanings could produce a response bias toward an incorrect "Yes" response in Experiment 2 and, hence, semantic categorization responses would be delayed for those words. In particular, if the ambiguous words with less related meanings involved some of these problematic items, the processing time disadvantage for this word type would be explained by the response bias created by these subordinate meanings in the decision-making process.

To address this issue, we asked another group of 30 students from Waseda University to rate the typicality as an exemplar of the Living Thing category for all the meanings of the items used in Experiment 2. In a questionnaire, all the words used in Experiment 2 were paired with their dictionary definitions and listed in a random order. The definitions of ambiguous words were those given by more than 5 participants (10%) when 50 participants were asked to write down all the meanings of those ambiguous words in the previous norming study. Thus, although unambiguous words and fillers were presented only once in the questionnaire, ambiguous words were presented more than once (i.e., once for each of the definitions). Each word and its definition were accompanied by a 7-point scale from 1 (Nonliving) to 7 (Living). The participants were asked to rate the typicality of these word meanings by circling the appropriate number on the scale.

The mean typicality ratings for the dominant meanings were quite comparable for the two types of ambiguous words: 2.10 for the ambiguous words with more related meanings and 2.17 for the ambiguous words with less related meanings, $t(38) = .30$. More importantly, the mean ratings for the subordinate meanings were also quite comparable for the two types of ambiguous words. In particular, when we calculated mean ratings for the subordinate meanings for each ambiguous word, the average ratings were 2.24 for the ambiguous words with more related meanings (ranging from 1.60 to 3.20) and 2.19 for the ambiguous words with less related meanings (ranging from 1.42 to 3.29), $t(38) = .37$. As such, these ratings indicate that the processing time disadvantage for the ambiguous words with less related meanings in Experiment 2 cannot be accounted for in terms of a response bias created by subordinate meanings.[4]

Second, Forster and Hector (2002), Hino, Lupker, and Pexman (2005), Pecher, Zeelenberg, and Wagenmakers (2005), and Rodd (2004) recently reported that semantic categorization performance was affected by meanings possessed by orthographic neighbors of the target stimuli. Thus, it may be possible that the processing time disadvantage observed for the ambiguous words with less related meanings in Experiment 2 was due to the existence of orthographic neighbors with Living Thing meanings.

To address this possibility, we further asked another group of 30 students from Waseda University to rate the typicality as an exemplar of a Living Thing category of all the orthographic neighbors of the experimental words used in Experiment 2. In this questionnaire, all the experimental words and their orthographic neighbors were listed in a random order. Each word was, then, accompanied by a 7-point scale from 1 (Nonliving) to 7 (Living). The participants were asked to rate the typicality of these words by circling the appropriate number on the scale.

Based on these ratings, we classified an orthographic neighbor as Living if its mean rating was more than 4.0. According to this classification scheme, 10 ambiguous words with less related meanings had orthographic neighbors with Living Thing meanings, 7 ambiguous words with more related meanings had neighbors with Living Thing meanings and 5 unambiguous words had neighbors with Living Thing meanings. To further estimate the impact of these neighbors on target processing, we calculated the summed frequencies of the Living and Nonliving Thing neighbors for each item because Hino et al. (2005) reported that the performance in a semantic categorization task with the Living Thing category was strongly correlated with these values (see also Pecher et al., 2005). The summed frequencies of the Living and Nonliving Thing neighbors were 2.25 and 52.80 for the ambiguous words with less related meanings, 3.95 and 56.95 for the ambiguous words with more related meanings, and 6.65 and 33.6 for the unambiguous words. Both the summed frequencies of the Living Thing neighbors, $F(2, 57) = .38$, $MSE = 259.71$, and the summed frequencies of the Nonliving Thing neighbors, $F(2, 57) = .64$, $MSE = 4870.65$, were comparable across the three word groups. Based on these results, it seems unlikely that the ambiguity disadvantage only for the ambiguous words with less related meanings in Experiment 2 could be accounted for in terms of the meanings of the orthographic neighbors.

The significant interaction between Word Type and Experiment in the combined analysis of response latencies in Experiments 2 and 3 clearly indicates that the nature of the ambiguity effect (especially for ambiguous words with less related meanings) is modulated by the type of semantic category used in a semantic categorization task. One possible explanation for this interaction would be to suggest that normal semantic coding was involved in Experiment 2 but not in Experiment 3.

Some researchers have suggested that, in semantic categorization tasks, different decision strategies are implemented depending on the size of the category used (e.g., Forster, 2004; Landauer & Freedman, 1968). In particular, when a small category is used, participants may first attempt to generate all the exemplars of that category. The generated candidates would, then, be compared to the target and participants would respond "Yes" if a match was found. A "No" response would be made if no match was found. As such, normal lexical/semantic processing would not necessarily be required for targets in tasks with small categories, thus, explaining why no word frequency effect is often observed. When a large category is used, on the other hand, it would not be possible to generate all the exemplars in that category. In such a circumstance, participants would have to engage in normal lexical/semantic processing for targets and the decisions would be made after retrieving the meaning of the target. A significant word frequency effect does tend to be observed with large categories.

The results of Experiments 2 and 3 could be explained in terms of this type of idea. In particular, the candidate search strategy may have been used in Experiment 3 (with the small Vegetable category) but

[4] The mean typicality rating for the meanings of the unambiguous words was 2.14, which was quite comparable with the ratings for the dominant and subordinate meanings of the two types of ambiguous words. The mean rating for the meanings of the 60 fillers (i.e., the words requiring a "Yes" response) was 6.07.

not in Experiment 2 (with the large Living Thing category) due to the sizes of the categories. As a result, the ambiguity disadvantage would have only been observed in Experiment 2 in which normal lexical/semantic processing was required for the target stimuli. If so, one could explain the Word Type by Experiment interaction in the combined analysis of response latencies in Experiments 2 and 3 while maintaining the assumption that the processing time disadvantage for the ambiguous words with less related meanings was due to semantic coding.

To further examine this candidate search strategy account, therefore, we conducted an additional semantic categorization task in Experiment 4. Using the same set of experimental items as in Experiments 2 and 3 (in the "No" trials), participants were asked to decide whether a given word is the name of either a Vegetable or an Animal. By combining these two categories, the number of concepts requiring a positive response becomes considerably larger than that in Experiment 3. Thus, a candidate search strategy would be substantially more difficult to employ, making it more likely that normal lexical/semantic processing would be required before making decisions. If so, as in Experiment 2, the ambiguity disadvantage should emerge for the ambiguous words with less related meanings.

Experiment 4

Method

Participants

Twenty undergraduate students from Waseda 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 any of the previous experiments.

Stimuli

Stimuli were the same 60 Katakana words with Nonliving Thing meanings used in the previous experiments. To evaluate the typicality of these words as exemplars of the Animal category, the 120 ambiguous words with Nonliving Thing meanings and 120 unambiguous words with Nonliving Thing meanings used in the previous typicality ratings were once again used. These 240 words were mixed with 90 Katakana names of Animals and randomly listed in a questionnaire. In the questionnaire, each word was accompanied by a seven-point scale ranging from *Unlikely to be an Animal* (1) to *Likely to be an Animal* (7). An additional 42 students from Chukyo University were asked to rate the typicality as an exemplar of the Animal category by circling the appropriate number on the scale. No difference was detected for the "Animal" ratings across the three word groups, $F(2, 57) = 2.16$, $MSE = 0.05$. The mean "Animal" ratings were 1.84 for the ambiguous words with more related meanings, 1.95 for the ambiguous words with less related meanings, and 1.81 for the unambiguous words. Note also that the "Vegetable" ratings (collected in Experiment 3) were also comparable across the three word groups because these words were the same as those used in Experiment 3.

In addition to these 60 experimental words, 60 filler Katakana words were further selected. That is, 30 Katakana Animal names and 30 Katakana Vegetable names were selected and added to the stimulus set to create "Yes (either Animal or Vegetable)" trials in addition to the 60 experimental words (for "No (neither Animal nor Vegetable)" trials). The mean "Animal" rating for the 30 Animal names was 6.67 and the mean "Vegetable" rating for the 30 Vegetable names was 6.20. The mean word length and syllabic length of these 60 fillers were 3.45 and 3.40, respectively.

Procedure

Participants were asked to decide whether or not a word appearing on the video monitor is either a name of a Vegetable or a name of an Animal and to press either a "Yes (either Animal or Vegetable)" or a "No (neither Animal nor Vegetable)" key on a response box connected to the computer. The equipment used in this experiment was the same as in Experiment 1. Eighteen practice trials were presented prior to the 120 experimental trials. The practice items consisted of 5 Katakana names of Animals, 4 Katakana names of Vegetables, and 9 Katakana words with neither Animal nor Vegetable meanings. None of these items were used in the experimental trials. In all other respects, the procedure was identical to that in Experiments 2 and 3.

Results

As before, response latencies were classified as outliers if they were out of the 2.5 *SD* range from the cell mean in each condition. Seventy-six data points (3.17%) were classified as outliers and excluded from the statistical analyses. In addition, there were 57 error responses (2.38%). These trials were excluded from the latency analysis. Mean response latencies for correct responses and mean error rates were calculated across both subjects and items. The mean response latencies and error rates from the subjects' analysis are presented in Table 4.

In the analyses of response latencies, the main effect of Word Type was not significant in either analysis, $F_s(2, 38) = 0.36$, $MSE = 680.69$; $F_i(2, 57) = 0.44$, $MSE = 1369.31$ ($minF'(2, 93) = .20$, ns; Clark, 1973). Similarly, the main effect of Word Type was not significant in the analyses of error rates, $F_s(2, 38) = 1.56$, $MSE = 7.09$; $F_i(2, 57) = 0.58$, $MSE = 26.16$ ($minF'(2, 83) = .42$, ns; Clark, 1973).

Table 4
Mean response latencies in millisecond and percentage errors for each stimulus group in Experiment 4 (semantic categorization task with Animal and Vegetable categories)

Word type	RT	ER	Ambiguity effect		Relatedness effect	
			RT	ER	RT	ER
Ambig./More Rel.	559 (14.27)	1.76 (0.84)	0	+0.03	+6	+1.31
Ambig./Less Rel.	565 (16.68)	3.07 (0.78)	−6	−1.28		
Unambig.	559 (13.39)	1.79 (0.56)				
Filler	542 (15.50)	3.84 (0.64)				

Note. RT and ER stand for mean reaction time and error rate, respectively. Standard error of the mean is in parenthesis ().

Discussion

The size of the target category in Experiment 4 was larger than that in Experiment 3 because both Animal and Vegetable exemplars required a positive response. Therefore, according to the candidate search strategy account, it is more likely that participants would rely on normal lexical/semantic processing for the target words before making decisions in this experiment. Nonetheless, as in Experiment 3, but in contrast to Experiment 2, we failed to observe an ambiguity disadvantage. These results suggest that an account based on the use of the candidate search strategy is probably not a viable account of the results in either Experiment 3 or Experiment 4.

To further evaluate this conclusion, we conducted a multiple regression analysis based on the item means of response latencies for the experimental items (in the negative trials) from Experiment 4. Word length, word frequency, number-of-meanings ratings, typicality ratings for the Animal category, and typicality ratings for the Vegetable category were simultaneously entered as predictor variables. This analysis revealed that word frequency, $beta = -.28$, $t(54) = -2.47$, $p < .025$ and Vegetable typicality ratings, $beta = .46$, $t(54) = 3.60$, $p < .01$, explained significant proportions of the variance in response latencies in the negative trials. The significant frequency effect in this analysis reinforces the conclusion that responding was being based on normal lexical/semantic processing of the target words in Experiment 4.

These same types of analyses provide evidence that the candidate search strategy also was not being employed in Experiments 2 and 3. If participants employed this strategy, response latencies for the "Yes" decisions could be modulated by typicality ratings if typical candidates are more likely to be higher in the search set than atypical candidates. In contrast, the response latencies for the "No" decisions would be solely determined by the category size and, hence, response latencies should be independent of the typicality ratings of the non-exemplar targets. In these analyses, response latencies in the positive and negative trials were separately entered into the multiple regression analyses. Word length, word frequency and typicality ratings (the Living Thing ratings in Experiment 2 and the Vegetable ratings in Experiment 3) were simultaneously entered as predictor variables.

The regression analyses for the response latencies from Experiment 2 revealed that only the Living Thing typicality ratings explain a significant amount of variance in the response latencies on both the positive trials, $beta = -.31$, $t(56) = -2.38$, $p < .025$, and the negative trials, $beta = .38$, $t(56) = 3.15$, $p < .01$. In the analysis of response latencies on the positive trials of Experiment 3, word frequency, $beta = -.24$, $t(56) = -2.37$, $p < .025$, and the Vegetable typicality ratings, $beta = -.59$, $t(56) = -5.82$, $p < .001$, explained significant amounts of variance. In the negative trials of Experiment 3, the Vegetable ratings were the only significant predictor variable, $beta = .39$, $t(56) = 3.10$, $p < .01$. In essence, typicality effects were detected on both the positive and negative trials in both experiments.

Based on these results, it is hard to conclude that participants were using a candidate search strategy in either Experiment 2 or Experiment 3. Thus, it seems unlikely that the lack of an ambiguity disadvantage in Experiments 3 and 4 was due to the use of a candidate search strategy in those experiments. Rather, the significant typicality and frequency effects in the regression analyses appear to suggest that normal lexical/semantic processing was involved and, hence, the meanings were retrieved and evaluated for the target words in both the positive and negative trials in all our semantic categorization experiments.

Assuming that normal lexical/semantic processing is driving responses in all three experiments, the question is: how would it be possible to account for the ambiguity disadvantage observed only in Experiment 2? The most obvious answer, of course, is that the different results in Experiment 2 are likely due to the different decision operations which are presumably required when a broad, ill-defined semantic category is used.

The distinction between the decision-making processes in Experiment 2 versus Experiments 3 and 4 becomes a bit clearer when a couple of additional facts are considered. First, when no ambiguity disadvantage was observed for ambiguous words with less related meanings (in Experiments 3 and 4), the overall response latencies were noticeably shorter than when the ambiguity disadvantage emerged (in Experiment 2). Second, the typicality ratings for the (identical) experimental items in the "No" trials were somewhat higher for the Living Thing category (2.17 on average) than for the Vegetable (1.74 on average) or Animal (1.87 on average) categories. Thus, it's clearly the case that the Living Thing decisions in Experiment 2 were harder to make than the Vegetable decisions in Experiment 3 and the "either Animal or Vegetable" decisions in Experiment 4.

What these results suggest is that the decision-making process must have been much more complicated in Experiment 2 than in Experiments 3 and 4, a conclusion consistent with a logical analysis of the differences between experiments as well. Because both Vegetable and Animal categories are narrow categories with relatively clear boundaries, it would be possible that both Vegetable and Animal decisions could be made by checking only a small set of core semantic features against the semantic information activated by the target word. As a result, this operation would be done reasonably quickly regardless of the number of meanings a word has. On the other hand, because the Living Thing category is a fairly broad category, it is unlikely that checking a small set of semantic features would have led to accurate decision making. Instead, it appears that more analytic semantic processing was involved. In particular, participants may need to evaluate all the activated features of the target word for the likelihood that they would be features of a Living Thing. In the Nonliving Thing trials, ambiguous words, particularly, those with less related meanings, would have more unique features to evaluate and, as the relatively higher ratings indicate, some of those features are at least somewhat consistent with those of Living Things. A reasonable hypothesis, therefore, is that an ambiguity disadvantage will only be observed when the semantic categorization decision involves this type of analytic processing.

To address this proposal, in Experiment 5, we conducted a semantic categorization task with a somewhat different category. The category we selected was the Human category (i.e., is the word the name of a Human position, occupation or group?). The Human category was chosen because there appears to be no well-defined set of features that specifies the category (in contrast to the Animal and Vegetable categories) and, as will be described later, because the typicality ratings for the words used for the "No" trials in the previous experiments were relatively higher (2.44 on average) for the Human category than for the Animal (1.87 on average) or Vegetable categories (1.74 on average). That is, although the experimental words are clearly not exemplars of the Human category, there are some features of these words that are not inconsistent with the concept of Human. Semantic categorization decisions, therefore, should be more difficult and require more analytic processing in the Human decision task in Experiment 5 than in either Experiment 3 or 4. Thus, if our analysis of the necessary circumstances for producing an ambiguity disadvantage for words with less related meanings is correct, this effect should re-emerge in the Human decision task.

Experiment 5

Method

Participants

Fifty-one 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 any of the previous experiments.

Stimuli

Once again, stimuli were the same 60 Katakana words with Nonliving Thing meanings used in the previous experiments. To evaluate the typicality of these words as exemplars of the Human category, typicality ratings for the Human category were collected for these words. As in the previous typicality ratings, the same 120 ambiguous words and 120 unambiguous words were mixed with 90 Katakana names of Human positions, occupations and groups (e.g., WAITER, SINGER, BABY) and listed in a questionnaire. In the questionnaire, these 330 Katakana words were randomly ordered and each word was accompanied by a seven-point scale ranging from *Unlikely to be Human* (1) to *Likely to be Human* (7). Forty students from Chukyo University were asked to rate the typicality of each word as an exemplar of Human category by circling the appropriate number on the scale. The typicality ratings for the Human category was comparable for the three word groups, $F(2, 57) = 0.66$, $MSE = 0.18$. The mean "Human" ratings were 2.37 for the ambiguous words with more related meanings, 2.42 for the ambiguous words with less related meanings, and 2.52 for the unambiguous words.

Note also that the mean typicality rating for the Human category (2.44) was significantly higher than that for the Vegetable category (1.74), $t(59) = 9.22$, $p < .001$, as well as being higher than that for the Living Thing category (2.17), $t(59) = 4.82$, $p < .001$. (The typicality ratings for the Living Thing category was also higher than that for the Vegetable category, $t(59) = 6.67$, $p < .001$.) As such, these ratings suggest that these items involve more features that are not inconsistent with the Human category relative to the Vegetable category and, hence, more analytic processing would be required in making negative decisions for these items in a Human decision task than in a Vegetable decision task.

In addition to these 60 items for the experimental "No" trials, 60 Katakana names of Human positions, occupations, and groups were selected to create "Yes" trials. The typicality ratings of these fillers for the Human category were all more than 4.10, with a mean of 5.74. These fillers were all between 2 and 5 Katakana characters in length. The mean character length and syllabic length of the 60 fillers were 3.50 and 3.42, respectively.

Procedure

Participants were asked to decide whether or not a presented word was a name of a Human position, occupation, or group by pressing one of two keys on the keyboard. The equipment used in this experiment was the same as in Experiments 2 and 3. Two keys flanking the space key were assigned to the "Yes" and "No" responses. Eighteen practice trials were given prior to the 120 experimental trials. None of the items in the practice trials were used in the experimental trials. In all other respects, the procedure was identical to that in Experiments 2, 3, and 4.

Results

Participants' data were discarded if they had more than 15% errors. Thus, 5 participants' data were discarded and 46 participants' data were submitted to the statistical analyses.

As before, response latencies were classified as outliers if they were out of the 2.5 *SD* range from the cell mean in each condition. Thus, 157 outliers (2.84%) were excluded from the statistical analyses. In addition, there were 324 error responses (5.87%) in total. These trials were excluded from the latency analysis. Mean response latencies for correct responses and mean error rates were calculated across both subjects and items. The mean response latencies and error rates from the subjects' analysis in this experiment are presented in Table 5.

In the analyses of response latencies, the main effect of Word Type was significant in the subjects' and the items' analyses, $F_s(2, 90) = 29.68$, $MSE = 450.74$, $p < .001$; $F_i(2, 57) = 3.17$, $MSE = 1864.05$, $p < .05$ ($minF'(2, 85) = 2.86$, $p < .10$; Clark, 1973). Planned comparisons revealed that response latencies were slower for ambiguous words with less related meanings than for any other word types: the ambiguity disadvantage was significant for ambiguous words with less related meanings in the subjects' analysis, $t_s(45) = 6.61$, $p < .001$, and marginally significant in the items' analysis, $t_i(38) = 1.98$, $p < .06$. The relatedness advantage was also significant, $t_s(45) = 7.58$, $p < .001$; $t_i(38) = 2.23$, $p < .05$. In contrast, response latencies were comparable for the ambiguous words with more related meanings and for the unambiguous words, $t_s(45) = 1.28$; $t_i(38) = 0.29$.

In the analyses of error rates, the main effect of Word Type was significant in both the subjects' and the items' analyses, $F_s(2, 90) = 13.82$, $MSE = 14.32$, $p < .001$; $F_i(2, 57) = 4.57$, $MSE = 19.90$, $p < .025$ ($minF'(2, 127) = 3.43$, $p < .05$; Clark, 1973). Planned comparisons revealed that more errors were observed for the two types of ambiguous words: the ambiguity disadvantages were significant for ambiguous words with less related meanings, $t_s(45) = 3.62$, $p < .001$; $t_i(38) = 2.88$, $p < .01$, and for ambiguous words with more related meanings, $t_s(45) = 5.98$, $p < .001$; $t_i(38) = 2.84$, $p < .01$. Error rates were comparable for the two types of ambiguous words, $t_s(45) = 1.67$; $t_i(38) = 0.96$.

Discussion

Based on the typicality ratings for the items used in the negative trials in the present experiment, the expectation was that making decisions, particularly, "No" decisions, would be harder in the semantic categorization task with the Human category than in the semantic categorization tasks with Animal and Vegetable categories. As a result, the semantic categorization task with the Human category would require more analytic processing in making decisions. Consistent with this expectation, the overall response latencies were somewhat longer in Experiment 5 (648 ms) than those in Experiments 3 (569 ms on average) and 4 (556 ms on average), although they were somewhat shorter than those in Experiment 2 (725 ms on average). More importantly, a significant processing time disadvantage for the ambiguous words with less related meanings was observed in this experiment.[5]

Together with the results of Experiments 2, 3, and 4, the present results strongly suggest that the processing time disadvantage for ambiguous words with less related meanings is produced only when the semantic categorization decisions involve more complicated, analytic processing. When the category used in the task is relatively small (as in the Vegetable decision task), participants

[5] One may argue that the responses should be slower in Experiment 4 than in Experiment 3 if more features would have to be checked when two categories were used in the semantic categorization task. Nonetheless, the mean response latencies were numerically shorter in Experiment 4 than in Experiment 3. It is possible that at least some of the participants in Experiment 4 employed a simplified decision strategy. That is, after the experiment, some participants in Experiment 4 reported that they made decisions based mainly on whether it is "bitable" (i.e., one can bite it away with his/her tooth if it is either animal or vegetable). Such a simplified strategy may have produced the faster responses even when two categories were combined in Experiment 4.

Table 5
Mean response latencies in millisecond and percentage errors for each stimulus group in Experiment 5 (semantic categorization task with Human category)

Word type	RT	ER	Ambiguity effect		Relatedness effect	
			RT	ER	RT	ER
Ambig./More Rel.	636 (10.79)	4.56 (0.68)	+7	−4.11*	+32*	−1.59
Ambig./Less Rel.	668 (11.63)	2.97 (0.68)	−25*	−2.52*		
Unambig.	643 (10.29)	0.45 (0.22)				
Filler	643 (12.80)	10.39 (0.69)				

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.

appear to be able to focus their attention on a small set of core semantic features and their decisions can be made based on whether or not these features are activated. This type of decision can be made relatively quickly and the decision times appear to be independent of the number of activated features/meanings. When a broader category is used (as in the Living Thing and Human decision tasks), on the other hand, these categories consist of more complicated combinations of smaller sub-categories, such as combinations based on family resemblances (e.g., Rosch & Mervis, 1975). Consequently, the decisions are more difficult because participants cannot focus their attention on a small set of core semantic features. Instead, participants have to examine all the activated features/meanings to decide whether the presented word falls into the category in question. In such circumstances, the decision times are modulated by the number of activated features/meanings potentially putting ambiguous words at a disadvantage.

In addition, because there would, presumably, be a larger degree of overlap in the semantic representations for ambiguous words with related meanings than for ambiguous words with unrelated meanings, there would be more sets of features/meanings to analyze when ambiguous words have unrelated meanings. For example, when deciding whether PAPER is living, once the decision is made that "a writing material" is not living, there would be no reason to also consider whether "a content written on that material" is living. Thus, the processing time disadvantage, in comparison to unambiguous words, would be minimal for ambiguous words with more related meanings. On the other hand, when deciding whether BANK is living, even if the decision is made that "a financial institution" is not living, this decision would be essentially irrelevant to deciding whether "the edge of a river" is not living. As such, it would take longer to decide whether all the meanings fall into the semantic category in question when the ambiguous words have unrelated meanings. That is, for words with unrelated meanings, essentially all of the meanings would need to be individually analyzed. As a result, an ambiguity disadvantage would be observed but only for the ambiguous words with less related meanings.

The suggestion here, therefore, is that the processing time disadvantage for ambiguous words with less related meanings in semantic categorization tasks is not due to the semantic-coding process, which would be common to all the semantic categorization tasks regardless of the category used, but instead is due entirely to the decision-making process, which would be specific to the nature of the category used. If this suggestion is correct, then our results present a challenge to the view that the ambiguity disadvantage (and the relatedness advantage) in semantically-based tasks is due to the speed of semantic coding being modulated by the nature of the orthographic-to-semantic mappings (e.g., Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Gottlob et al., 1999; Hino et al., 2002; Piercey & Joordens, 2000; Rodd et al., 2002, 2004). This issue will be considered further in the General discussion.

General discussion

As noted by a number of researchers (e.g., Borowsky & Masson, 1996; Gottlob et al., 1999; Kawamoto, 1993; Piercey & Joordens, 2000), PDP models generally predict that the speed of semantic coding is modulated by the consistency of the orthographic-to-semantic mappings. Ambiguous words are assumed to possess one-to-many mappings from orthography to semantics, whereas unambiguous words are assumed to possess one-to-one mappings. Thus, the feedforward relationships from orthography to semantics are more inconsistent for ambiguous words than for unambiguous words. As a result, the speed of semantic coding should be slower for ambiguous words than for unambiguous words. Results supporting this prediction have been reported recently by Hino et al. (2002).

In addition, as suggested by Azuma and Van Orden (1997; see also Locker et al., 2003; Rodd et al., 2002), if the speed of semantic coding is modulated by the nature of the feedforward relationships from orthography to semantics, the speed of semantic coding may also be modulated by the relatedness of the meanings for ambiguous words. That is, if one assumes that related meanings share semantic units, the feedforward mappings from orthography to semantics would be more consistent when the meanings of ambiguous words are related to each other than when ambiguous words involve unrelated meanings. Thus, the speed of semantic coding would be faster for ambiguous words with related meanings than for ambiguous words with unrelated meanings.

Semantic categorization results

Assuming that semantic categorization performance is sensitive to the speed of semantic coding, the specific expectations, according to PDP models, are that: (a) response latencies in a semantic categorization task should be faster for unambiguous words than for the two types of ambiguous words and (b) for the two types of ambiguous words, response latencies should be faster for ambiguous words with more related meanings than for ambiguous words with unrelated meanings. This particular pattern of results, however, was not observed in any of our semantic categorization experiments.

Indeed, the nonexistence of an ambiguity disadvantage for ambiguous words with more related meanings in Experiments 2, 3, 4, and 5 and the lack of an ambiguity disadvantage for ambiguous words with less related meanings in Experiments 3 and 4 pose a major challenge for models based on PDP principles. If semantic categorization performance reflects the speed of semantic coding, we should have been able to observe a consistent ambiguity disadvantage across semantic categorization experiments. Our inability to do so, coupled with Forster's (1999) inability to do so, suggests that we need to reconsider exactly how semantic categorization performance reflects the speed of semantic coding.

Our use of the identical word stimuli in the "No" trials in Experiments 2, 3, 4 and 5 highlights this problem. Because the stimuli are identical, the semantic-coding process (but not the decision process) required in the negative trials should have been essentially the same across these experiments. Although some researchers (e.g., Forster, 2004; Landauer & Freedman, 1968) have suggested the possibility that semantic coding could be bypassed when small semantic categories are used, the significant typicality and frequency effects in the multiple regression analyses in the negative trials of Experiments 2, 3, and 4 appear to eliminate this possibility. Nonetheless, the ambiguity disadvantage for ambiguous words with less related meanings and the relatedness advantage were observed only in Experiments 2 and 5. Thus, it is unlikely that these effects were due to the common semantic-coding process. Instead, the different results in these experiments appear to be due to the different operations required during the task-specific, decision-making process, the process following semantic coding in all our semantic categorization experiments.

To further consider what operations are responsible for producing the ambiguity disadvantage and the relatedness advantage during the decision-making process, we focused on the following facts: while overall response latencies were relatively shorter when no ambiguity disadvantage emerged (in Experiments 3 and 4), a significant ambiguity disadvantage for the ambiguous words with less related meanings was accompanied by relatively longer response latencies (in Experiments 2 and 5). In addition, the ambiguity disadvantage was observed only when the typicality ratings for the experimental items in the negative trials were relatively higher (in spite of the fact that none of the items were members of the category in question). These facts appear to suggest that what is required for producing an ambiguity disadvantage is that the decisions must be somewhat difficult to make. That is, the decisions can't be made by checking small sets of semantic features and, thus, more analytic processing is required. The relatively higher typicality ratings for the items used in the negative trials presumably led to more difficult decision-making in general and, hence, to the longer response latencies that were observed in Experiments 2 and 5 (relative to those in Experiments 3 and 4).

When narrower categories were used in a semantic categorization task, latencies were, presumably, short because the decisions can probably be based on the existence (or nonexistence) of only a small set of target features (i.e., Experiments 3 and 4). Hence, the number of irrelevant features activated by the target word plays little role in the decision-making process. As a result, neither ambiguity nor relatedness of meanings affects the process. On the other hand, when decisions can't be made by looking for a small set of target features, participants would need to evaluate all the activated features (or meanings) of the target word for the likelihood that they belong to the category in question (e.g. the Living Thing category). In this situation, the number of features (or meanings) activated by the target word would matter. In particular, when an ambiguous word involves unrelated meanings, the semantic analysis needs to consider all the meanings and, hence, the decision-making process takes measurably longer. When the ambiguous word has only related meanings, however, the process is a bit simpler because, most of the time, the semantic analysis could essentially be done for all meanings at once. Thus, in comparison to unambiguous words, a measurable ambiguity disadvantage would only arise for ambiguous words with less related meanings, as was observed in Experiments 2 and 5.

Relatedness-of-meanings effects

In contrast to the results of the semantic categorization tasks, we observed a clear ambiguity advantage in the lexical decision task as has been reported many times in the past (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino et al., 1998, 2002; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas et al., 1988; Millis & Button, 1989; Pexman & Lupker, 1999; Rubenstein et al., 1970, 1971). In addition, in this task, there was no sign of a relatedness-of-meanings effect. That is, the size of the ambiguity advantage was similar for the two types of ambiguous words regardless of the relatedness of the meanings.

Based on these results, it is apparent that the lexical decision performance in our task was not sensitive to the nature of semantic coding, that is, the nature of the orthographic-to-semantic mappings possessed by words, as characterized by PDP models. Instead, the ambiguity advantage observed in our task appears to be more easily accounted for as a semantic feedback effect as suggested by Hino et al. (2002). Because ambiguous words possess multiple meanings, the amount of semantic activation would be greater for ambiguous words than for unambiguous words and, hence, the amount of semantic feedback would also be greater for ambiguous words. As a result, the orthographic processing required in making lexical decisions would receive more support from semantic feedback for ambiguous words, producing an ambiguity advantage in lexical decision.

In contrast to the ambiguity effect, we failed to observe a relatedness-of-meanings effect in our lexical decision task. The lack of this effect suggests that at least the initial semantic activation of ambiguous words is not modulated by the relatedness of meanings because if the amount of semantic activation were modulated by relatedness of meanings, the amount of semantic feedback would also have been modulated by this factor (e.g., Locker et al., 2003). We also failed to observe a relatedness effect in the semantic categorization tasks when narrower semantic categories (e.g., the Vegetable category) were used (as in Experiments 3 and 4). In fact, we only observed a significant relatedness effect when a broader semantic category was used in the semantic categorization tasks. Along with our task analysis presented above, these facts appear to suggest that the relatedness-of-meanings effect arises only when more analytic semantic processing is required during the decision-making process (as in Experiments 2 and 5). A reasonable conclusion based on these findings would be that the initial semantic activation is independent of the relatedness of meanings for ambiguous words, a conclusion which is also consistent with the recent findings reported by Klein and Murphy (2001, 2002).

In contrast to our lexical decision results, Azuma and Van Orden (1997) and Rodd et al. (2002) reported a significant relatedness effect in their lexical decision experiments when they used pseudohomophones as nonword stimuli. In Azuma and Van Orden's study, a significant relatedness advantage was observed in their few meaning condition in which ambiguous words possessed less than 5 meanings. Similarly, Rodd et al. reported a processing time advantage for words with related meanings (along with a processing time disadvantage for words with unrelated meanings). However, in neither of these studies was there a relatedness effect when using pronounceable nonwords in a lexical decision task. Thus, our lexical decision results are not technically inconsistent with Azuma and Van Orden's or Rodd et al.'s. In the end, those authors specifically argued that the relatedness effects in their lexical decision experiments (with pseudohomophones) were due to the semantic-coding process playing a large role in the task expressly due to the use of pseudohomophones (although Rodd et al. did observe a relatedness effect in their Experiment 3, an auditory lexical decision task that did not involve pseudohomophones).

An implication of Azuma and Van Orden's (1997) and Rodd et al.'s (2002) position, however, is that there should be a relatedness effect in our semantic categorization task, a task that requires the semantic-coding process. The lack of such an effect provides a very clear challenge to that position. For example, because there were only two ambiguous words having more than 4 meanings in our stimulus set (based on the 50 participants' responses in our norming study), most of our ambiguous words would have belonged in the few meaning condition in Azuma and Van Orden's study. It is precisely these types of words in Azuma and Van Orden's study that produced a relatedness advantage.

One way to attempt to explain this discrepancy would be to argue that our semantic categorization data do not reflect the speed of semantic coding because our data are based on the negative trials. To account for the performance on negative trials in a lexical decision task, a number of researchers have suggested that the negative responses are made based on a time deadline (e.g., Borowsky & Masson, 1996; Coltheart et al., 1977; Grainger & Jacobs, 1996; Rodd et al., 2004). If the negative responses in our semantic categorization tasks were also based on a deadline, their latencies would, indeed, not reflect the speed of semantic coding.

We find this possibility unlikely, however, based on the following facts. First, if the negative responses are made based on a time deadline, the mean response latencies should always be slower on negative trials than on positive trials. Although the negative responses were slower than the positive responses in Experiments 2 and 4, mean response latencies were comparable for the positive and negative trials in Experiment 5 and the negative responses were even faster than the positive responses in Experiment 3. Second, we observed significant typicality effects in our multiple regression analyses of the response latencies on negative trials in

Experiments 2, 3, and 4. Further, in the regression analysis of Experiment 4, we observed a significant word frequency effect in the negative trials. If negative responses were based on a time deadline, there would be no reason to expect typicality and frequency effects in the negative trials. As such, the significant typicality and frequency effects in the negative trials clearly suggest that normal lexical/semantic processing was involved for target words in all of our experiments.

A second way to attempt to explain this discrepancy would be to consider whether it might be possible to apply Rodd et al.'s (2004) account, an account which allows for different results as a function of the nature of the task, to the present situation. This account was proposed to explain the lexical decision data reported in Rodd et al. (2002). It is based on the idea that there are changes in the processing difficulty of the various types of ambiguous and unambiguous words as a function of the degree to which semantic processing is engaged. A changing data pattern among our various word types as the nature of the task changed was, of course, the pattern we observed. Thus, at first glance, this approach might seem to have some promise.

Rodd et al.'s (2004) model assumes that related meanings share semantic features but unrelated meanings do not. When this model was trained using ambiguous words with unrelated meanings, it developed attractor basins for these meanings at different locations in semantic space. As in the other PDP simulations described earlier, due to the competition created by these meanings, the model took longer to settle at the semantic level when words possessed more unrelated meanings. In contrast, because related meanings share semantic features, the model developed one broader attractor basin for multiple related meanings of an ambiguous word. Therefore, the network state could move into the attractor basin more quickly when a word possessed more related meanings. When it is necessary to fully settle on a semantic pattern, however, having multiple related meanings does hurt processing (in comparison to unambiguous words). This situation results because, although related meanings share features, those meanings are not identical and, thus, there is some competition involved in settling on these partially different patterns. The implication is that if a task could be accomplished before completing semantic coding, the model predicts a processing time advantage for ambiguous words having related meanings (in comparison to unambiguous words). On the other hand, if a task requires participants to fully complete semantic coding before responding, the model predicts a processing time disadvantage for ambiguous words having related meanings (in comparison to unambiguous words). In either situation, however, ambiguous words with unrelated meanings should be processed more slowly than either unambiguous words or ambiguous words with related meanings (although the latter difference would be reduced when complete semantic coding is required).

The results of the lexical decision task in Experiment 1, showing faster latencies for ambiguous words with unrelated meanings than for unambiguous words (and no relatedness advantage) are, of course, inconsistent with this model. As noted, however, Rodd et al. (2002, 2004) have argued that unless the nonwords are pseudohomophones, lexical decisions are not contingent on the nature of semantic processing and, hence, the results of Experiment 1 would be irrelevant to their model. The results of the other experiments would not be irrelevant to their model, however. Here the predictions are clear: (1) a relatedness advantage should be most evident when participants can respond before semantic coding is completed and (2) a diminished (or null) relatedness advantage should emerge when semantic coding has to be completed before responding and (3) there should always be an ambiguity disadvantage for ambiguous words with unrelated meanings.

In terms of the processing apparently involved in our categorization experiments, our results and these predictions appear to be quite inconsistent with one another. As discussed, based on the overall latencies and logical considerations, one could argue that, if anything, more complete settling was required in Experiments 2 and 5 (the Living Thing and Human decisions) than in Experiments 3 and 4 (the experiments with the Vegetable and Animal categories). Therefore, according to the model, it is Experiments 3 and 4 in which one would be most likely to find a relatedness advantage. In fact, it is in Experiments 2 and 5 where the advantage actually arose (relative to ambiguous words with less related meanings). And, of course, the prediction that unambiguous words would always have an advantage over ambiguous words with unrelated meanings is contradicted by the results of Experiments 3 and 4. Thus, the general patterns observed in the present experiments do not appear to be explainable within the framework of this model.

Alternative accounts of the present results

Our data, as well as those reported by Pexman et al. (2004) and Forster (1999), suggest that ambiguity disadvantages in semantic categorization and relatedness judgment tasks are likely not due to the semantic-coding process as that process is conceptualized within PDP models. Before considering how those models would need to be changed to allow them to accommodate our data, an obvious question to ask is whether these data could be explained by a semantic-coding process as it might be conceptualized within a lexical model based on localist assumptions (i.e., a model in which there is a lexical level between the orthographic and semantic levels, Becker, 1980; Forster, 1976; Morton, 1969; McClelland & Rumelhart, 1981; Paap, Newsome, McDonald, & Schvaneveldt, 1982).

As noted, if one assumes a lexical representation for each word meaning regardless of relatedness of meanings, the speed of selecting a lexical representation would be probabilistically faster for words with multiple lexical representations (i.e., ambiguous words) than for words with a single lexical representation (i.e., unambiguous words) (e.g., Jastrzembski, 1981; Rubenstein et al., 1970, 1971). As such, this type of model can account for our lexical decision data.

Where the model would have problems, however, is in explaining our semantic categorization data. Presumably, because each meaning has its own lexical entry, once a lexical representation had been selected, the meaning representation attached to that entry could simply be activated regardless of whether those semantic representations are assumed to be localist or distributed (e.g., Taft, 2003, 2004; Taft & Kougious, 2004). Hence, this type of model would not predict an ambiguity disadvantage due to competition during semantic coding. Nonetheless, it still would predict an ambiguity disadvantage in our tasks. That is, to make "No" decisions on the negative trials, participants would have to ensure that there are no meanings that fall into the semantic category in question. After the semantic representation for the first lexical entry had been checked, each of the other lexical entries with the same spelling would need to be accessed and their semantic representations would have to be checked. As a result, the response latencies on the negative trials should increase depending on the number of meanings a word has, producing a processing time disadvantage for ambiguous words compared to unambiguous words in our semantic categorization tasks. Thus, the model would not be able to explain the results of Experiments 3 and 4.

When considering whether semantic representations should be thought of as being distributed or local, it should also be noted that there are data which favor the distributed representation assumption. For example, Pexman, Lupker, and Hino (2002) recently reported that both lexical decision and naming performance are affected by the number of semantic features a word has (see also Grondin, Lupker, & McRae, in press). Based on norms provided by McRae, de Sa, and Seidenberg (1997), Pexman et al. manipulated the number of features for their unambiguous word stimuli. In their lexical decision and naming tasks, response latencies were faster for words with a high number of features than for words with a low number of features. Assuming that each word meaning is represented as a local semantic representation, there would be no reason to expect that lexical decision and naming performance would be modulated by the number of semantic features factor. Thus, these results are more consistent with models postulating distributed representations of word meanings. In particular, if there is feedback activation from the semantic level to the orthographic (and phonological) level, orthographic (and phonological) processing should be faster when words are represented by more semantic features due to the greater amount of semantic feedback (see also Strain, Patterson, & Seidenberg, 1995, for a similar account of an imageability effect). Given these considerations it would appear to be premature to abandon the distributed representation assumption, at least with respect to semantic representations. Thus, we next consider some alternatives that maintain that assumption to evaluate whether our data can be explained by any of those proposals.

In general, within a PDP framework, the semantic-coding process must be somewhat noisy in that certain semantic features will not achieve a high level of activation early in processing while others (i.e., ones that actually are not features of the concept) may be activated early in processing but will fall out over time. Given this scenario, it's possible that the semantic-coding process could make use of the rapidly available features allowing it to be short-circuited in easy tasks such as those in Experiments 3 and 4. That is, when a well-defined category is used in a semantic categorization task, it may be possible to make accurate decisions by checking only those features that are strongly activated very early in processing. In such a situation, a "Yes" decision could certainly be made quite easily. With respect to "No" decisions, it is not impossible that a decision could be made well before semantic coding is completed if there was no suggestion among the activated semantic units that a positive response might ultimately be called for. If so, comparable response latencies across the three word types in the "No" trials in Experiments 3 and 4 would follow (although "No" trials may be a bit error prone). In addition, because the "No" responses should be slower if some degree of activation were detected for the relevant semantic units, this account would also explain the typicality effect observed in the negative responses in Experiments 3 and 4.

In contrast, when the category used in a semantic categorization task cannot be characterized by a small set of features, it seems much more likely that the semantic-coding process would have to be completed (or nearly completed) before making decisions. Thus, semantic categorization performance would be more sensitive to the nature of orthographic-to-semantic mappings. Hence, a processing time disadvantage would be observed due to the inconsistent orthographic-to-semantic mappings, as was found for the ambiguous words with less related meanings in Experiments 2 and 5.

These principles are essentially those embodied in Piercey and Joordens's (2000) PDP account. This type of account would also be able to explain the results of our lexical decision task. That is, assuming that lexical decision operations are executed based on the total orthographic/semantic activation before completing semantic coding and if semantic activation is modulated by the number of meanings a word has, an ambiguity advantage would be expected along with the lack of a relatedness effect.

Where this type of account runs in to a problem, however, is in explaining the lack of a processing time disadvantage for the ambiguous words with more related meanings when the semantic-coding process is assumed to be completed before responding (in Experiments 2 and 5). That is, even if meanings are related, they are not identical and, hence, they should show some degree of semantic competition in a semantic categorization task (see Rodd et al.'s, 2004, simulation). If one were to assume instead that the semantic representations for ambiguous words with more related meanings were essentially the same as the semantic representations for unambiguous words as suggested by Caramazza and Grober (1976), part of the problem might be solved, however, other problems would be created. In particular, it would raise the question of why there was an ambiguity advantage for ambiguous words with more related meanings in the lexical decision task in Experiment 1. That is, if the semantic representations for ambiguous words with more related meanings and for unambiguous words are essentially the same, there would be no reason to expect processing differences for these two types of words in the lexical decision task. As such, at least at present, it doesn't appear that our results can be explained using this type of account.

The problem with this account, as with others discussed above, is that it is based on the principle that there is a competition during the semantic-coding process. Therefore, these types of accounts inevitably predict an ambiguity disadvantage independent of the nature of the task. What is needed, therefore, is a model in which no competition is created during the semantic-coding process even when the distributed representation assumption is maintained at the semantic level.

A hint for how this might be accomplished is provided by Joordens and Besner's (1994) report that they could increase their model's ability to settle into a correct semantic pattern (i.e., avoid the "blend state" problem) by increasing the number of semantic units. In particular, the rate of correct settling for ambiguous words was 30.8% with 100 semantic units. With 1000 units, the model's performance improved to 48.8%. By increasing the number of semantic units, therefore, the competition created by multiple meanings was weakened. As such, any processing time difference due to the semantic-coding process for the ambiguous and unambiguous words would, presumably, also decrease.

To further minimize the processing time difference during semantic coding for ambiguous and unambiguous words (within PDP models), one could also propose a framework based on the idea that connection weights producing one orthographic-to-semantic association are not modulated by learning another association even when it shares an orthographic pattern. For example, if different patterns of activation are created at the hidden layer for each meaning of an ambiguous word, the connection weights between the hidden units and the semantic units would be different for each meaning. If so, it would be possible to create a PDP model in which competition is substantially reduced during semantic coding even for words with inconsistent orthographic-to-semantic mappings.

One may further wish to consider the possibility that the multiple meanings of a word are mapped onto different sets of semantic units (i.e., different semantic spaces). If these sets of semantic units were either only weakly connected or unconnected with each other, the processing time disadvantage for ambiguous words would be minimal because the orthographic-to-semantic connections adjusted to produce one relationship (e.g., for the dominant meaning) would be completely separate from the connections adjusted to produce the other relationship (e.g., for the subordinate meaning(s)) and, hence, there would be little, if any, competition during either learning or semantic activation.

Note that within this type of architecture, not only would ambiguity not create orthographic–semantic inconsistency/competition but, in addition, the overall amount of semantic activation generated would, nonetheless, still be larger for ambiguous words (because more semantic units are activated for additional meanings). If one then makes the assumptions we have made about the central role of the decision-making process, we would have a model that would be able to account for the results reported in all of the present experiments. Note also that these ideas would be broadly consistent with some neurological findings that different neural regions are associated with retrieval of different types of semantic information (e.g., Damasio, 2001; Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Tranel, Damasio, & Damasio, 1997, 1998).

Although the present research is certainly consistent with the notion that different meanings are represented in different semantic spaces, there are probably a number of other ways to implement similar ideas within a PDP framework. Any of these implementations would, of course, involve some modification in the representational scheme at the semantic level. The question of how best to make this modification, while retaining the basic distributed representation assumption, is an important issue for future research.

Appendix A

Experimental items used in all the experiments along with their English translations

More related		Less related	
Ambiguous word			
カバー	cover compensation singing other person's song a book cover	クラッカー	thin, crisp biscuit firecracker
コンタクト	contact lens contact, touch	コート	wearing coat court in sports coating with paint
サイン	signature signal, motion	サンド	sandwich place tightly between two other things sand, tiny grains of crushed rocks
スーパー	grocery store super, splendid	ジュース	juice, soft drink deuce (in tennis)
スペース	space, room cosmos, universe blank, margin line space, distance	スチール	steel still (picture) base steal (in baseball)
セット	set, a suite setting up set, a match having one's hair set studio set	スパイク	spiked shoes spiking in volleyball
タイプ	type, kind, class typewriter type, prototype print or copy with a typewriter	ソース	sauce source
ドラム	drum, a musical instrument something shaped like a drum	チェック	check, investigate check, a pattern made of squares
ネット	net network	トラック	truck track in a stadium track to record information on magnetic media
パス	passing a ball pass in playing cards pass, a ticket passing an examination	パイ	pie, baked pie tiles used in mahjong breast pi, circle ratio
ヒット	hit in baseball hit, a success	ハイライト	highlight on TV programs Hi-lite, a name of a cigarette highlight on pictures
ブリッジ	bridge (to cross a river) bridge in wrestling	バレー	volleyball ballet valley
ブルー	blue color blue, depression	ベース	base, basis base in baseball bass (guitar)
プレート	plate, a dish plate, a board the earth's crust	ボーリング	Bowling boring (to dig)
プレス	press, place with force the press, news, journalism make smooth with a hot iron	ポスト	post, pillar box post, position afterward post, pole
ブロック	concrete block block in volleyball	ラップ	clear-plastic wrap rap music

(*continued on next page*)

Appendix A (*continued*)

More related		Less related	
	(toy) building block		lap time
	block, a square		
	block, a group		
	block, a lump		
フロント	front desk (in a hotel)	リーチ	reach in mahjong
	front, the foremost part		reach, length of the arm
			Reach, a name of a toothbrush
ポイント	the point, the gist	リスト	list
	point, a score		wrist
	point, a place		
ホーム	home	リボン	ribbon, a strip of cloth
	home base		Ribbon, a name of a magazine
	platform		ribbon used in rhythmic gymnastics
	one's hometown		printer ribbon
リング	(finger) ring	ルート	route, a way to go
	Ring, a title of a drama		square root
	ring, a circle		connections, contacts
	ring for boxing		
Unambiguous word			
アイデア	idea		
ギフト	gift		
コラム	column		
サンダル	sandals		
シーズン	season		
シューズ	shoes		
スイッチ	switch		
スポット	spot		
ズボン	pants		
ゼロ	zero		
バイク	motorbike		
ブック	book		
プラン	plan		
ベンチ	bench		
ポット	pot		
ホリデー	holiday		
マーケット	market		
ランク	rank		
ランプ	lamp		
リサーチ	research		

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. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d'Arcaise, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9–32). Hillsdale, HJ: Lawrence Erlbaum Associates.

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., Ferraro, R. F., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–221). Hillsdale, NJ: Lawrence Erlbaum Associates.

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition, 8*, 493–512.

Besner, D. (1983). Basic decoding components in reading: Two dissociable feature extraction processes. *Canadian Journal of Psychology, 37*, 429–438.

Besner, D., & Joordens, S. (1995). Wrestling with ambiguity—further reflections: Reply to Masson and Borowsky (1995) and Rueckl (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 515–519.

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.

Caramazza, A., & Grober, E. (1976). Polysemy and the structure of the subjective lexicon. In C. Rameh (Ed.), *Georgetown University roundtable on languages and linguistics. Semantics: Theory and application* (pp. 181–206). Washington, DC: Georgetown Univ. Press.

Clark, H. H. (1973). The Language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335–359.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). New York: Academic Press.

Damasio, H. (2001). Words and concepts in the brain. In J. Branquinho (Ed.), *The foundations of cognitive science* (pp. 109–120). Oxford, UK: Oxford University Press.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. (1996). A neural basis for lexical retrieval. *Nature, 380*, 499–505.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*, 429–446.

Ferraro, F. R., & Hansen, C. L. (2002). Orthographic neighborhood size, number of word meanings, and number of higher frequency neighbors. *Brain and Language, 82*, 200–205.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walk (Eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North-Holland.

Forster, K. I. (1999). *Beyond lexical decision: Lexical access in categorization tasks.* Paper presented at the 40th Annual Meeting of the Psychonomic Society: Los Angeles, CA.

Forster, K. L. (2004). Category size effects revisited: Frequency and masked priming effects in semantic categorization. *Brain and Language, 90*, 276–286.

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

Forster, K. I., & Hector, J. (2002). Cascaded versus noncascaded models of lexical and semantic processing: The turple effect. *Memory & Cognition, 30*, 1106–1117.

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.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*, 518–565.

Grondin, R., Lupker, S. J., & McRae, K. (in press). Shared features dominate the number-of-features effect. *Proceedings of the twenty-eighth annual meeting of the Cognitive Science Society*. Lawrence Erlbaum Associates: Mahwah, NJ.

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. (1998). The effects of word frequency for Japanese Kana and Kanji words in naming and lexical decision: Can the dual-route model save the lexical-selection account? *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1431–1453.

Hino, Y., & Lupker, S. J. (2000). The effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 166–183.

Hino, Y., Lupker, S. J., & Besner, D. (1998). *Polysemy effects in the naming of Japanese Katakana words and their Hiragana transcriptions.* Presented at the 39th Annual Meeting of the Psychonomic Society, Dallas, TX, USA.

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., & Pexman, P. M. (2005). *Effects of orthographic–semantic consistency for Japanese Kanji words.* Presented at the 46th Annual Meeting of the Psychonomic Society, Toronto, Ont., Canada.

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.

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.

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.

Landauer, T. K., & Freedman, J. L. (1968). Information retrieval from long-term memory: Category size and recognition time. *Journal of Verbal Learning and Verbal Behavior, 7*, 291–295.

Locker, L., Simpson, G. B., & Yates, M. (2003). Semantic neighborhood effects on the recognition of ambiguous words. *Memory & Cognition, 31*, 505–515.

Masson, M. E. J., & Borowsky, R. (1995). Unsettling questions about semantic ambiguity in connectionist models: Comments on Joordens and Besner (1994). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 509–514.

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word-frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance, 14*, 693–706.

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

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126*, 99–130.

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.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76*, 165–178.

National Language Research Institute (1971). *Denshikeisanki niyoru shinbun no goi-chosa (II)* [Studies on the vocabulary of modern newspapers (Volume 2), Vocabulary tables classified from the lexicological and grammatical point of view]. Tokyo: Shuei Shuppan (in Japanese).

National Language Research Institute (1993). Bunrui Goi Hyou (Floppy Disk Version) [Thesaurus (Floppy Disk Version)]. Tokyo: Shuei Shuppan (in Japanese).

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation–verification model for letter and word recognition: The word-superiority effect. *Psychological Review, 89*, 573–594.

Pecher, D., Zeelenberg, R., & Wagenmakers, E. (2005). Enemies and friends in the neighborhood: Orthographic similarity effects in semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 121–128.

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.

Pexman, P. M., Lupker, S. J., & Hino, Y. (2002). The impact of feedback semantics in visual word recognition: Number of features effects in lexical decision and naming tasks. *Psychonomic Bulletin & Review, 9*, 542–549.

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.* Hillsdale, NJ: Erlbaum, pp. 824–829.

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.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*, 191–201.

Rodd, J. (2004). When do leotards get their spots? Semantic activation of lexical neighbors in visual word recognition. *Psychonomic Bulletin & Review, 11*, 434–439.

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.

Rosch, E., & Mervis, C. B. (1975). Family resemblance studies in the internal structure of categories. *Cognitive Psychology, 7*, 573–605.

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.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1140–1154.

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: Kluwer.

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

Taft, M., & Kougious, P. (2004). The processing of morpheme-like units in monomorphemic words. *Brain and Language, 90*, 9–16.

Tranel, D., Damasio, H., & Damasio, A. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia, 35*, 1319–1327.

Tranel, D., Damasio, H., & Damasio, A. (1998). The neural basis of lexical retrieval. In R. W. Parks & D. S. Levine, et al. (Eds.), *Fundamentals of neural network modeling: Neuropsychology and cognitive neuroscience* (pp. 271–296). Cambridge, MA: The MIT Press.

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

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.