Journal of Experimental Psychology:
Learning, Memory, and Cognition
2002, Vol. 28, No. 4, 686–713

0278-7393/02/$5.00 DOI: 10.1037//0278-7393.28.4.686

Ambiguity and Synonymy Effects in Lexical Decision, Naming, and Semantic Categorization Tasks: Interactions Between Orthography, Phonology, and Semantics

Yasushi Hino
Chukyo University

Stephen J. Lupker
The University of Western Ontario

Penny M. Pexman
University of Calgary

In this article, ambiguity and synonymy effects were examined in lexical decision, naming, and semantic categorization tasks. Whereas the typical ambiguity advantage was observed in lexical decision and naming, an ambiguity disadvantage was observed in semantic categorization. In addition, a synonymy effect (slower latencies for words with many synonyms than for words with few synonyms) was observed in lexical decision and naming but not in semantic categorization. These results suggest that (a) an ambiguity disadvantage arises only when a task requires semantic processing, (b) the ambiguity advantage and the synonymy disadvantage in lexical decision and naming are due to semantic feedback, and (c) these effects are determined by the nature of the feedback relationships from semantics to orthography and phonology.

Over the past 30 years, a central question in reading research has been the following: How do semantic factors influence reading processes? One means of addressing this question has been to look for semantic effects in isolated word-recognition tasks. For example, as has been reported by a number of researchers, words with multiple meanings (e.g., *bank*, *lean*) are typically responded to faster than words with fewer meanings (e.g., *food*, *tent*) in both lexical decision and naming tasks (e.g., Gottlob, Goldinger, Stone, & Van Orden, 1999; Hino & Lupker, 1996; Hino, Lupker, Sears, & Ogawa, 1998; Jastrzembski, 1981; Kellas, Ferraro, & Simpson, 1988; Lichacz, Herdman, LeFevre, & Baird, 1999; Millis & Button, 1989; Rubenstein, Garfield, & Millikan, 1970).

As critics have noted, however, not only have some of the reported "ambiguity" effects been relatively small but also some researchers (e.g., Borowsky & Masson, 1996; Forster & Bednall, 1976; Gernsbacher, 1984) have failed to observe any effects at all. Thus, the reality of ambiguity effects has not been universally accepted (e.g., Azuma & Van Orden, 1997; Clark, 1973; Gernsbacher, 1984; Gottlob et al., 1999; Rueckl, 1995). For example, Clark (1973) argued that it is important to treat items as a random factor in analyses of variance (ANOVAs) when analyzing data from word-recognition research so that the results can be generalized beyond the stimulus set used in that particular experiment. When Clark reanalyzed Rubenstein et al.'s (1970) data with items as well as subjects as random factors, he failed to observe a significant ambiguity effect. Thus, Clark concluded that the ambiguity effects observed by Rubenstein et al. (1970) were due simply to the idiosyncratic nature of their stimulus materials.

As a number of researchers (Cohen, 1976; Keppel, 1976; Raaijmakers, Schrijnemakers, & Gremmen, 1999; J. E. K. Smith, 1976; Wike & Church, 1976) have pointed out, however, generalizability is not a statistical issue. The only way to determine whether any effect generalizes over items is through replication using new sets of items. In fact, since Rubenstein et al.'s (1970) first report of an ambiguity effect, this effect has been replicated by a number of researchers (e.g., Gottlob et al., 1999; Hino & Lupker, 1996; Hino et al., 1998; Jastrzembski, 1981; Kellas et al., 1988; Lichacz et al., 1999; Millis & Button, 1989) using many different stimulus sets, providing good support for the argument that the ambiguity advantage in lexical decision and naming tasks is, indeed, a real one.

Although reports of ambiguity effects are now pervasive in the literature, none of these reports makes an explicit distinction between words that are ambiguous because they have a number of unrelated meanings (homonyms) and words that are ambiguous because they have a number of related senses (polysemous words).

Yasushi Hino, Department of Psychology, Chukyo University, Toyota, Aichi, Japan; Stephen J. Lupker, Department of Psychology, The University of Western Ontario, London, Ontario, Canada; Penny M. Pexman, Department of Psychology, University of Calgary, Calgary, Alberta, Canada.

We thank Alan Kawamoto for providing us with the additional information needed to build our simulation of Kawamoto et al.'s (1994) model. We also thank Derek Besner and an anonymous reviewer for their comments on a draft of this article, as well as Akiko Uchida and Yoko Atoda for their assistance in the data collection.

Correspondence concerning this article should be addressed to Yasushi Hino, care of Shinji Yukumatsu, Institute for Advanced Studies in Artificial Intelligence, Chukyo University, 101 Tokodate, Kaizu-cho, Toyota, Aichi 470-0393, Japan. E-mail: hino@sccs.chukyo-u.ac.jp

On the other hand, there are some researchers (especially linguists and psycholinguists working on lexical semantics) who regard this distinction as an important one because they believe that homonymous meanings and polysemous senses are represented in different ways (see Klein & Murphy, 2001, for a review). For example, Caramazza and Grober (1976) suggested that whereas homonymous meanings are separately represented, polysemous senses are not. For polysemous words, what is represented is only a small number of abstract core meanings (possibly only one). When these words are accompanied by context, each sense that is consistent with that context is generated on the basis of the activated core meaning(s). According to this type of theory, it would be quite important to distinguish between homonymous meanings and polysemous senses. In particular, if the ambiguity advantage were due to having multiple meaning representations, this advantage should arise only for homonyms and not for polysemous words.

Klein and Murphy (2001) have recently examined the question of whether multiple senses for polysemous words are represented separately using a priming technique in their sensicality-judgment task. In this task, participants were asked to decide whether a given phrase makes sense by pressing one of two buttons. In Klein and Murphy's experiments, noun phrases were created by presenting polysemous words preceded by a modifier (e.g., *shredded paper*). The same polysemous word was accompanied by different modifiers in different phrases, so that the polysemous word in pairs of phrases denoted either the same sense (e.g., *wrapping paper–shredded paper*) or different senses (e.g., *daily paper–shredded paper*). The expectation was that if polysemous words (e.g., *paper*) are represented by only an abstract core meaning, phrases like *wrapping paper* and *daily paper*, for example, would facilitate the sensicality-judgment responses for *shredded paper* to the same extent. On the other hand, if polysemous senses are separately represented, facilitation would be expected only when the two phrases denoted the same sense of the polysemous word (e.g., *wrapping paper–shredded paper*). Klein and Murphy's results were consistent with the latter prediction. In addition, the priming effect sizes were similar when the target phrases involved polysemous words and when the target phrases involved homonyms. On the basis of these results, Klein and Murphy concluded that, like homonymous meanings, the polysemous senses of a word are represented separately in a reader's semantic system.

Following Klein and Murphy's (2001) conclusions, therefore, it seems reasonable to assume that both homonymous meanings and polysemous senses are represented in a similar fashion and, hence, in our experiments, we did not make any specific distinction between homonyms and polysemous words. Thus, throughout this article we simply refer to both homonyms and polysemous words as *ambiguous words*, words with multiple meanings.

Ambiguity Effects in Lexical Decision and Naming Tasks

As noted above, a number of researchers have replicated the finding that words with multiple meanings are responded to faster than words with fewer meanings in both lexical decision and naming tasks. These ambiguity effects, at least in the lexical decision task, were initially explained within the framework of classical lexical models (e.g., Becker, 1980; Forster, 1976; McClelland & Rumelhart, 1981; Morton, 1969; Paap, Newsome, McDonald, & Schvaneveldt, 1982). In these types of models, orthographic representations are first constructed on the basis of the visual input, and a lexical unit is then selected on the basis of those orthographic representations. At this point, sufficient lexical information would be available to allow a positive response. Semantic information would become available only after selecting the appropriate lexical unit. Hence, according to the simple versions of these models, semantic information should not affect the lexical selection process. Rubenstein et al. (1970) and Jastrzembski (1981) modified this description by assuming that ambiguous words were represented by multiple lexical units, whereas unambiguous words were represented by a single lexical unit. As a result, there would be an increased probability of rapidly selecting an appropriate unit for an ambiguous word than for an unambiguous word, producing a processing-time advantage for ambiguous words.

In contrast, Balota, Ferraro, and Connor (1991) suggested that ambiguity effects could be explained within a classical lexical framework if the assumption was made that there was feedback activation between the semantic and lexical levels (e.g., McClelland & Rumelhart, 1981). Instead of assuming multiple lexical units for ambiguous words, Balota et al. assumed that ambiguous words are represented by single lexical units that are linked to multiple semantic units. When activation is instigated at the semantic level, the amount of semantic activation tends to be greater for ambiguous words than for unambiguous words because ambiguous words are more densely represented at the semantic level. Thus, the feedback activation from semantic units to lexical units would also be greater for ambiguous words, resulting in faster lexical selection for ambiguous words than for unambiguous words.

In a further investigation of these issues, Hino and Lupker (1996) noted that if the lexical selection process were involved in both lexical decision and naming tasks and if ambiguity effects were due simply to lexical selection, ambiguity effects should be fairly similar across different tasks. What Hino and Lupker (1996) reported, however, were somewhat different patterns of ambiguity effects in their lexical decision and naming tasks. Specifically, identical ambiguity effects were observed for both high- and low-frequency words in the lexical decision task (as was also reported by Rubenstein et al., 1970). In Hino and Lupker's (1996) naming task, on the other hand, the ambiguity effect was limited to low-frequency words. A similar interaction between ambiguity and frequency in naming was also reported by Lichacz et al. (1999).

As Hino and Lupker (1996) noted, however, it may be possible to reconcile these results with a lexical selection account. As in Jastrzembski's (1981) and Balota et al.'s (1991) models, one could assume that the lexical selection process is sensitive to both frequency and ambiguity. Thus, a Frequency $\times$ Ambiguity interaction would be expected to arise during the lexical selection process. For example, one could argue that because lexical selection for high-frequency words is much faster than that for low-frequency words, the impact of ambiguity would be much greater for low-frequency words. Thus, the interaction that Hino and Lupker (1996) observed between frequency and ambiguity in a naming task would follow.

In a lexical decision task, however, the situation is somewhat different. That is, as has been suggested by a number of researchers (e.g., Balota, 1990; Balota & Chumbley, 1984; Besner, 1983; Besner & McCann, 1987; McCann, Besner, & Davelaar, 1988; Seidenberg, Waters, Barnes, & Tanenhaus, 1984), a lexical decision task also involves a decision-making process following lexical selection, a process that is based on stimulus familiarity. Further, as Gernsbacher (1984) has argued, stimulus familiarity is correlated with both ambiguity and frequency. Thus, this decision-making process would, presumably, be sensitive to both factors. As a result, the observed pattern of results in lexical decision would reflect not only the impact of lexical selection but also the impact of this decision-making process. In particular, it is possible that the relationship between frequency and ambiguity could change from an interactive one to an additive one if a larger ambiguity effect were produced for high-frequency words than for low-frequency words during the decision-making process. As such, the different patterns of results in lexical decision and naming tasks could be explained within a lexical selection account.

To evaluate this possibility, Hino and Lupker (1996) examined ambiguity effects in a go/no-go naming task. In this task, participants are asked to name a stimulus aloud only if it is a word. Because this task requires participants to make implicit lexical decisions, both the lexical selection and decision-making processes are required. Thus, even though the task is a naming task, if ambiguity effects are due to the contributions of both the lexical selection process and the decision-making process, the ambiguity effects observed here should have mirrored those in the lexical decision task.

On the other hand, if ambiguity effects in the lexical decision and naming tasks were due not to lexical selection but rather to task-specific processes, because the go/no-go naming task involves both a decision-making process (which is specific to lexical decision) and a phonological coding process (which is specific to naming) in a quasi-sequential order, the expected ambiguity-effect sizes should have been similar to the sum of the effects in the lexical decision and naming tasks. The results were consistent with the latter predictions. That is, ambiguity effects were observed for both high- and low-frequency words, and the effect size was larger for low-frequency words than for high-frequency words. (For low-frequency words, the 42-ms effect in go/no-go naming was similar to the sum of the 13-ms effect in lexical decision and the 21-ms effect in naming. For high-frequency words, the 16-ms effect in go/no-go naming was similar to the sum of the 13-ms effect in lexical decision and the 1-ms effect in naming.) On the basis of these results, Hino and Lupker (1996) argued that it is unlikely that ambiguity effects can be adequately explained by classical lexical models, even expanded versions like those of Rubenstein et al. (1970) or Balota et al. (1991).

Consistent with Hino and Lupker's (1996) suggestions, some researchers have recently proposed nonlexical accounts of ambiguity effects based on parallel distributed processing (PDP) models. These PDP models (e.g., Plaut, 1997; Plaut & McClelland, 1993; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg, 1992; Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990) do not involve either a lexicon or a lexical selection process. Instead, words are assumed to produce patterns of activation over sets of units representing orthographic, phonological, and semantic features. These units are connected with each other and, through a learning process, these connections come to be weighted in a way that reflects the appropriate relationships among units.

More specifically, according to the PDP models, when a word is viewed, orthographic units are first activated, and this activation spreads to phonological and semantic units through these weighted connections. As such, the phonological and semantic coding processes are simply the network computing the activation pattern in either the phonological or semantic units given the activation pattern in the orthographic units. In addition, because the weights on connections depend on the nature of the relationships between the input and output units (i.e., consistency), the speed and accuracy of phonological and semantic coding is expected to be modulated by the nature of the relationships between units.

For example, because (heterophonic) homographs are words with multiple meanings that have different pronunciations (e.g., *lead*, *wind*), these homographs would involve the mapping of a single orthographic code onto two phonological codes, as well as the mapping of a single orthographic code onto two semantic codes. Because of the one-to-many feedforward relationships between orthographic codes and phonological codes, PDP models predict that the speed of phonological coding should be slower for homographs than for nonhomographs (e.g., Seidenberg & McClelland, 1989). Consistent with these predictions, a homograph disadvantage has typically been reported in naming tasks (e.g., Gottlob et al., 1999; Kawamoto & Zemblidge, 1992; Seidenberg et al., 1984).

Three PDP Accounts of Ambiguity Effects

If one-to-many relationships between the input and output codes always produce a disadvantage in the computation of output codes, one would expect that semantic coding should also be slower for ambiguous words than for unambiguous words. That is, because ambiguous words involve the mapping of a single orthographic code onto multiple semantic codes, the one-to-many feedforward mapping would be expected to produce a cost in terms of the time needed to settle on a semantic code. As previously noted, however, ambiguous words are typically responded to faster than unambiguous words in both lexical decision and naming tasks. Thus, Joordens and Besner (1994) argued that the ambiguity advantage, if such an effect indeed exists, presents a major challenge for PDP frameworks.

As a number of researchers (Borowsky & Masson, 1996; Kawamoto, 1993; Kawamoto, Farrar, & Kello, 1994; Masson & Borowsky, 1995; Rueckl, 1995) have pointed out, however, Joordens and Besner's (1994) argument is valid only if one makes the assumption that semantic coding must be completed in order to respond accurately in the experimental task. If not, then the fact that ambiguous words take longer to settle at the semantic level may be irrelevant. In fact, as has been suggested by a number of researchers (e.g., Balota, 1990; Balota & Chumbley, 1984, 1985; Besner, 1983; Besner & McCann, 1987; Hino & Lupker, 1996, 1998, 2000; McCann, Besner, & Davelaar, 1988; Pexman & Lupker, 1999; Seidenberg & McClelland, 1989), what is most important in the lexical decision task is the decision-making operation that is typically based on the familiarity of the orthographic codes, and what is necessary in the naming task is the phonological-coding process, which then drives the production of overt pronunciation responses. Thus, in neither task would it necessarily be the case that semantic coding would play a major role.

Working on the assumption that lexical decisions are made on the basis of orthographic processing, Kawamoto et al. (1994) were able to simulate ambiguity effects in lexical decision tasks using a PDP network model. In their simulations, the time (the number of cycles) taken to settle on an orthographic code was taken as a measure of lexical decision performance. In line with the above discussion, for ambiguous words, a single orthographic pattern was associated with multiple semantic patterns, whereas for unambiguous words, a single orthographic pattern was associated with a single semantic pattern. Because the one-to-many associations were somewhat difficult for the network to learn, the connections between orthographic and semantic units were weaker for ambiguous words than for unambiguous words. Thus, as suggested by Joordens and Besner (1994), the time to settle on a semantic code was, in fact, somewhat slower for ambiguous words than for unambiguous words. More important, however, although Kawamoto et al. failed to simulate an ambiguity advantage when the model was trained by the Hebbian learning algorithm, when it was trained by the least mean square error-correction learning algorithm, the weaker connections between orthographic and semantic units for ambiguous words were compensated for by the model establishing stronger connections among orthographic units for those words. That is, whereas the mean absolute values of the weights on the orthographic-to-semantic connections were smaller when the model was trained with ambiguous words rather than with unambiguous words, these values on the orthographic-to-orthographic connections were larger when the model was trained with ambiguous words rather than with unambiguous words. As a result of the stronger connections among orthographic units for ambiguous words than for unambiguous words, the time to settle on an orthographic code was faster for ambiguous words than for unambiguous words. On the basis of this simulation, Kawamoto et al. suggested that ambiguity effects in lexical decision tasks are due to the faster orthographic processing for ambiguous words.

Kawamoto et al.'s (1994) account has been challenged recently by Borowsky and Masson (1996). Borowsky and Masson used different types of nonwords in their two lexical decision experiments: orthographically legal nonwords and orthographically illegal nonwords. In their experiments, the ambiguity effect was significant only when the legal nonwords were used. Similarly, Pexman and Lupker (1999) examined ambiguity effects in two lexical decision experiments using legal nonwords and pseudohomophones. Pexman and Lupker reported that the ambiguity effect was larger with pseudohomophones than with legal nonwords. According to Borowsky and Masson, the fact that ambiguity effects increase when a deeper level of processing is required for the nonwords (pseudohomophones require deeper processing than legal nonwords, which in turn require deeper processing than illegal nonwords) suggests that the locus of ambiguity effects is semantic rather than orthographic.

As a result, Borowsky and Masson (1996) proposed a somewhat different simulation of ambiguity effects using their PDP model. Working with the idea that lexical decisions are based on stimulus familiarity, Borowsky and Masson computed the sum of *energy* at the orthographic and semantic levels as a measure of stimulus familiarity. This energy is a metric representing the network's activity toward a basin of attraction (i.e., a learned pattern of activation). The energy value would decrease (take on a larger negative value) when the network's activity approached a basin of attraction. In Borowsky and Masson's simulation of lexical decision tasks, word decisions were made when the summed energy reached a (negative) criterion value.

In Borowsky and Masson's (1996) simulations, semantic units were initially set to random states and updated across cycles. Similar to Kawamoto et al.'s (1994) simulations, semantic units settled into a stable state more slowly for ambiguous words than for unambiguous words. In addition, in nearly 92% of trials involving ambiguous words, semantic units settled into blend states, in which both meanings were partially activated. However, when the sum of energy at the orthographic and semantic levels was measured, the summed energy value reached a criterion value faster for ambiguous words than for unambiguous words. According to Borowsky and Masson, the reason the network moves into a basin of attraction more quickly for ambiguous words is because the distance from the initial random state to a basin of attraction is, on average, smaller when there are multiple basins of attraction than when there is a single basin of attraction.

In contrast, Hino and Lupker (1996) explained ambiguity effects in terms of semantic-feedback activation. As Kawamoto et al. (1994) and, to some extent, Borowsky and Masson (1996) did, Hino and Lupker (1996) assumed that lexical decisions are based primarily on the orthographic familiarity of stimuli. In contrast, in naming tasks, the assumption was that phonological codes play the central role, because pronunciation responses are required. What Hino and Lupker (1996) suggested was that ambiguity effects in lexical decision tasks are due to feedback activation from the semantic level to the orthographic level, whereas ambiguity effects in naming tasks are due to feedback activation from the semantic level to the phonological level. Accordingly, because ambiguous words activate multiple semantic codes, ambiguous words produce a greater amount of semantic activation than do unambiguous words. Thus, the amount of feedback activation from the semantic level to the orthographic or phonological level is greater for ambiguous words. As a result, both the orthographic processing required in making lexical decisions and the phonological coding required in naming receive more support from the semantic feedback for ambiguous words than for unambiguous words.

The crux of Hino and Lupker's (1996) account was that ambiguity effects are due to feedback activation from semantics to either orthography or phonology. In contrast, Kawamoto et al.'s (1994) account was based on the strength of connections at the orthographic level. Although Kawamoto et al.'s model involves semantic-feedback connections to orthographic units, semantic-feedback activation was not the source of ambiguity effects. Also in contrast to Hino and Lupker's (1996) position, Borowsky and Masson (1996) suggested that ambiguity effects are due to a faster decrease in energy values at the orthographic and semantic levels because of the expected proximity of at least one semantic representation for ambiguous words. Note also that in Borowsky and Masson's model, semantic feedback is not assumed to update orthographic activation, because this model assumes that orthographic units are updated by external inputs only. In addition, because Borowsky and Masson failed to observe a significant ambiguity effect in their naming experiment, when they simulated naming performance using their model, they weakened the influence of semantic feedback to phonological units in order to simulate their null ambiguity effect. As such, according to Borowsky and Masson's model, the influence of semantic feedback on both orthographic and phonological processing is minimal.

The purpose of the present research was to evaluate Hino and Lupker's (1996) feedback account and to contrast it with Borowsky and Masson's (1996) proximity account and Kawamoto et al.'s (1994) orthographic account.

Ambiguity Effects in a Semantic Categorization Task

According to Hino and Lupker's (1996) account, the ambiguity advantage should be limited to tasks in which performance is affected by semantic feedback. That is, when a task is accomplished mainly on the basis of orthographic or phonological processing (as with the lexical decision and naming tasks), processing, and hence task performance, would be affected by semantic feedback. If a task requires meaning determination, however, the responses would be made based mainly on the results of semantic processing. Thus, task performance would not be sensitive to feedback from the semantic level to other levels. Rather, task performance should be more sensitive to the speed of meaning determination, that is, the speed of settling at the semantic level with that speed being most affected by the nature of the feedforward relationships from orthography to semantics. Therefore, if a task does require meaning determination, as Joordens and Besner (1994) have argued, an ambiguity disadvantage should be observed because of the one-to-many feedforward mappings between orthography and semantics. In addition, as noted above, an ambiguity disadvantage was actually observed in both Kawamoto et al.'s (1994) and Borowsky and Masson's (1996) simulations when the time needed to settle on a semantic code was measured. As such, all three accounts appear to predict an ambiguity disadvantage when a task requires meaning determination. (As discussed in the General Discussion, however, this prediction of the PDP models is not as straightforward as it seems. Nonetheless, for present purposes we continue to assume that all the PDP models predict a processing disadvantage at the semantic level for ambiguous words.)

Consistent with this prediction, some researchers have, in fact, reported that fixation times are longer for ambiguous words than for unambiguous words when the words are presented in a neutral sentential context (e.g., Duffy, Morris, & Rayner, 1988; Rayner & Duffy, 1986). Also consistent with this prediction is the fact that reaction times are slower for ambiguous words than for unambiguous words in association-judgment tasks (Gottlob et al., 1999; Piercey & Joordens, 2000), tasks in which two words are presented sequentially and the decision is whether the words are related. Unfortunately, some of the results in this task are open to alternative interpretations. For example, when the ambiguous word is presented first and the stimulus onset asynchrony is long (as it was in Piercey & Joordens's, 2000, experiment and in Experiment 2 of Gottlob et al., 1999), the participant has sufficient time to select and focus on one of the multiple meanings of the ambiguous word before the second word (unambiguous) appears. Thus, the delay in responding to the second word may reflect the fact that on some of the trials the selected meaning was not the one that was related to the meaning of the second word. Experiment 3 of Gottlob et al. (1999), in which the ambiguous words were presented second, avoids this particular criticism.

There is also a second interpretation problem, however. In all cases, the crucial results came from trials in which an ambiguous word or an unambiguous word was paired with a related word, hence requiring a "yes" response. To create the ambiguous word trials, each ambiguous word was paired with a word that was related to only one meaning of that ambiguous word. Thus, it is possible that the ambiguity disadvantage was caused not because it is more difficult to determine the meaning of ambiguous words but because the other activated meanings of the ambiguous words produced a bias toward a "no" response (these meanings would not have been related to the meaning of the paired word). That is, it is possible that the ambiguity disadvantage observed in the association-judgment tasks was not due to the meaning-determination process but rather due to the decision-making process.

In the present experiments, the task selected to tap semantic memory was a semantic categorization task (in which participants determine whether the word is the name of a living or nonliving object). This task requires participants to determine whether a meaning of a word falls into a certain semantic category, and thus PDP models would predict an ambiguity disadvantage in this task. In addition, to address the response-bias problem mentioned above, we selected only ambiguous words in which all the meanings fall into the same semantic category (i.e., the nonliving-object category). In contrast, when these same words are used in a lexical decision experiment, the models outlined above would all predict the standard ambiguity advantage, although these models differ in terms of how this ambiguity advantage is produced. To examine these predictions, therefore, we first conducted lexical decision (Experiment 1) and semantic categorization experiments (Experiment 2) using the same ambiguous and unambiguous words.

Synonymy Effects and the Nature of Semantic-Feedback Relationships

The central distinction between Hino and Lupker's (1996) account and the others is that it is an account based on feedback activation from semantics to orthography. Thus, the best way to discriminate between this account and the others would be to determine whether there is other evidence of feedback activation affecting task performance.

One can find such evidence, for example, in the results reported by Peterson and Savoy (1998). In Peterson and Savoy's task, participants were asked to name a picture aloud. A visually presented word target was occasionally presented following a picture, and on those trials, participants were asked to name the target word aloud, instead of naming the picture. In these experiments, Peterson and Savoy observed identical priming effects for target words that are phonologically related to the dominant and secondary names of the pictures (e.g., *count* was phonologically related to the dominant name of the picture *couch* and *soda* was phonologically related to *sofa*, the secondary name of the same picture). These results clearly indicate that activated semantic codes (due to a picture prime) do feed activation to other levels (i.e., the phonological level).

Along these same lines, some researchers have argued that lexical decision performance is modulated by the degree of feedback consistency between phonology and orthography (e.g., Stone, Vanhoy, & Van Orden, 1997; Taft & Van Graan, 1998; Ziegler, Montant, & Jacobs, 1997; but see Peereman, Content, & Bonin, 1998, for a criticism). For example, Pexman and colleagues (Pexman & Lupker, 1999; Pexman, Lupker, & Jared, 2001; Pexman, Lupker, & Reggin, 2002) have suggested that the homophone disadvantage in lexical decision tasks (e.g., Davelaar, Coltheart, Besner, & Jonasson, 1978; Rubenstein, Lewis, & Rubenstein, 1971) is due to the one-to-many feedback mappings between phonology and orthography. Homophones are words that are identically pronounced but have two spellings (i.e., *made* and *maid*). After the initial automatic phonological activation (e.g., Lesch & Pollatsek, 1993; Lukatela, Frost, & Turvey, 1998; Perfetti, Bell, & Delaney, 1988; Taft & Van Graan, 1998; Van Orden, 1987; Van Orden, Johnston, & Hale, 1988; Van Orden et al., 1990), the one-to-many feedback relationships for homophones should then create competition among the orthographic codes for the different spellings of the homophone. This competition would give rise to a disadvantage in terms of orthographic processing and, hence, longer lexical decision latencies for homophones than for nonhomophones, as is typically observed.

If Pexman et al.'s (2001) account is correct, orthographic processing should be affected in a similar manner by one-to-many feedback relationships between semantics and orthography. That is, semantic feedback should cause orthographic confusion if there was a one-to-many feedback relationship between semantics and orthography. This type of relationship exists for words with synonyms. That is, because synonyms are words possessing identical meanings (e.g., *brochure*, *pamphlet*), words with synonyms would possess one-to-many relationships between semantics and orthography, and hence the feedback from the activated semantic units would be shared by the orthographic representations of all the synonyms. As a result, orthographic processing (and lexical decision responses) should be slower for words with many synonyms than for words with few synonyms.

Consistent with this prediction, Pecher (2001) recently reported that reaction times were slower for words with a familiar synonym than for words without a familiar synonym in both lexical decision and naming tasks. The present experiments provide a more extensive, and better controlled, follow-up to Pecher's initial report. For example, in Pecher's experiments, the influence of ambiguity was controlled only in Experiment 3. Even there, however, this was done by removing "all words that had two or more unrelated meanings according to Webster's dictionary (1989)" (Pecher, 2001, p. 548). As some researchers have noted, counting meanings listed in a dictionary does not provide a good measure of the number of meanings readers actually think of (e.g., Gernsbacher, 1984; Millis & Button, 1989). Thus, it is not impossible that the synonymy factor was confounded with ambiguity in Pecher's experiments. In addition, no attempt was made to control the degree of spelling-to-sound consistency. This factor also may have been confounded with synonymy, which would have had a major impact in Pecher's naming experiment.

The Present Experiments

Our experiments were conducted using Japanese materials. Ambiguity was measured by obtaining number-of-meanings (NOM) ratings, in which participants were asked to decide whether a given character string possessed no meaning, only one meaning, or more than one meaning, using a 3-point scale (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Kellas et al., 1988). Across different synonymy conditions ambiguity was controlled on the basis of these ratings. The degree of character-to-sound consistency was controlled in two different ways. First, some of our experiments involved Katakana words. Because Katakana is a shallow orthography, its character-to-sound relationships are consistent in all experimental conditions. Second, in those experiments in which Kanji words were used, the summed frequencies of phonological friends and enemies were matched as closely as possible across word groups (i.e., across synonymy conditions).

We also conducted semantic categorization tasks using the same stimuli. As previously noted, because semantic categorization tasks require meaning determination, the responses would be made mainly on the basis of the results of semantic processing. As such, feedback from the semantic level to other levels should be essentially irrelevant in the semantic categorization task. Thus, if the synonymy effect really were due to semantic feedback to orthography and phonology, this effect should not emerge in semantic categorization tasks.

In summary, according to Hino and Lupker's (1996) feedback account, both orthographic and phonological processing are assumed to be affected by semantic-feedback activation. Thus, both an ambiguity advantage and a synonymy disadvantage were expected in both lexical decision and naming tasks. In addition, assuming that semantic coding is sensitive to the nature of orthographic-to-semantic mappings, this account predicts an ambiguity disadvantage and no synonymy effect in a semantic categorization task.

On the other hand, without postulating feedback activation from semantics to orthography and by minimizing the effect of semantic feedback to phonology as in Borowsky and Masson's (1996) account, there would be no reason to expect a processing disadvantage due to synonymy in lexical decision, naming, or semantic categorization tasks. In fact, because there are multiple basins of attraction for words with synonyms at the orthographic level, it is possible that this model could even predict a synonymy advantage in lexical decision.

In contrast, Kawamoto et al.'s (1994) model involves semantic feedback to orthographic units, and the orthographic units' activation is driven by both external inputs and the inputs from other units (including semantic units). Thus, it appears that it would be possible to produce a synonymy disadvantage in lexical decision if the semantic feedback were strong enough to influence the time to settle on an orthographic code. To evaluate this possibility, we attempted to simulate the effects of synonymy in lexical decision and semantic categorization using Kawamoto et al.'s model. In these simulations, two orthographic patterns were associated with a single semantic pattern for a word with a synonym, whereas a single orthographic pattern was associated with a single semantic pattern for a word without a synonym (i.e., a control word). Lexical decision performance was evaluated by determining the time to settle on an orthographic code. Semantic categorization performance was evaluated by determining the time to settle on a semantic code. In all other respects, we followed Kawamoto et al.'s description of their model as closely as possible.[1]

This model did produce a synonymy disadvantage in lexical decision. As shown in Figure 1, the time to settle on an orthographic code was slower for a word with a synonym than for a word without a synonym. In addition, when the semantic-feedback connections were removed (i.e., when only the orthographic-to-orthographic connections were used), this synonymy disadvantage disappeared. In fact, the orthographic activation developed faster for a word with a synonym (see Figure 2). This result clearly indicates that the model's ability to produce a synonymy disadvantage was due to semantic feedback.

At the same time, however, the intact model also produced a synonymy disadvantage in semantic categorization. As shown in Figure 3, the time to settle on a semantic code was slower for a word with a synonym than for a word without a synonym. This synonymy disadvantage in semantic categorization appears to be due to the weaker orthographic-to-semantic connections for a word with a synonym because the mean absolute values of the weights on the orthographic-to-semantic connections were .090 for words without a synonym and .057 for words with a synonym. As such, Kawamoto et al.'s (1994) model predicts a synonymy disadvantage not only in lexical decision but also in semantic categorization.

A brief summary of these predictions is shown in Table 1. As this table shows, the three accounts predict different results in terms of the effects of synonymy in lexical decision, naming, and semantic categorization tasks. To evaluate these predictions, we conducted a lexical decision task (in Experiment 3), a semantic categorization task (in Experiment 4), and a naming task (in Experiment 5) using the same stimuli.

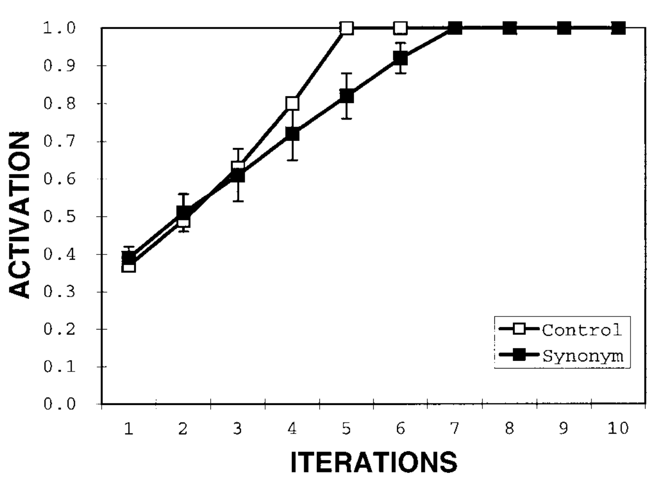

Figure 1. Time course of activation of the orthographic units for Kawamoto et al.'s (1994) model trained with synonyms and the model trained with control words. Orthographic units settled more slowly for synonyms.

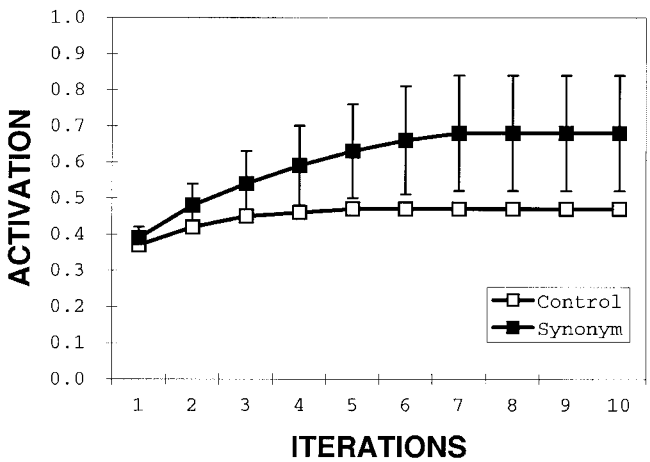

Figure 2. Time course of activation of the orthographic units for Kawamoto et al.'s (1994) model trained with synonyms and the model trained with control words when only the orthographic-to-orthographic connections were used. Orthographic units settled more rapidly for synonyms.

Experiments 1 and 2

Method

Participants. Sixty-four undergraduate students from Chukyo University, Toyota, Aichi, Japan, participated in these experiments for course credit. Thirty students participated in Experiment 1, and the rest participated in Experiment 2. All were native Japanese speakers who had normal or corrected-to-normal vision.

Stimuli. Two hundred forty Katakana words were selected from word-frequency norms for Katakana-written words (National Language Research Institute, 1971, table of loan words listed in order of their frequencies). Thirty participants were asked to rate the NOM possessed by these words. As previously noted, the procedure used for collecting these ratings was identical to that used by Borowsky and Masson (1996), Hino and Lupker (1996), and Kellas et al. (1988). The 240 Katakana words and 60 Katakana nonwords were randomly ordered and listed in a questionnaire. Each item

[1] Following Kawamoto et al.'s (1994) description of their model, the model consisted of four orthographic units and eight semantic units. These units were fully interconnected and trained using the least mean square learning algorithm. To simulate the effects of synonymy, we trained two separate models. Both involved only two stimulus patterns. That is, two orthographic patterns were associated with a single semantic pattern for synonyms, whereas a single orthographic pattern was associated with a single semantic pattern for a control word (i.e., a word without a synonym). In learning trials, all weights on connections between units were initially set to zero and modified after each trial, with the learning constant set to .02. The number of learning trials was equated between the synonyms and the control word. That is, for the model trained with synonyms, each synonym was trained in 20 trials, so that there were 40 learning trials in total. For the model trained with the control word, a single stimulus pattern (for a control word) was trained in 20 learning trials. To examine the performance of each of these models, we presented to the model only orthographic patterns that were scaled by a factor of .25. The decay constant was set to .2. For each test stimulus, the model continued processing until all the units were saturated.

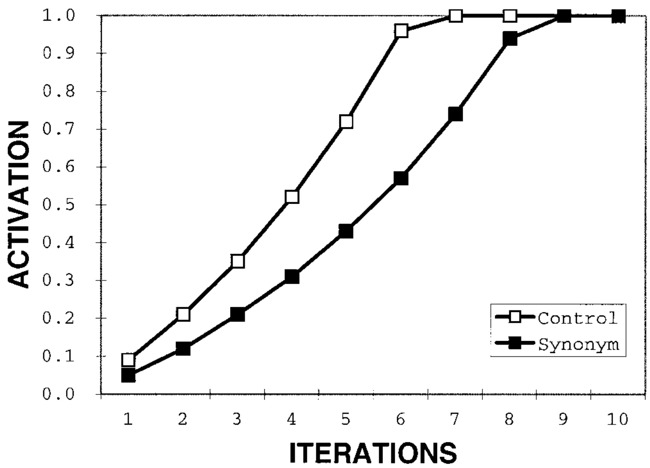

Figure 3. Time course of activation of the semantic units for Kawamoto et al.'s (1994) model trained with synonyms and the model trained with control words. Semantic units also settled more slowly for synonyms.

was accompanied by a 3-point scale. The participants were asked to decide whether the item had *no meaning* (0), *only one meaning* (1), or *more than one meaning* (2) and to circle the appropriate number on the scale.

The Katakana words were classified as ambiguous if their NOM rating was greater than or equal to 1.5, whereas the words were classified as unambiguous if their rating was less than 1.25. In addition, these words were selected for use as targets in the experiment only when all the definitions listed in a Japanese dictionary (Kindaichi, Kindaichi, Kenbou, Shibata, & Yamada, 1974) referred to members of the nonliving-object category. As a consequence, 61 ambiguous and 86 unambiguous Katakana words were selected.

To ensure that these ambiguous words did not involve living-object meanings, we also collected meaning-type ratings. Another 32 participants were asked to decide whether these ambiguous words consisted of *all nonliving-object meanings* (0), *nonliving- and living-object meanings* (1), or *all living-object meanings* (2). The 61 ambiguous Katakana words were randomly ordered and listed in a questionnaire and accompanied by a 3-point scale ranging from 0 to 2. The participants were asked to circle the appropriate number on the scale.

In addition to the 61 ambiguous and 86 unambiguous Katakana words (with nonliving-object meanings), 91 Katakana words with living-object meanings were further selected from the same frequency norms (National Language Research Institute, 1971). Experiential familiarity ratings and typicality ratings were collected for these 238 Katakana words. A different set of 30 participants was asked to rate the experiential familiarity of these words. The 238 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, another 30 participants were asked to rate how typical the meanings of these words are as members of living-object or nonliving-object categories. These 238 words were, once again, randomly ordered and listed in a questionnaire. In the questionnaire, each word was accompanied by a 7-point scale ranging from *nonliving object* (1) to *neutral* (4) to *living object* (7). The participants were asked to rate the typicality as a member of living- or nonliving-object category by circling the appropriate number on the scale.

On the basis of these ratings, 28 ambiguous and 28 unambiguous Katakana words were finally selected. The statistical characteristics of these words are given in Table 2. These words were all between two and five characters in length. They were all low-frequency words, with frequency counts less than or equal to 20 per 940,533. The frequency counts were matched as much as possible for ambiguous words and unambiguous words, $t(54) = 1.24$. The NOM ratings were, however, significantly higher for the ambiguous words (1.68 on average) than for the unambiguous words (1.10 on average), $t(54) = 21.42$, $p < .05$. Word length and the number of syllables (moras) were matched between the two word groups. Experiential familiarity ratings and typicality ratings were also equated as much as possible between the two word groups. There were no significant differences between the two word groups on experiential familiarity ratings, $t(54) = 0.17$, or typicality ratings, $t(54) = 0.35$. In addition, ortho-

Table 1
A Brief Summary of the Predictions From the Three Parallel Distributed Processing Accounts and the Results of the Present Experiments

Account	Lexical decision	Naming	Semantic categorization
	Ambiguity effect		
Kawamoto et al. (1994)	Advantage	Advantage[a]	Disadvantage
Borowsky & Masson (1996)	Advantage	No effect	Disadvantage
Hino & Lupker (1996)	Advantage	Advantage	Disadvantage
Experimental results	Advantage	Advantage	Disadvantage
Experiments	1, 3	5	2, 4
	Synonymy effect		
Kawamoto et al. (1994)	Disadvantage	Disadvantage[a]	Disadvantage
Borowsky & Masson (1996)	No effect or advantage	No effect	No effect
Hino & Lupker (1996)	Disadvantage	Disadvantage	No effect
Experimental results	Disadvantage	Disadvantage	No effect
Experiments	3, 9	7	4, 10

[a] Because Kawamoto et al.'s (1994) original model does not involve phonological units, it is not possible to predict naming performance. However, if phonological units were added to their model, it is quite likely that this modified model would produce an ambiguity advantage and a synonymy disadvantage for naming performance (i.e., the time to settle on a phonological code).

Table 2
Stimulus Characteristics for the Word Stimuli in Experiments 1 and 2

Word type	Freq.	WL	SL	N	FAM	NOM	T	MT
Ambiguous	5.43	3.39	3.32	3.43	4.55	1.68	1.83	0.52
Unambiguous	7.50	3.39	3.32	3.64	4.59	1.10	1.86	

Note. Mean typicality rating for the 56 filler Katakana words used in Experiment 2 was 5.65. Freq. = mean word frequency; WL = word length; SL = syllabic length; N = orthographic neighborhood size; FAM = experiential familiarity rating; NOM = number-of-meanings rating; T = typicality rating; MT = meaning-type rating.

graphic neighborhood sizes were calculated for these words (e.g., Coltheart, Davelaar, Jonasson, & Besner, 1977). That is, the number of words created by replacing one character was counted for each word by using a computer-based dictionary with 36,780 word entries (National Language Research Institute, 1993, sakuin.dat). The mean orthographic neighborhood sizes were comparable between the two word groups, $t(54) = 0.23$. Further, the meaning-type ratings for the ambiguous words were all less than 1 with a mean of .52. Thus, it is unlikely that these words involve (subordinate) living-object meanings. These words and their English translations are listed in Appendix A.

In addition to these experimental word stimuli, 56 Katakana nonwords were created by replacing one character from actual Katakana words, and these nonwords were added to the stimulus set used in the lexical decision task of Experiment 1. The mean character length of these nonwords was 3.36, ranging from two to five. The mean syllabic length of these nonwords was 3.32, ranging from two to five. In the semantic categorization task of Experiment 2, 56 filler Katakana words with living-object meanings were further selected and added to the stimulus set to create living-object trials in addition to the 56 experimental word stimuli with nonliving-object meanings (for nonliving-object trials). The mean word length and syllabic length of these fillers were 3.45 and 3.38, respectively. The mean typicality rating for these fillers was 5.65.

Procedure. Participants were tested individually in a normally lit room. In Experiment 1, participants were asked to make a word–nonword discrimination for a stimulus appearing on a video monitor (PC-TV455; NEC Corporation, Tokyo, Japan) by pressing either a *word* or a *nonword* key on a keyboard. Two keys flanking the space key were used as the *word* and *nonword* keys, respectively (*XFER* and *NFER* keys on an NEC Japanese keyboard). The "word" response was made using the participant's dominant hand.

In Experiment 2, participants were asked to decide whether a word appearing on the video monitor falls into the living-object category or the nonliving-object category by pressing either a *living-object* or a *nonliving-object* key on the keyboard. As in Experiment 1, two keys flanking the space key were used as the *living-object* and *nonliving-object* keys, respectively. The "living-object" 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. Twenty-four practice trials were given prior to the 112 experimental trials in both experiments. The practice items used in Experiment 1 were 12 Katakana words and 12 Katakana nonwords, whereas the practice items used in Experiment 2 were 12 Katakana words with living-object meanings and 12 Katakana words with nonliving-object meanings. None of these items were used in the experimental trials. During the practice 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.

In both experiments, 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. Each participant was seated in front of the video monitor at a distance of about 50 cm and asked to respond to the stimulus by pressing one of two keys on the keyboard. 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 keypress and whether the response was correct were automatically recorded by a computer (PC-9801FA; NEC Corporation, Tokyo, Japan) on each trial. The intertrial interval was 2 s.

Results

Experiment 1 (lexical decision task). Outliers were excluded from the statistical analyses on the basis of Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion. A total of 85 data points (2.53%) were excluded in this fashion. Mean lexical decision latencies for correct responses and mean percentage errors were calculated across subjects.[2]

In all our experiments, significant effects reported were based on a .05 alpha level. Mean lexical decision latencies for ambiguous words and unambiguous words were 592 ms (*SEM* = 13.33) and 611 ms (*SEM* = 13.73), respectively. This 19-ms ambiguity advantage was significant, $t(29) = 4.28$. In addition, mean percentage errors for the ambiguous and unambiguous words were 4.21% (*SEM* = 0.83) and 8.34% (*SEM* = 1.17), respectively. The 4.13% difference in percentage errors was also significant, $t(29) = 3.77$. Mean lexical decision latency and percentage error for the 56 nonwords were 678 ms (*SEM* = 14.82) and 10.42% (*SEM* = 1.46), respectively.

[2] In many of our experiments, the results of the items analyses did not confirm the results of the subjects analyses because the items analyses were not sufficiently powerful to pick up effects of the sizes reported here. We do not, however, regard this as a problem. The items we used were not randomly selected but were selected on the basis of an extensive set of criteria (see Tables 2, 3 and 6). As such, treating items as a random factor would violate many of the assumptions of the ANOVA model with the impact being to further reduce the power of the analysis. Thus, as Wike and Church (1976), Raaijmakers, Schrijnemakers, and Gremmen (1999) and others (Cohen, 1976; Keppel, 1976; J. E. K. Smith, 1976) have argued, items analyses would clearly be inappropriate in the present situation. Therefore, we only report the results of the subjects analyses, and our conclusions are based on those results. We will be happy to provide the results of the items analyses for any of our experiments for any interested readers.

Experiment 2 (semantic categorization task). As in Experiment 1, outliers were excluded from the statistical analyses on the basis of Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion. A total of 92 data points (2.42%) were excluded in this fashion. Mean response latencies for correct responses and mean percentage errors were calculated across subjects.

In the semantic categorization task, mean response latencies were 737 ms ($SEM = 22.18$) for ambiguous words and 707 ms ($SEM = 14.55$) for unambiguous words. This 30-ms ambiguity disadvantage was significant, $t(33) = 2.05$. In addition, mean percentage errors for ambiguous words and unambiguous words were 10.92% ($SEM = 1.89$) and 10.13% ($SEM = 2.05$), respectively. The 0.97% difference was not significant, $t(33) = 0.57$. Mean response latency and percentage error for the 56 fillers were 635 ms ($SEM = 13.38$) and 10.55% ($SEM = 1.06$), respectively.

Discussion

In the lexical decision task, the typical ambiguity advantage was observed. Lexical decision responses were faster and more accurate for ambiguous words than for unambiguous words. Using the same items, however, an ambiguity disadvantage was observed in the semantic categorization task. That is, response latencies were slower for ambiguous words than for unambiguous words in the semantic categorization task.

As previously noted, Gottlob et al. (1999) also reported an ambiguity advantage in their naming task but an ambiguity disadvantage in their association-judgment task. In the present semantic categorization task, participants were, presumably, required to complete meaning determination in order to respond accurately, because participants were asked to decide whether the meaning of a given word falls into a certain semantic category. In such a situation, task performance would have been most sensitive to the nature of the feedforward relationships between orthography and semantics, and, hence, the one-to-many feedforward mappings from orthography to semantics for ambiguous words should have delayed meaning determination, as all three accounts suggest.

The next question, therefore, was which of these explanations of the ambiguity advantage in the lexical decision task is the most viable. Kawamoto et al. (1994) suggested that the advantage is due to the stronger connections among orthographic units, whereas Borowsky and Masson (1996) suggested that the ambiguity advantage is a proximity effect due to having multiple basins of attraction in the semantic level for ambiguous words. Alternatively, according to Hino and Lupker (1996), the ambiguity advantage in lexical decision is due to enhanced semantic feedback for ambiguous words.

Experiments 3 and 4

If, as suggested by Hino and Lupker (1996), semantic feedback does exist and the ambiguity advantage in lexical decision is due to semantic feedback to the orthographic level, the impact of this semantic feedback should be modulated by the nature of the relationships between semantics and orthography. In particular, as discussed above, orthographic competition would be created by semantic-feedback activation if there were one-to-many feedback relationships from semantics to orthography. Thus, a processing disadvantage is expected at the orthographic level if a word has synonyms. That is, the orthographic processing required for making lexical decisions would be slower for words with synonyms than for words with no synonyms.

At the same time, however, this synonymy disadvantage should not be observed in a semantic categorization task according to the feedback account. Because a semantic categorization task would require participants to complete meaning determination, task performance should really be sensitive only to the activation at the semantic level and not to the problems created by feedback to the orthographic level. As such, there would be no reason to expect synonymy to affect semantic categorization performance.

Alternatively, without assuming semantic feedback, that is, if ambiguity effects were explained in a different fashion as in the account offered by Borowsky and Masson (1996), there would be no reason to expect a synonymy disadvantage in any task. Further, consistent with the feedback account, the simulations using Kawamoto et al.'s (1994) model predict a synonymy disadvantage in lexical decision because of the one-to-many feedback relationships for synonyms. However, this model also produced a synonymy disadvantage in terms of the time needed to settle on a semantic code, therefore it predicts a synonymy disadvantage in semantic categorization as well. As such, in order to discriminate between the feedback account and these alternative accounts, we examined the effects of synonymy for ambiguous and unambiguous words using lexical decision and semantic categorization tasks.

Method

Participants. Fifty undergraduate students from Chukyo University participated in these experiments for course credit. Twenty-six students participated in Experiment 3, and the rest participated in Experiment 4. All were native Japanese speakers who had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Stimuli. From the 240 Katakana words that were used in the NOM rating tasks in Experiments 1 and 2, 57 ambiguous and 98 unambiguous words with nonliving-object meanings were selected. The NOM ratings for the 57 ambiguous words were all greater than or equal to 1.5, whereas the ratings for the 98 unambiguous words were all less than 1.25. These 155 Katakana words were randomly ordered and listed in a questionnaire. Thirty-two participants were asked to count and record the number of synonyms (words possessing the same meanings) for these words.

The number of synonyms (NOS) factor was manipulated on the basis of these NOS counts. Because the NOS ratings were highly correlated with the NOM ratings ($r = .73$), ambiguous words were divided into the medium-NOS and large-NOS groups, whereas unambiguous words were divided into the small-NOS and medium-NOS groups, respectively. Words were classified as small-NOS if the NOS ratings were less than 0.7, medium-NOS if the NOS ratings were more than 0.7 but less than 1.2, and large-NOS if the NOS ratings were more than 1.2. Thus, the NOS ratings were significantly higher for words in the large-NOS group (1.55 on average) than for words in the medium-NOS group (1.01 on average) in the ambiguous word condition, $t(28) = 7.24$, $p < .05$. Similarly, the NOS ratings were significantly higher for words in the medium-NOS group (0.94 on average) than for words in the small-NOS group (0.45 on average) in the unambiguous word condition, $t(28) = 10.75$, $p < .05$. However, the NOS ratings were comparable between the ambiguous and unambiguous medium-NOS groups, $t(28) = 1.26$.

Each word group consisted of 15 Katakana words with nonliving-object meanings. The statistical characteristics of these words are given in Table 3. All these words were three or four characters in length and consisted of three or four syllables (moras). Across the four word groups, word frequency, word length, syllabic length, orthographic neighborhood size, experiential familiarity ratings, and typicality ratings were equated as much as possible (all Fs < 1.0). In addition, NOM ratings were equated between the two ambiguous word groups, $t(28) = 0.72$, and between the two unambiguous word groups, $t(28) = 1.26$, respectively. Meaning-type ratings were also equated as much as possible between the two ambiguous word groups, $t(28) = 0.00$. The experiential familiarity ratings, typicality ratings, NOM ratings, and meaning-type ratings used in selecting these items were those collected in Experiments 1 and 2. These words and their English translations are listed in Appendix B.

In addition to these experimental word stimuli, 60 Katakana nonwords were created by replacing one character from actual Katakana words. These nonwords were used in the lexical decision task of Experiment 3. The mean character length of these nonwords was 3.40, ranging from three to four. The mean syllabic length of these nonwords was 3.37, ranging from three to four. In addition, 60 filler Katakana words with living-object meanings were further selected and used in the semantic categorization task of Experiment 4. The mean word length and syllabic length of these fillers were 3.62 and 3.57, respectively, ranging from two to five. The mean typicality ratings for the 60 fillers was 5.94. In both experiments, the stimulus set consisted of 120 items.

Procedure. In Experiment 3, participants were asked to make a lexical decision to the stimulus appearing at the center of the video monitor. In Experiment 4, participants were asked to decide whether a word appearing on the video monitor falls into living-object category or nonliving-object category. In both experiments, 16 practice trials were given prior to the 120 experimental trials. The practice items were 8 Katakana words and 8 Katakana nonwords in Experiment 3 and 8 Katakana words with living-object meanings and 8 Katakana words with nonliving-object meanings in Experiment 4. None of these items were used in the experimental trials. In all other respects, the procedure in Experiment 3 was identical to that in Experiment 1, and the procedure in Experiment 4 was identical to that in Experiment 2.

Results

Experiment 3 (lexical decision task). As in the previous experiments, outliers were excluded from the statistical analyses on the basis of Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion. A total of 78 data points (2.50%) were excluded in this fashion. Mean lexical decision latencies for correct responses and mean error rates were calculated across subjects. The mean lexical decision latencies and percentage errors are presented in Table 4. Planned comparisons were used to examine the differences between the two ambiguous word groups, between the two unambiguous word groups, and between the ambiguous and unambiguous word groups with medium-NOS, respectively.

The 16-ms difference in lexical decision latencies between the ambiguous words with medium-NOS and the ambiguous words with large-NOS was significant, $t(25) = 3.49$, although the 0.54% difference in percentage errors was not, $t(25) = 0.50$. The 14-ms difference in lexical decision latencies between the unambiguous words with small-NOS and the unambiguous words with medium-NOS was also significant, $t(25) = 2.30$, as was the 3.69% difference in percentage errors, $t(25) = 2.32$. As such, lexical decision latencies were slower for words with more synonyms than for words with fewer synonyms in both the ambiguous-word and unambiguous-word conditions. In addition, lexical decision responses were less accurate for words with more synonyms than for words with fewer synonyms in the unambiguous-word condition. Further, for medium-NOS words, the 29-ms ambiguity advantage in lexical decision latencies, $t(25) = 4.85$, and the 4.49% advantage in percentage errors, $t(25) = 3.66$, were both significant.

Experiment 4 (semantic categorization task). Outliers were again excluded from the statistical analyses on the basis of Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion. A total of 72 data points (2.50%) were excluded in this fashion. Mean response latencies for correct responses and mean error rates were calculated across subjects. The mean response latencies and percentage errors are also presented in Table 4. As in Experiment 3, planned comparisons were used to examine the differences between the two ambiguous word groups, between the two unambiguous word groups, and between the ambiguous and unambiguous word groups with medium-NOS, respectively.

Neither the 6-ms difference in response latencies, $t(23) = 0.75$, nor the 1.61% difference in percentage errors, $t(23) = 0.86$, was significant between the ambiguous words with medium-NOS and the ambiguous words with large-NOS. Similarly, neither the 6-ms difference in response latencies, $t(23) = 0.60$, nor the 1.17% difference in percentage errors, $t(23) = 0.85$, was significant between the unambiguous words with small-NOS and the unambiguous words with medium-NOS. However, the 24-ms difference

Table 3
Stimulus Characteristics for the Word Stimuli in Experiments 3–6

Ambiguity and NOS	Freq.	WL	SL	N	FAM	NOM	T	MT	NOS
Unambiguous									
Small	12.80	3.40	3.40	4.47	4.75	1.05	1.85		0.45
Medium	10.20	3.40	3.33	3.33	4.76	1.08	1.84		0.94
Ambiguous									
Medium	11.80	3.40	3.40	4.67	4.75	1.67	1.83	0.54	1.01
Large	9.93	3.40	3.33	4.00	4.76	1.70	1.85	0.54	1.55

Note. Mean typicality rating for the 60 filler Katakana words used in Experiment 4 was 5.94. NOS = number-of-synonyms rating; Freq. = mean word frequency; WL = word length; SL = syllabic length; N = orthographic neighborhood size; FAM = experiential familiarity rating; NOM = number-of-meanings rating; T = typicality rating; MT = meaning-type rating.

Table 4
Response Latencies in ms and Percentage Errors (PEs) for the Four Word Groups in Experiments 3 and 4

	Synonymy						
	Small		Medium		Large		
Ambiguity	RT	PE	RT	PE	RT	PE	Synonymy effect
Experiment 3 (lexical decision task)							
Ambiguous							
M			543	3.85	559	4.43	−16
SEM			10.39	0.66	12.92	1.11	
Unambiguous							
M	558	4.65	572	8.34			−14
SEM	10.88	1.04	12.74	1.31			
Ambiguity effect			+29				
Experiment 4 (semantic categorization task)							
Ambiguous							
M			713	7.30	707	5.69	+6
SEM			18.10	1.79	18.49	1.26	
Unambiguous							
M	695	8.81	689	7.64			+6
SEM	17.55	1.61	17.50	1.36			
Ambiguity effect			−24				

Note. In Experiment 3, mean lexical decision latency and percentage error for the 60 nonwords were 630 ms (*SEM* = 16.06) and 7.37% (*SEM* = 1.07), respectively. In Experiment 4, both the ambiguous and unambiguous words fell into the nonliving-object category. Mean response latency and percentage error for the 60 fillers, which fell into the living-object category, were 638 ms (*SEM* = 12.76) and 8.34% (*SEM* = 1.12), respectively. RT = reaction time.

in response latencies was significant between the unambiguous words with medium-NOS and the ambiguous words with medium-NOS, $t(23) = 2.62$, although the 0.34% difference in percentage errors was not, $t(23) = 0.21$. Thus, as in Experiment 2, an ambiguity disadvantage was observed in the semantic categorization task.

Discussion

Consistent with the results of Experiments 1 and 2, in the medium-NOS condition, an ambiguity advantage was observed in the lexical decision task (Experiment 3), whereas an ambiguity disadvantage was observed in the semantic categorization task (Experiment 4). More important, lexical decision latencies were slower for words with more synonyms than for words with fewer synonyms in both the ambiguous and unambiguous conditions. As such, we successfully replicated Pecher's (2001) results. In the semantic categorization task, however, performance was not affected by the number of synonyms factor.

These results are, therefore, consistent with the feedback account (while being inconsistent with the alternative accounts). As noted above, assuming that lexical decision tasks require participants to engage in orthographic processing such as decision-making operations based on orthographic familiarity of stimuli, lexical decision performance should be affected by the nature of feedback relationships from semantics to orthography. That is, when a word possesses one-to-many feedback relationships (i.e., words with many synonyms), the initially activated semantic code would provide feedback activation to multiple orthographic codes (i.e., orthographic codes for all the synonyms). As a result, orthographic competition would be created, and hence, responding would be delayed. When a word possesses a one-to-one feedback relationship from semantics to orthography (i.e., words with few synonyms or, ideally, words with no synonyms), the initially activated semantic code would produce feedback activation only to the orthographic code for that word. Because there would be less orthographic competition, orthographic processing would be faster for words with few synonyms than for words with many synonyms. As such, the synonymy disadvantage in lexical decision can be easily explained in terms of the feedback account.

In the semantic categorization task, however, no such effects were expected. In order to determine whether a word falls into a specific semantic category (e.g., the living object category), participants would be required to complete meaning determination. Thus, task performance should be sensitive only to the nature of feedforward relationships from orthography to semantics and not to the nature of feedback relationships. Therefore, task performance would be unaffected by the one-to-many feedback relationships from semantics to orthography for words with synonyms, resulting in no synonymy effect in the semantic categorization task.

An additional point that should be made here is that these results provide a possible explanation for why, at times, ambiguity effects have been difficult to find (e.g., Clark, 1973; Forster & Bednall, 1976; Gernsbacher, 1984). As previously noted, the NOS ratings were highly correlated with the NOM ratings ($r = .73$). That is, there was a strong tendency for words with more meanings to have more synonyms. However, whereas the ambiguity factor facilitates lexical decision performance, the synonymy factor produces an inhibitory effect on lexical decision performance. Given that the synonymy factor was not controlled in any of the previous examinations of ambiguity effects, previous failures to observe significant ambiguity effects may have been, at least in part, because of the fact that any positive effects of ambiguity were counteracted by the inhibitory effects of synonymy.

Experiments 5 and 6

The synonymy disadvantage in the lexical decision task provides support for our claim that orthographic processing (and, hence, lexical decision performance) is affected by semantic-feedback activation and that the ambiguity advantage in lexical decision tasks is due to stronger semantic feedback for ambiguous words. In addition, Hino and Lupker (1996) also reported an ambiguity advantage in naming and suggested that this result was also due to semantic feedback, specifically feedback to the phonological level. If so, a similar synonymy disadvantage would also be expected in a naming task. In fact, Pecher (2001) reported a synonymy disadvantage in her naming experiment, although the degree of spelling-to-sound consistency was not controlled in her stimulus set.

In order to address this issue, Experiment 5 was a naming task using the same Katakana words used in Experiments 3 and 4. Because the initial sounds were not equated across word groups, a delayed-naming task was used in Experiment 6 to examine whether articulation onset differences across word groups had any effect on naming latencies. Because our stimuli were all Japanese Katakana words, their character-to-sound relationships were consistent. Thus, any effects observed in this task cannot be attributed to differences in the degree of character-to-sound consistency across word groups.

Method

Participants. Forty-eight undergraduate students from Chukyo University participated in these experiments for course credit. Twenty-four students participated in a standard naming task in Experiment 5. The rest participated in a delayed-naming task in Experiment 6. 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-object meanings used in Experiments 3 and 4.

Procedure. In Experiment 5, participants were asked to name a word aloud into a microphone when it appeared on the video monitor, as quickly and as accurately as possible. The microphone was connected to a voice key interfaced to the computer. Naming latency was measured from the onset of the word stimulus to the onset of the vocal response.

In Experiment 6, participants were instructed that a word would appear on the video monitor and that after a delay it would be surrounded by brackets. They were asked to name the word aloud as quickly and as accurately as possible as soon as the brackets appeared (brackets appeared after 1,500 ms). On each trial, a fixation point and a word stimulus were presented in the same manner as in Experiments 3 and 4. The delayed-naming latency was measured from the onset of the brackets to the onset of the participant's vocal response.

In both experiments, the vocal response terminated the stimulus presentation. During the experimental trials, an experimenter was in a different room than the participant but was able to check the participants' vocal responses through audio–video monitors to record errors. Sixteen practice trials (involving stimuli not used in the experiment proper) were given prior to the 60 experimental trials. In all other respects, the procedure was identical to that of Experiments 3 and 4.

Results

Experiment 5 (standard naming task). A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some extraneous sound triggered the voice key. There were five (0.35%) mechanical errors. These mechanical errors were excluded from the data analyses. As in the previous experiments, outliers were excluded from the statistical analyses on the basis of Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion; 29 data points (2.01%) were removed in this fashion. Mean naming latencies for correct responses and mean percentage errors were calculated across subjects. The mean naming latencies and percentage errors are presented in Table 5.

As in Experiments 3 and 4, planned comparisons were used to examine the differences between the two ambiguous word groups, between the two unambiguous word groups, and between the ambiguous and unambiguous word groups with medium-NOS. The 2-ms difference in naming latencies was not significant between the ambiguous words with medium-NOS and the ambiguous words with large-NOS, $t(23) = 0.56$, nor was the 0.28% difference in percentage errors, $t(23) = 0.44$. Naming latencies, $t(23) = 0.30$, and percentage errors, $t(23) = 1.40$, were also comparable for the unambiguous words with small-NOS and the unambiguous words with medium-NOS. However, the 13-ms difference in naming latencies was significant between the unambiguous words with medium-NOS and the ambiguous words with medium-NOS, $t(23) = 3.70$, although the percentage errors were comparable, $t(23) = 1.12$.

Experiment 6 (delayed-naming task). A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some extraneous sound triggered the voice key. There were 11 mechanical errors (0.92%). These mechanical errors were excluded from the data analyses. In addition, outliers were excluded from the statistical analyses on the basis of Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion. Thus, 31 additional data points (2.15%) were removed from the data analyses. Mean delayed-naming latencies for correct responses and mean percentage errors were calculated across subjects. The mean delayed-naming latencies and percentage errors are presented in Table 5.

In the analyses of delayed-naming latencies, the 2-ms difference between the two ambiguous word groups was not significant, $t(23) = 0.44$, nor was the 3-ms difference between the two unambiguous word groups, $t(23) = 0.56$. In addition, the 3-ms difference between the ambiguous words with medium-NOS and the unambiguous words with medium-NOS was not significant, $t(23) = 0.67$. Percentage errors were too low to conduct statistical analyses.

Table 5
Mean Naming Latencies in ms and Percentage Errors (PEs) for the Four Word Groups in Experiments 5 and 6

	Synonymy						
	Small		Medium		Large		
Ambiguity	RT	PE	RT	PE	RT	PE	Synonymy effect
Experiment 5 (standard naming task)							
Ambiguous							
M			457	1.41	459	1.13	+2
SEM			10.28	0.57	11.82	0.53	
Unambiguous							
M	470	1.47	470	0.58			0
SEM	11.85	0.60	11.37	0.40			
Ambiguity effect			+13				
Experiment 6 (delayed-naming task)							
Ambiguous							
M			341	0.00	343	0.00	−2
SEM			14.32	1.79	16.43	0.00	
Unambiguous							
M	347	0.28	344	0.28			+3
SEM	14.29	0.28	15.70	0.28			
Ambiguity effect			+3				

Note. RT = reaction time.

Discussion

Delayed-naming performance was comparable across word groups. These results appear to indicate that it is unlikely that the performance in the standard naming task was affected by articulation-onset differences across word groups.

In the standard naming task, a small ambiguity advantage was observed, which was consistent with the results in Gottlob et al. (1999), Hino and Lupker (1996), Hino et al. (1998), and Lichacz et al. (1999). However, no synonymy disadvantage was observed for either ambiguous words or unambiguous words. Thus, contrary to Hino and Lupker's (1996) feedback account, these results suggest that semantic-feedback activation does not affect phonological processing (and, hence, naming performance).

A point to keep in mind, however, is that there may be a sensitivity problem in this experiment. Note, for example, that Strain, Patterson, and Seidenberg (1995) reported that semantic-feedback activation affects naming performance only for slowly named stimuli. That is, Strain et al. observed a significant imageability effect in naming only for low-frequency inconsistent words. Strain et al.'s account of their imageability effect was also based on semantic feedback. In particular, imageable words are presumed to be represented more richly at the semantic level. Thus, more semantic units would be activated for imageable words than for less imageable words. As a result, semantic-feedback activation would be greater for imageable words than for less imageable words, producing an imageability advantage. However, this advantage would be limited to situations in which the orthographic-to-phonological conversion is slow. That is, when the orthographic-to-phonological conversion is fast enough, the pronunciation response could be initiated before semantic feedback noticeably affected phonological processing. Thus, the imageability effect would be limited to the slowest stimuli (i.e., low-frequency inconsistent words).

If Strain et al.'s (1995) suggestion is correct, it is quite possible that the lack of a synonymy effect in our naming experiment is due to the fact that our stimuli were Katakana words, which are words from a shallow orthography with completely consistent character-to-sound relationships. Because the orthographic-to-phonological conversion is fast for Katakana words, the impact of semantic feedback, and hence the size of all effects, would be limited. Note, for example, that the ambiguity-effect size was much smaller in the naming task (13-ms effect) than in the lexical decision task (29-ms effect).

If, as we suggest, the lack of a synonymy effect in the naming task was due to the fast orthographic-to-phonological conversion for our Katakana words, it should be possible to observe a significant synonymy effect if we were to use stimuli that take longer to name (i.e., stimuli with less consistent character-to-sound relationships). Fortunately, there is a set of Japanese words of this sort. The Japanese writing system consists of two types of scripts that differ in terms of their character-to-sound relationships: Kana (including Katakana and Hiragana scripts) and Kanji. Kana is considered a shallow orthography because each Kana character corresponds to a single syllable (mora). As a result, character-to-sound correspondences are consistent when words are written in Kana. On the other hand, because each Kanji character generally possesses multiple pronunciations (i.e., On-reading and Kun-reading pronunciations), Kanji is considered a deep orthography. Thus, the character-to-sound relationships are much less consistent for words written in Kanji than for words written in Katakana, and hence naming latencies are generally slower for Kanji words than for Katakana words. For example, even when the number of syllables, word frequency, and experiential familiarity were equated, Hino and Lupker (1998) reported much slower naming latencies for Kanji words (671 ms for their low-frequency Kanji words) than for Katakana words (528 ms for their low-frequency Katakana words).

The fact that orthographic-to-phonological conversion takes longer for Kanji words leads to the prediction that there would be more chance of observing semantic-feedback effects, particularly synonymy effects, when participants are naming Kanji words. Thus, in Experiment 7, a naming task was conducted using Kanji words in order to look for a synonymy effect. Once again, because it was difficult to equate the initial sounds across word groups, a delayed-naming task was also conducted using the same Kanji words in Experiment 8.

Because each Kanji character generally denotes a specific meaning, it was difficult to find Kanji words with multiple meanings. As such, we found it impossible to simultaneously manipulate ambiguity and synonymy in these experiments. In order to ensure that ambiguity was at least controlled across the synonymy conditions, we collected NOM ratings for these Kanji words.

Experiments 7 and 8

Method

Participants. Forty-eight undergraduate students from Chukyo University participated in these experiments for course credit. Twenty-four students participated in each experiment. All were native Japanese speakers who had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Stimuli. Two hundred six 2-character Kanji words were collected from word-frequency norms for Japanese words (National Language Research Institute, 1970). The frequency counts of these words were all less than 50 per 940,533. One hundred forty of these Kanji words possessed nonliving-object meanings, whereas the rest of the Kanji words had living-object meanings.

For these 206 Kanji words, experiential familiarity ratings and typicality ratings were collected. To obtain the experiential familiarity ratings, we asked 42 participants to rate the experiential familiarity of these words. The 206 Kanji 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 experiential familiarity by circling the appropriate number on the scale.

For the typicality ratings, another 42 participants were asked to rate how typical the meanings of these words are as a member of living-object or nonliving-object categories. The 206 Kanji words were, once again, randomly ordered and listed in a questionnaire. In the questionnaire, each word was accompanied by a 7-point scale ranging from *nonliving object* (1) to *neutral* (4) to *living object* (7). The participants were asked to rate the typicality as a member of living- or nonliving-object category by circling the appropriate number on the scale.

In addition, in order to manipulate the synonymy factor for these Kanji words, we collected NOS ratings. Another 42 participants were asked to count and record the number of synonyms (words possessing the same meanings) for these 206 Kanji words.

To control ambiguity across synonymy conditions, we also collected NOM ratings. In addition to the 206 Kanji words, 50 two-character Kanji nonwords were created by randomly pairing two Kanji characters. These 256 stimuli were randomly ordered and listed in a questionnaire. Each item was accompanied by a 3-point scale ranging from 0 (*no meaning*) to 1 (*only one meaning*) to 2 (*more than one meaning*). The participants were asked to circle the appropriate number on the scale.

On the basis of these ratings, 20 Kanji words with small-NOS and 20 Kanji words with large-NOS were selected from the 140 Kanji words with nonliving-object meanings. The statistical characteristics of the two word groups are given in Table 6. The NOS ratings for the Kanji words with small-NOS were all less than 0.60, whereas the NOS ratings for the Kanji words with large-NOS were all more than 0.80, so that the mean rating for the Kanji words with large-NOS (1.08) was significantly larger than the mean rating for the Kanji words with small-NOS (0.38), $t(38) = 8.89$, $p < .05$. Word length and the number of syllables (moras) were identical for the two word groups. In addition, word frequency, $t(38) = 1.05$; orthographic neighborhood size, $t(38) = 0.07$; experiential familiarity ratings, $t(38) = 0.23$; typicality ratings, $t(38) = 1.07$; and the NOM ratings, $t(38) = 0.81$, were equated as much as possible between the two word groups. Summed Kanji-character frequency was computed for each Kanji word on the basis of norms from the National Language Research Institute (1963). The summed Kanji-character frequency was also equated as much as possible between the two word groups, $t(38) = 0.61$.

Further, because naming latencies for Kanji words are affected by the degree of character-to-sound consistency (e.g., Fushimi, Ijuin, Patterson, & Tatsumi, 1999), summed frequency for phonological friends and summed frequency for phonological enemies were computed for each Kanji word. For each Kanji word, its orthographic neighbors were first generated, and then each neighbor was classified as a phonological friend or enemy on the basis of whether the shared Kanji character in the neighbor is pronounced the same as that in the target Kanji word. When the same Kanji character is pronounced identically to that in the target Kanji word, the neighbors were classified as phonological friends, whereas when the same Kanji character is pronounced in different ways, these neighbors were classified as phonological enemies. The summed frequencies for these friends and enemies were computed on the basis of norms from the National Language Research Institute (1970). To ensure that Kanji words in the two word

Table 6
Stimulus Characteristics for the Word Stimuli in Experiments 7–10

Word group	Freq.	WL	SL	N	KCF	PF	PE	FAM	NOM	T	NOS
Small NOS	6.45	2.0	3.4	47.00	591.15	351.50	64.15	3.94	1.03	2.00	0.38
Large NOS	6.95	2.0	3.4	46.30	522.05	361.05	60.40	4.01	1.05	2.10	1.08

Note. Freq. = mean word frequency; WL = word length; SL = syllabic length; N = orthographic neighborhood size; KCF = summed Kanji character frequency; PF = summed frequency of phonological friends; PE = summed frequency of phonological enemies; FAM = experiential familiarity rating; NOM = number-of-meanings rating; T = typicality rating; NOS = number-of-synonyms rating.

groups possess similar degrees of character-to-sound consistency, we equated summed frequency of phonological friends, $t(38) = 0.09$, and summed frequency of phonological enemies, $t(38) = 0.12$, as much as possible between the two word groups. The 40 Kanji words are listed in Appendix C along with their English translations.

Procedure. A standard naming task was conducted using the 40 Kanji words in Experiment 7. The procedure of this experiment was identical to that in Experiment 5. That is, participants were asked to name a Kanji word aloud into a microphone when it appeared on the video monitor, as quickly and as accurately as possible. Experiment 8 was a delayed-naming task using the same 40 Kanji words. The procedure was identical to that in Experiment 6. Participants were asked to name a Kanji word aloud as quickly and as accurately as possible as soon as it was surrounded by brackets, which were presented 1,500 ms after the onset of the Kanji word.

Results

Experiment 7 (standard naming task). A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some extraneous sound triggered the voice key. There were two (0.21%) mechanical errors, and these were excluded from the data analyses. In addition, outliers were excluded from the data analyses using Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion: 26 additional data points (2.71%) were excluded in this fashion. Mean naming latencies for correct responses and mean percentage errors were calculated across subjects. The mean naming latencies and percentage errors are presented in Table 7.

Mean naming latencies were 21 ms slower for Kanji words with large-NOS than for Kanji words with small-NOS, which was significant, $t(23) = 2.07$. Percentage errors were 2.44% higher for Kanji words with large-NOS than for Kanji words with small-NOS, which was marginally significant, $t(23) = 1.77$, $p < .10$.

Experiment 8 (delayed-naming task). A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some extraneous sound triggered the voice key. There were five mechanical errors (0.52%), and these were excluded from the data analyses. In addition, outliers were again excluded from the data analyses using Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion: 21 additional data points (2.19%) were removed from the data analyses. Mean delayed-naming latencies for correct responses and mean percentage errors were calculated across subjects. The mean delayed-naming latencies and percentage errors are presented in Table 7. Both the delayed-naming latencies, $t(23) = 0.03$, and percentage errors, $t(23) = 1.04$, were comparable between the two word groups.

Discussion

In the standard naming task of Experiment 7, naming latencies for Kanji words (662 ms on average) were much slower than those for Katakana words in Experiment 5 (470 ms on average for unambiguous words). More important, a significant synonymy disadvantage was observed for Kanji words in the standard naming task but not in the delayed-naming task. Because the synonymy manipulation produced no effect on delayed-naming performance, it is unlikely that the synonymy effect in the standard naming task was due to articulation-onset differences between word groups. In addition, because the degree of character-to-sound consistency was also equated across word groups, it is unlikely that the observed effect was confounded with consistency. Thus, these results support the claim that semantic-feedback activation affects phonological processing when the orthographic-to-phonological conversion is relatively slow.

Table 7
Mean Response Latencies in ms and Percentage Errors (PEs) for Words With Small Number of Synonyms and Words With Large Number of Synonyms in Experiments 7–10

	Synonymy				
	Small		Large		
Task Type	RT	PE	RT	PE	Synonymy effect
Experiment 7					
Standard naming					
M	651	3.36	672	5.80	−21
SEM	13.97	0.72	17.21	1.38	
Experiment 8					
Delayed naming					
M	335	0.84	335	1.50	0
SEM	12.39	0.39	11.49	0.57	
Experiment 9					
Lexical decision					
M	569	5.94	582	8.56	−13
SEM	10.29	0.94	12.42	1.28	
Experiment 10					
Semantic categorization					
M	612	1.28	610	1.96	+2
SEM	18.81	0.55	17.99	0.69	

Note. In Experiment 9, mean lexical decision latency and percentage error for the 40 nonwords were 666 ms (*SEM* = 19.43) and 8.25% (*SEM* = 1.02), respectively. In the semantic categorization task of Experiment 10, experimental words (words with small NOS and words with large NOS) fall into the nonliving-object category. Mean response latency and percentage error for the 40 fillers were 570 ms (*SEM* = 15.28) and 5.66% (*SEM* = 0.74), respectively. RT = reaction time; NOS = number of synonyms.

Experiments 9 and 10

If the effect observed for Kanji words in the standard naming task really was due to the synonymy manipulation and to semantic-feedback activation, a similar effect would also be expected in a lexical decision task but not in a semantic categorization task. To evaluate this possibility, we conducted a lexical decision task (Experiment 9) and a semantic categorization task (Experiment 10) using the same Kanji words.

Method

Participants. Fifty-two undergraduate students from Chukyo University participated in these experiments for course credit. Twenty-eight students participated in Experiment 9, and the rest participated in Experiment 10. 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 40 two-character Kanji words used in Experiments 7 and 8. In addition, 40 Kanji nonwords were created by randomly pairing two Kanji characters each with single pronunciation to use in the lexical decision task in Experiment 9. The number of syllables for these Kanji nonwords was matched with those of the 40 Kanji words. In Experiment 10, in addition to the 40 Kanji words, 40 two-character Kanji words with living-object meanings were selected as fillers from the 66 Kanji words used in the ratings. The syllabic lengths of these fillers were matched with those of the 40 experimental Kanji words. The mean typicality rating for the fillers was 6.40.

Procedure. A lexical decision task was conducted in Experiment 9. Participants were asked to make a word–nonword discrimination for a Kanji character string appearing at the center of the video monitor. A semantic categorization task was conducted in Experiment 10. Participants were asked to decide whether a Kanji word appearing on the video monitor falls into the living-object category or the nonliving-object category. In both experiments, 20 practice trials were given prior to the 80 experimental trials. The practice items were 10 Kanji words and 10 Kanji nonwords in Experiment 9, and 10 Kanji words with living-object meanings and 10 Kanji words with nonliving-object meanings in Experiment 10. None of these items were used in the experimental trials. In all other respects, the procedure in Experiment 9 was identical to that in Experiments 1 and 3. The procedure in Experiment 10 was identical to that in Experiments 2 and 4.

Results

Experiment 9 (lexical decision task). As in the previous experiments, outliers were excluded from the data analyses using Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion: 56 data points (2.50%) were excluded in this fashion. Mean lexical decision latencies for correct responses and mean percentage errors were calculated across subjects. The mean lexical decision latencies and percentage errors are presented in Table 7.

Mean lexical decision latencies were 13 ms slower for Kanji words with large-NOS than for Kanji words with small-NOS, a difference that was significant, $t(27) = 2.10$. Percentage errors were 2.62% higher for Kanji words with large-NOS than for Kanji words with small-NOS, which was marginally significant, $t(27) = 1.71$, $p < .10$.

Experiment 10 (semantic categorization task). Outliers were once again excluded from the data analyses using Van Selst and Jolicoeur's (1994) nonrecursive 2.5-standard-deviations cutoff procedure with moving criterion: 51 data points (2.66%) were excluded in this fashion. Mean response latencies for correct responses and mean percentage errors were calculated across subjects. The mean response latencies and percentage errors are presented in Table 7.

Response latencies and percentage errors were comparable between the two word groups. Thus, neither the 2-ms difference in response latencies, $t(23) = 0.41$, nor the 0.68% difference in percentage errors, $t(23) = 0.71$, was significant.

Discussion

Consistent with our expectations, a significant synonymy disadvantage was again observed in the lexical decision task but not in the semantic categorization task. Together with the results of Experiments 7 and 8, these data provide support for our claim that feedback activation from semantics is the cause of these synonymy effects and that the impact of semantic-feedback activation is determined by the nature of feedback relationships from semantics. Orthographic processing (and, hence, lexical decision performance) is influenced by semantic-feedback activation and, therefore, by the nature of the feedback relationships from semantics to orthography. Phonological processing (and, hence, naming performance) is also affected by semantic-feedback activation and, therefore, by the nature of feedback relationships from semantics to phonology. However, the effect of semantic-feedback activation in naming experiments is detectable only if the orthographic-to-phonological conversion is relatively slow.

General Discussion

A now standard finding in the literature is that reaction times are faster for words with multiple meanings than for words with fewer meanings in lexical decision and naming tasks (e.g., Gottlob et al., 1999; Hino & Lupker, 1996; Hino et al., 1998; Jastrzembski, 1981; Kellas et al., 1988; Lichacz et al., 1999; Millis & Button, 1989; Pexman & Lupker, 1999; Rubenstein et al., 1970). Early accounts of these effects were based on classical lexical models with ambiguity effects being assumed to arise during the lexical selection process (e.g., Balota et al., 1991; Jastrzembski, 1981; Rubenstein et al., 1970).

If the lexical selection process is involved in both lexical decision and naming tasks and if ambiguity effects are due simply to lexical selection, the implication is that ambiguity effects should be similar across different tasks. Hino and Lupker (1996), however, reported different patterns of ambiguity effects in their lexical decision and naming tasks. In addition, in their go/no-go naming task, ambiguity effect sizes were similar to the sum of the effect sizes in lexical decision and naming tasks. On the basis of these results, Hino and Lupker (1996) argued that it is unlikely that these ambiguity effects can be explained within a classical lexical framework and, hence, suggested a nonlexical account of ambiguity effects based on a PDP framework.

Hino and Lupker (1996) explained ambiguity effects as being due to semantic-feedback activation. Assuming that lexical decisions are made primarily on the basis of orthographic processing, whereas phonological coding plays the central role in naming, Hino and Lupker (1996) suggested that ambiguity effects in lexical decision are due to feedback activation from the semantic level to the orthographic level, whereas the effects in naming are due to feedback activation from the semantic level to the phonological level. Because ambiguous words activate multiple semantic codes, ambiguous words would produce a greater amount of semantic activation than would unambiguous words. Thus, the amount of feedback activation from the semantic level to the orthographic or phonological level would be greater for ambiguous words. As a result, both the orthographic processing required in making lexical decisions and the phonological coding required in naming would receive more support from semantic feedback for ambiguous words than for unambiguous words.

Kawamoto et al. (1994) suggested a different PDP-based account of ambiguity effects. Ambiguous words involve inconsistent one-to-many mappings between orthography and semantics, and thus the orthographic-to-semantic connections are weaker for ambiguous words than for unambiguous words in their model. These weaker connections are compensated for by having stronger connections for ambiguous words at the orthographic level. As a result, orthographic processing would be faster for ambiguous words than for unambiguous words because of these stronger connections among orthographic units.

Borowsky and Masson (1996) proposed a third PDP-based account of ambiguity effects. According to Borowsky and Masson, the network moves into a basin of attraction (and, hence, into more negative energy values) more quickly for ambiguous words. The reason is that there are multiple basins of attraction for ambiguous words, whereas there is only a single basin of attraction for unambiguous, single meaning, words. Thus, the distance from the initial random state to a basin of attraction should, on average, be smaller for ambiguous words than for unambiguous words.

The purpose of the present research was to evaluate these accounts by contrasting Hino and Lupker's (1996) feedback account with the other two accounts. All three accounts were based on PDP frameworks and, as Joordens and Besner (1994) argued, these accounts predict that the speed of semantic coding should be slower for ambiguous words than for unambiguous words because ambiguous words involve one-to-many feedforward mappings from orthography to semantics. Thus, ambiguous words should show a disadvantage in any semantically based task. To examine this prediction, we first conducted lexical decision and semantic categorization tasks using the same ambiguous and unambiguous words. The results were consistent with the prediction. That is, an ambiguity disadvantage was observed in semantic categorization (in Experiments 2 and 4), whereas an ambiguity advantage was observed in lexical decision (in Experiments 1 and 3).

To discriminate among the three accounts, we further examined the effects of synonymy in lexical decision and semantic categorization tasks. According to Hino and Lupker's (1996) feedback account, the effect of semantic-feedback activation should be modulated by the nature of the feedback relationships from semantics to orthography (or phonology). Because ambiguous words involve many-to-one mappings from semantics to orthography, the semantic feedback to the orthographic level would be stronger than that for unambiguous words, words possessing one-to-one feedback mappings. In contrast, if words have synonyms, these words would have one-to-many mappings from semantics to orthography. In such a situation, the activated semantic code would produce feedback activation to multiple orthographic codes, creating orthographic competition and delaying orthographic processing. As such, the feedback account predicts a synonymy disadvantage in a lexical decision task. On the other hand, because semantic-feedback activation to orthography (or phonology) would be irrelevant in making responses based on the result of semantic processing, no synonymy disadvantage would be expected in a semantic categorization task.

Consistent with these predictions, a synonymy disadvantage was observed in the lexical decision task (in Experiments 3 and 9) but not in the semantic categorization task (in Experiments 4 and 10). Thus, these results provide further support for the feedback account.

In contrast, these results were not consistent with Borowsky and Masson's (1996) account. As previously noted, because Borowsky and Masson's model does not postulate semantic feedback to the orthographic level, there is no reason to expect a synonymy disadvantage in a lexical decision task. In addition, because there are multiple basins of attraction for synonyms at the orthographic level, the energy value at the orthographic level might even be reduced faster for words with synonyms than for words without synonyms. Therefore, this model might actually predict a synonymy advantage in a lexical decision task, in contrast to the present results.

Kawamoto et al.'s (1994) model also failed to predict our experimental results, specifically our results in the semantic categorization task. Kawamoto et al.'s model involves semantic-feedback connections to orthographic units, and consequently the orthographic units' activation can be influenced by semantic feedback. Thus, consistent with our lexical decision results, the model did predict a synonymy disadvantage at the orthographic level. The time needed to settle on an orthographic code was slower for words with synonyms than for words without synonyms. At the same time, however, the time to settle on a semantic code was also slower for words with synonyms than for words without synonyms. Thus, in contrast to the present results, this model predicted a synonymy disadvantage in a semantic categorization task.

Our analysis is based on the assumption that the additional lexical activation (due to semantic feedback) for words with synonyms produces competition and, hence, a processing-time disadvantage in lexical decision tasks. In contrast, there is at least one major model (Grainger & Jacobs's, 1996, multiple read-out model) that assumes just the opposite, that additional lexical activation produces a processing-time advantage at least in a certain situation. Grainger and Jacobs (1996) developed their model (based on McClelland & Rumelhart's, 1981, interactive-activation model) in an effort to explain both inhibitory and facilitory effects due to orthographic neighbors in lexical decision tasks. According to Grainger and Jacobs's model, a positive lexical decision can be made either when a single lexical unit is activated over its activation threshold (i.e., their M criterion) or, most relevant to the present discussion, when total lexical activity reaches a certain activation criterion (i.e., their Σ criterion).

The degree to which these two criteria play a role in any given lexical decision task is determined by the nature of the nonwords. When the nonwords are very wordlike, they would activate a number of lexical units (i.e., those units corresponding to their orthographic neighbors). Thus, it would be difficult to discriminate words from nonwords on the basis of total lexical activity. In such a situation, a positive lexical decision could be made only when a single lexical unit had been activated over the threshold (i.e., through the use of the M criterion). A major implication of using the M criterion is that because the activation of a lexical unit is assumed to be inhibited when other lexical units are activated, the existence of higher frequency neighbors should slow lexical activation for the target word (e.g., Grainger, O'Regan, Jacobs, & Segui, 1989). When nonwords are less wordlike, on the other hand, the total lexical activity would be higher for words than for nonwords, and thus total lexical activity would provide a reliable basis for making lexical decisions (i.e., the Σ criterion would be used). A major implication of using the Σ criterion is that words possessing more neighbors should be responded to more rapidly because those words create more total lexical activity than words with few neighbors (e.g., Andrews, 1989; Sears, Hino, & Lupker, 1995).

If the assumption of feedback activation from the semantic units to the lexical units, as suggested by Balota et al. (1991), were incorporated into Grainger and Jacobs's (1996) model, a synonymy disadvantage would be expected when positive decisions were made on the basis of a lexical unit being activated over the threshold (i.e., when nonwords were wordlike). On the other hand, however, this model would predict a synonymy advantage when positive decisions were based on total lexical activity (i.e., when nonwords were less wordlike).

In addition, similar predictions would also be made with respect to homophone effects in lexical decision tasks if the assumption of feedback activation from phonological units to lexical units were incorporated into the model. That is, a homophone disadvantage would be expected when positive decisions were based on a lexical unit's activation, whereas a homophone advantage would be expected when positive decisions were based on total lexical activity. Although Pexman and Lupker (1999) reported that the homophone disadvantage does increase when nonwords are more wordlike (i.e., pseudohomophones) than when nonwords are less wordlike (i.e., pronounceable nonwords), a homophone advantage in the lexical decision task with less wordlike nonwords has never been observed. Indeed, Edwards and Pexman (2001) presented consonant strings as nonwords and still found a modest homophone disadvantage. Similarly, in our lexical decision tasks, we always observed a synonymy disadvantage. Thus, at this point, there is no evidence that either a synonymy advantage or a homophone advantage could be produced in lexical decision tasks.

Hino and Lupker (1996) observed an ambiguity advantage in their naming task, although only for low-frequency words, and hence suggested that phonological coding is also influenced by semantic-feedback activation. If so, it should also be possible to observe a synonymy disadvantage in a naming task. Although we failed to observe a synonymy disadvantage in the naming of Katakana words (in Experiment 5), a significant synonymy disadvantage was observed in the naming of Kanji words (in Experiment 7).

Strain et al. (1995) argued that semantic-feedback effects in naming would be limited to slowly named stimuli because when the orthographic-to-phonological conversion is fast enough, the pronunciation response could be initiated before semantic feedback noticeably affected phonological processing. Katakana is a shallow script with consistent character-to-sound correspondences, whereas Kanji is a deep orthography, and each Kanji character generally possesses at least two pronunciations (i.e., On-reading and Kun-reading pronunciations). Thus, the orthographic-to-phonological conversion would be slower for Kanji words than for Katakana words. In fact, naming latencies were slower for Kanji words in Experiment 7 than for Katakana words in Experiment 5. Therefore, the significant synonymy disadvantage for Kanji words only is consistent with Strain et al.'s suggestions, as is the fact that Hino and Lupker (1996) found an ambiguity effect for low-frequency words only. Thus, we conclude that phonological processing (and, hence, naming performance) is affected by the nature of the feedback relationships from semantics to phonology, although the effects of semantic feedback are limited to stimuli for which the orthographic-to-phonological conversion is relatively slow.

Measures of Ambiguity and Synonymy

A recurring issue in ambiguity research is the question of how one measures ambiguity (i.e., the number of meanings a word has) accurately. We measured ambiguity using subjective ratings rather than counting the number of definitions listed in a dictionary because Gernsbacher (1984) argued that counting the number of definitions listed in an unabridged dictionary (e.g., Jastrzembski, 1981) overestimates the NOM. In fact, Gernsbacher reported that her informal survey revealed that even well-educated participants, on average, could report only 3 definitions for the word *fudge*, 2 for the word *gauge*, and 1 for the word *cadet*, although these words have 15, 30, and 15 dictionary definitions, respectively.

Millis and Button (1989) used three different measures of ambiguity and examined which, if any, of those measures produced a significant ambiguity effect in lexical decision tasks. First, as with Rubenstein et al.'s (1970) procedure, Millis and Button asked participants to write down the first meaning of each word, and they took the total NOM that appeared across participants as a measure of ambiguity (first-meaning metric). Millis and Button also used two other measures that were derived from the task of asking participants to write down all the meanings they could think of for each word: the total NOM generated across participants (total-meaning metric) and the average NOM generated over participants (average-meaning metric). In their lexical decision experiments, ambiguity effects were significant only when the total-meaning metric and the average-meaning metric of ambiguity were used. Thus, Millis and Button's results suggest that the total-meaning and average-meaning metrics are better measures of ambiguity than the first-meaning metric.

Following Millis and Button's (1989) suggestions, for our NOM ratings, we asked participants to count the NOM for each word, and the average NOM across participants was taken as a measure of ambiguity. This type of ambiguity measure has been used in a number of recent studies examining ambiguity effects (e.g., Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino et al., 1998; Kellas et al., 1988; Lichacz et al., 1999). Nonetheless, to further evaluate the validity of our ambiguity measure, we attempted to compare our measure with a measure based on a Japanese dictionary.

As previously noted, Klein and Murphy (2001) suggested that homonymous meanings (i.e., unrelated meanings) and polysemous senses (i.e., related meanings) appear to be represented in a similar manner. Thus, we saw no reason to discriminate between these two concepts and, therefore, simply counted the number of definitions (both meanings and senses) listed in a Japanese dictionary (Kindaichi et al., 1974) for each word used in our experiments. Because some Katakana words were included in two different stimulus sets (see Appendixes A and B), there were 38 ambiguous Katakana words, 43 unambiguous Katakana words, and 40 unambiguous Kanji words in our experimental items. For these 121 items, our NOM ratings were positively correlated with the number of dictionary definitions ($r = .62$, $p < .05$). This result appears to provide additional support for the validity of our NOM ratings.

In addition, we also used subjective ratings for our synonymy measure. That is, we first asked participants to count the number of synonyms for each word in the NOS ratings, and the average number of synonyms across participants was taken as a measure of synonymy. We used the NOS ratings because we could not find any other appropriate measures of synonymy for our stimuli. For example, it would certainly be possible to count the number of entries listed in a thesaurus. However, it is quite likely that this would be an inaccurate measure of the number of synonyms because a thesaurus generally lists not only synonyms but also antonyms, words with related but different meanings, and so on. Nonetheless, because, at least to our knowledge, this is the first attempt to measure the number of synonyms using subjective ratings, it would be important to at least evaluate the validity of our synonymy measure against a more objective measure.

For this purpose, we conducted a survey based on a thesaurus (National Language Research Institute, 1993). For our experimental items used in Experiments 3 (Katakana words) and 7 (Kanji words), 83 items were found in the thesaurus. These items were randomly listed in a questionnaire. In this questionnaire, each experimental item was accompanied by a list of items found in the thesaurus. We asked 30 participants to consider each of the 83 experimental items and to circle the items in the list that have identical meanings. The number of circled items in the list was counted for each experimental item and averaged over participants. We compared these scores with our NOS ratings for the 83 items to evaluate the validity of our synonymy measure based on the NOS ratings. Our NOS ratings were positively correlated with these scores ($r = .46$, $p < .05$). Thus, once again, this result appears to provide support for the validity of our synonymy measure based on the subjective ratings (i.e., NOS ratings).

Feedback Account for Other Semantic Effects

The feedback account of Hino and Lupker (1996) is not only consistent with the present results; it also appears to be able to account for a number of other semantic effects in the literature. First, as previously noted, Strain et al. (1995) explained imageability effects as being due to semantic feedback. Because imageable words are represented by richer semantic information, more semantic units would be activated for imageable words than for less imageable words. Consequently, semantic-feedback activation would be stronger for imageable words, producing an imageability advantage. As such, Strain et al.'s account of imageability effects is consistent with the feedback account.

In addition, Pexman, Lupker, and Hino (in press) recently reported that both lexical decision and naming performances are affected by the number of semantic features a word has. On the basis of the norms provided by McRae, de Sa, and Seidenberg (1997; see also McRae & Cree, in press), Pexman, Lupker, and Hino (in press) manipulated the number of features (NOF) for their word stimuli. In their lexical decision and naming tasks, response latencies were faster for words with high NOF than for words with low NOF. These results are also consistent with the feedback account. That is, when word meanings are represented by more semantic features, more semantic units are activated. As a result, semantic-feedback activation to the orthographic and phonological levels would be greater, producing a processing advantage in both orthographically based and phonologically based tasks. As such, both imageability and NOF effects appear to be easily explained by Hino and Lupker's (1996) feedback account.

On the other hand, recent results reported by Azuma and Van Orden (1997) may be problematic for the feedback account. Azuma and Van Orden examined the effects of NOM and relatedness among meanings for ambiguous words using lexical decision tasks. In their norming study (see also Azuma, 1996), they asked participants to write down all the meanings each word had, and the NOM factor was manipulated on the basis of a count of the NOM provided by these participants. Words with six or more meanings were classified as many-meaning words, whereas words with four or less meanings were classified as few-meaning words. In addition, for each word, the dominant meaning was paired with each of the other (subordinate) meanings, and participants were asked to rate the relatedness of these meanings using a 7-point scale. On the basis of these ratings, words were classified as low-related words when the mean relatedness ratings were less than 3.0 and high-related words when the mean ratings were more than 3.5. When these words were presented with legal nonwords, neither ambiguity nor relatedness affected lexical decision performance. However, when pseudohomophones were used as nonwords, lexical decision latencies were slowest for few-meaning, low-related words and comparable for the other three word types, producing an ambiguity by relatedness interaction.

From Azuma and Van Orden's (1997) standpoint, the important component of these results is that with pseudohomophones there was a relatedness effect for few-meaning words (i.e., few-meaning, low-related words had longer latencies than did few-meaning, high-related words). That is, Azuma and Van Orden suggested that because the relatedness ratings did not take into account relatedness among subordinate meanings, these ratings provide a more reliable relatedness measure for few-meaning words than for many-meaning words. Given that there was a clear relatedness effect for few-meaning words, Azuma and Van Orden argued that relatedness of meanings was the important factor.

Azuma and Van Orden (1997) explained why the relatedness-of-meanings effect (for few meaning words) was observed only with pseudohomophones by assuming that what pseudohomophones do is force participants to rely more on semantic coding. Because high-related meaning words undoubtedly share semantic features, the orthographic-to-semantic mappings should be more consistent for words with high-related meanings. As a result, semantic coding should be faster for words with high-related meanings than for words with low-related meanings. Thus, when pseudohomophones are used and semantic coding is presumed to play a more major role, one would expect to see a processing-time advantage for high-related words (at least in the few-meanings condition).

What should be noted, however, is that this assumption about what pseudohomophones do is inconsistent with other data. For example, Pexman and Lupker (1999) reported that when lexical decision performance was compared when using legal nonwords versus pseudohomophones, ambiguity effects were larger with pseudohomophones. If pseudohomophones really do drive participants to make lexical decisions on the basis more of semantic, rather than orthographic, processing (as Azuma & Van Orden, 1997, have assumed), ambiguous words should have become more difficult to process in the pseudohomophone condition for Pexman and Lupker's participants. Thus, ambiguity effects should have decreased rather than increased in that condition. The fact that the effect increased when pseudohomophones were used, therefore, challenges Azuma and Van Orden's analysis of the effect of using pseudohomophones in lexical decision tasks. Rather, as suggested by Pexman and colleagues (e.g., Pexman & Lupker, 1999; Pexman, Lupker, & Reggin, 2002; Pexman et al., 2001), what pseudohomophones appear to do is to drive participants to engage in more extensive orthographic processing.

As a result, we are reluctant to accept Azuma and Van Orden's (1997) conclusions and would suggest the following alternative explanation (couched within the framework of the feedback account) for at least part of Azuma and Van Orden's results. Following Azuma and Van Orden's suggestion, if one assumes that highly related meanings share many semantic features, fewer semantic features would be activated by words with highly related meanings than by words with less related meanings. Thus, the amount of semantic activation would not differ much for unambiguous words and for ambiguous words with highly related meanings, making it somewhat difficult to observe an ambiguity effect in the highly related condition. In the low-related condition, however, the NOM would more directly reflect the amount of semantic activation and, hence, the strength of semantic feedback. As a result, an ambiguity effect would be more easily detectable in this condition.

The feedback account can, therefore, explain the simple main effects of ambiguity in the low- and high-related-meaning conditions. Where the account runs into difficulty is in explaining the simple main effects of relatedness of meanings. That is, if the amount of semantic activation is greater for words with low-related meanings than for words with highly related meanings, the feedback account predicts that orthographic processing would be faster for words with low-related meanings because of the stronger semantic feedback. As noted, Azuma and Van Orden (1997) reported no relatedness effect in the many-meaning condition and an advantage for high-related-meaning words in the few-meaning condition.

As previously noted, however, Klein and Murphy (2001) suggested that both homonymous meanings (i.e., unrelated meanings) and polysemous senses (i.e., related senses) are represented in a similar manner. If so, in contrast to what has been suggested by Azuma and Van Orden (1997), it is unlikely that the semantic activation is modulated by relatedness among meanings. In fact, in our recent attempt to replicate Azuma and Van Orden's results using their stimuli as well as a new set of items (Hino, Lupker, & Pexman, 2001), we repeatedly failed to observe their relatedness effect in lexical decision tasks. Thus, at present, it is not entirely clear whether relatedness among meanings really affects lexical decision performance. Undoubtedly, more research will be needed to resolve this issue.

Can the Feedback Account Be Mapped to Localist Frameworks?

Hino and Lupker's (1996) feedback account was based on a PDP framework. However, as previously noted, the idea of semantic feedback comes from Balota et al.'s (1991) explanation of semantic effects, an explanation that was couched in terms of McClelland and Rumelhart's (1981) interactive-activation model, which is a localist framework. Thus, a reasonable question is whether the present results could also be explained in terms of semantic-feedback activation within a localist framework. In general, the answer appears to be yes. For example, the synonymy disadvantage in our lexical decision and naming tasks would be easily explained in terms of the interactive-activation model if semantic feedback to the lexical units were assumed. That is, when an activated semantic unit feeds activation back to all the lexical units for synonyms, lexical selection would be slowed because of lateral inhibition at the lexical level.

Where this type of account would run into problems, however, is that as noted previously, it fails to account for Hino and Lupker's (1996) ambiguity effects in lexical decision, naming, and go/no-go naming tasks. In addition, if one assumes localist representations, it would appear to be difficult to account for Pexman, Lupker, and Hino's (in press) NOF effects in lexical decision and naming tasks. That is, assuming that each word meaning is represented as a single semantic unit, there would be no obvious reason to predict that the amount of semantic activation would be modulated by the number of semantic features a word has. Thus, when the feedback account is mapped onto the localist frameworks, its explanatory power appears to be more limited. It remains to be seen whether the localist frameworks could ultimately overcome these problems.

Other Related Issues

Following Joordens and Besner's (1994) suggestion, we have argued that the time to settle on a semantic code should be slower for ambiguous words in all the PDP models because ambiguous words would possess one-to-many mappings from orthography to semantics. Consistent with this suggestion, we observed an ambiguity disadvantage in our semantic categorization tasks (in Experiments 2 and 4). In addition, a similar ambiguity disadvantage was also reported in studies examining fixation times for ambiguous and unambiguous words in a neutral sentential context (e.g., Duffy et al., 1988; Rayner & Duffy, 1986) and in studies using association-judgment tasks (Gottlob et al., 1999; Piercey & Joordens, 2000), although, as previously noted, it is somewhat unclear whether the results of association-judgment tasks reflect the nature of the meaning-determination process.

Recently, however, Forster (1999) reported null ambiguity effects in semantic categorization tasks. Forster's (1999) task was somewhat different from ours because whereas our participants were asked to decide whether a given word is a name of a living object, in Forster's (1999) task, participants were asked to decide whether a given word is a name of an animal. That is, Forster (1999) used a somewhat narrower semantic category in his categorization task. To examine whether the range of semantic categories used in a semantic categorization task affects the sizes of ambiguity effects, we (Hino et al., 2001) recently examined two semantic categorization tasks using the same ambiguous and unambiguous words. Whereas participants were asked to decide whether a given word is a name of a living object in one task, participants in the other task were asked to decide whether a given word is a name of a vegetable. Consistent with our results, an ambiguity disadvantage was observed when a broader semantic category (living objects) was used. In contrast, consistent with Forster's (1999) results, no ambiguity effect was observed when a narrower semantic category (vegetables) was used.

If this empirical conclusion is accurate it raises the question of why the nature of the semantic task matters. Hino et al. (2001) used the same ambiguous and unambiguous words, and if the speed of semantic coding is modulated by the nature of the orthographic-to-semantic mappings, an ambiguity effect should emerge regardless of the type of semantic category used in these tasks. Obviously, a complete explanation of these results will need to take into account the processing differences between deciding that something is a member of a well-specified category versus deciding that something is a member of a more general category.

An additional issue that should be noted here is that, as Joordens and Besner (1994) pointed out, in PDP simulations with ambiguous words (see also Borowsky & Masson, 1996; Kawamoto et al., 1994), semantic units mostly settled into blend states in which both meanings were partially activated. It seems unlikely that this could be an accurate description of what the participants in our experiments were doing. That is, to respond accurately in any semantic categorization task, it would appear to be necessary to unambiguously determine at least one meaning of an ambiguous word, rather than simultaneously activating partial semantic information for all the meanings. Thus, at a general level, it is unclear whether any of these PDP simulations could be correctly reflecting how semantic information is activated. Clearly, more research is needed to resolve how readers activate meanings on the basis of orthographic inputs.

Thus far, we have argued that semantic feedback affects orthographic and phonological processing in an automatic fashion. What should also be noted is that there are studies suggesting that semantic feedback can be blocked in some situations (e.g., M. C. Smith & Besner, 2001; Stolz & Neely, 1995). For example, using a priming task, M. C. Smith and Besner (2001) asked their participants to either make lexical decisions or do a letter search for a target, depending on the color of the target (i.e., a blue target for a lexical decision response and a red target for a letter-search response). In their experiment, a standard semantic priming effect was observed when a lexical decision was required for targets, but no semantic priming effect was observed when letter search was required for targets, although letter search responses were faster for word targets than for nonword targets. On the basis of these results, M. C. Smith and Besner argued that semantic-feedback activation is blocked when letter search is required for target stimuli. If M. C. Smith and Besner's results were really due to a feedback activation block, any feedback account would need to incorporate an additional mechanism that would be responsible for enabling or disabling feedback activation. On the other hand, the nature of the processing differences between making a lexical decision and doing a letter search could provide at least a partial explanation of M. C. Smith and Besner's results. Clearly, further research is necessary to resolve these issues.

Interactions Between Orthography, Phonology, and Semantics

Our results suggest that when responding is based on orthographic processing (as in lexical decision tasks), task performance is affected by the nature of feedback relationships from semantics to orthography. Similarly, when responding is based on phonological processing (as in naming tasks), task performance is affected by the nature of feedback relationships from semantics to phonology. In particular, a processing advantage was observed for ambiguous words, words that involve many-to-one mappings from semantics to orthography and phonology. In contrast, when words possess one-to-many feedback mappings from semantics to orthography and phonology (i.e., words with synonyms), a processing disadvantage was observed.

These results changed dramatically when the task required participants to complete meaning determination, and the responding was based on the outcome of semantic processing (i.e., in semantic categorization tasks). In those situations, task performance appeared to be affected by the nature of feedforward relationships from orthography to semantics. That is, the one-to-many feedforward mappings for ambiguous words produced a processing disadvantage (but see Forster, 1999, and Hino et al., 2001). In contrast, the number of synonyms factor produced no observable effect.

As previously noted, these arguments can be easily extended to explain the interactions of orthography and phonology. That is, when responding is based mainly on orthographic processing, as in lexical decision tasks, a homophone disadvantage has been observed (e.g., Davelaar et al., 1978; Pexman & Lupker, 1999; Pexman et al., 2001; Rubenstein et al., 1971). As suggested by Pexman and colleagues (Pexman & Lupker, 1999; Pexman et al., 2001), because task performance is affected by the nature of feedback relationships from phonology to orthography, the homophone disadvantage is due to the one-to-many feedback relationships for homophones. On the other hand, when a task requires participants to complete phonological coding and responding is based on the results of phonological processing, as in naming tasks, there is no homophone effect (Pexman, Lupker, & Reggin, 2002). Instead, a homograph disadvantage is observed (i.e., words with multiple pronunciations, like *wind*, are harder to name than matched controls; e.g., Gottlob et al., 1999; Kawamoto & Zemblidge, 1992; Pexman, Lupker, & Reggin, 2002; Seidenberg et al., 1984). Because naming task performance is sensitive to the nature of feedforward relationships from orthography to phonology, the homograph disadvantage is, presumably, due to the one-to-many feedforward relationships for homographs.

As such, the interaction between orthography and phonology and the interaction between orthography and semantics appear to have similar characteristics. In addition, these characteristics appear to be captured by PDP frameworks if the nature of feedforward and feedback relationships are fully considered. Although there are a number of criticisms of this type of framework (e.g., Balota & Spieler, 1998; Besner, 1999; Besner, Twilley, McCann, & Seergobin, 1990; Fera & Besner, 1992; Spieler & Balota, 1997), at least at present PDP frameworks appear to provide a very useful basis for understanding the process of reading words.

References

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 802–814.

Azuma, T. (1996). Familiarity and relatedness ratings for 110 homographs. *Behavior Research Methods, Instruments, & Computers*, *28*, 109–124.

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, NJ: Erlbaum.

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., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, *24*, 89–106.

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: Erlbaum.

Balota, D. A., & Spieler, D. H. (1998). The utility of item-level analyses in model evaluation: A reply to Seidenberg and Plaut (1998). *Psychological Science*, *9*, 238–240.

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. (1999). Basic processes in reading: Multiple routines in localist and connectionist models. In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 413–458). Cambridge, MA: MIT Press.

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: The psychology of reading* (pp. 201–219). Hillsdale, NJ: Erlbaum.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review*, *97*, 432–446.

Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental 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 University 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.

Cohen, J. (1976). Random means random. *Journal of Verbal Learning and Verbal Behavior*, *15*, 261–262.

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.

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J. T. (1978). Phonological recoding and lexical access. *Memory & Cognition*, *8*, 195–209.

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

Edwards, J. D., & Pexman, P. M. (2001). What can homophone effects tell us about the nature of orthographic representation in visual word recognition? In J. D. Moore & K. Stenning (Eds.), *Proceedings of the twenty-third annual conference of the Cognitive Science Society* (pp. 261–266). Mahwah, NJ: Erlbaum.

Fera, P., & Besner, D. (1992). The process of lexical decision: More words about a parallel distributed processing model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 749–764.

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, November). *Beyond lexical decision: Lexical access in categorization tasks.* Paper presented at the 40th Annual Meeting of the Psychonomic Society, Los Angeles.

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

Fushimi, T., Ijuin, M., Patterson, K., & Tatsumi, I. F. (1999). Consistency, frequency, and lexicality effects in naming Japanese Kanji. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 382–407.

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.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics, 45,* 189–195.

Hino, Y., & Lupker, S. J. (1996). The 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., & Pexman, P. M. (2001, November). *Effects of polysemy and relatedness of meanings in lexical decision and semantic categorization tasks.* Paper presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL.

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.

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

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

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

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

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

Keppel, G. (1976). Words as random variables. *Journal of Verbal Learning and Verbal Behavior*, *15*, 263–265.

Kindaichi, K., Kindaichi, H., Kenbou, H., Shibata, T., & Yamada, T. (1974). *Shin-meikai kokugo jiten: Dai 2 ban* [New clearly resolved Japanese dictionary: The second edition]. Tokyo: Sanseido.

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

Lesch, M. F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 285–294.

Lichacz, F. M., Herdman, C. M., LeFevre, J., & Baird, B. (1999). Polysemy effects in word naming. *Canadian Journal of Experimental Psychology*, *53*, 189–193.

Lukatela, G., Frost, S. J., & Turvey, M. T. (1998). Phonological priming by masked nonword primes in the lexical decision task. *Journal of Memory and Language*, *39*, 666–683.

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: I. An account of basic findings. *Psychological Review*, *88*, 375–407.

McRae, K., & Cree, G. S. (in press). Factors underlying category specific semantic deficits. In E. M. E. Forde & G. W. Humphreys (Eds.), *Category-specificity in brain and mind.* East Sussex, UK: Psychology Press.

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. (1963). *Gendai-zasshi 90 shu no yougo youji: Dai 2 bunsatsu, Kanji-hyou* [Vocabulary and Chinese characters in 90 magazines of today: Vol. 2. Chinese characters frequency table]. Tokyo: Shuei Shuppan.

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

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

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

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. (2001). Perception is a two-way junction: Feedback semantics in word recognition. *Psychonomic Bulletin & Review*, *8*, 545–551.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language*, *39*, 151–174.

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language*, *27*, 59–70.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 539–557.

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. (in press). The impact of feedback semantics in visual word recognition: Number of features effects in lexical decision and naming tasks. *Psychonomic Bulletin & Review*.

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 139–156.

Pexman, P. M., Lupker, S. J., & Reggin, L. D. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 572–584.

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

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

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

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

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*, 416–426.

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.

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.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 876–900.

Seidenberg, M. S. (1992). Beyond orthographic depth in reading: Equitable division of labor. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 85–118). Amsterdam: Elsevier.

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

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

Smith, J. E. K. (1976). The assuming-will-make-it-so fallacy. *Journal of Verbal Learning and Verbal Behavior*, *15*, 262–263.

Smith, M. C., & Besner, D. (2001). Modulating semantic feedback in visual word recognition. *Psychonomic Bulletin & Review*, *8*, 111–117.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science*, *8*, 411–416.

Stolz, J. A., & Neely, J. (1995). When target degradation does and does not enhance semantic context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 596–611.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language*, *36*, 337–359.

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., & Van Graan, F. (1998). Lack of phonological mediation in a semantic categorization task. *Journal of Memory and Language*, *38*, 203–224.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*, 181–198.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 371–386.

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.

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *47*(A), 631–650.

Wike, E. L., & Church, J. D. (1976). Comments on Clark's "The language-as-fixed-effect fallacy." *Journal of Verbal Learning and Verbal Behavior*, *15*, 249–255.

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language*, *37*, 533–554.

Appendix A

Experimental Items and Their English Translations Used in Experiments 1 and 2

Ambiguous word		Unambiguous word	
アウト	out, in baseball	アイデア	idea
	outside		
オーバー	overcoat	アルミ	aluminum
	exaggeration		
クラウン	crown	イコール	equal
	name of a car		
ケース	case, for holding something	カード	card
	case, what actually exists or happens		
コンタクト	contact, touch	ガソリン	gas
	contact lens		
サークル	circle, shape	グレー	gray
	circle, as an organization		
サイクル	cycle	サンダル	sandals
	bicycle		
ショート	short	ショック	shock
	shortstop, in baseball		
スパイク	spike, in volleyball	スケート	skate
	spiked shoe		
セーブ	saving	ゼロ	zero
	save, in baseball		
ソフト	soft ice cream	ダーツ	darts
	softball		
	software		
ダウン	down	チェーン	chain
	a down jacket		
チェック	check, examine	デニム	denim
	check (clothes)		
チップ	tip, money for a service	テント	tent
	chip, as a small piece		
パス	passing, of ball or baton	ドーム	dome
	pass, for transportation		
バレー	ballet	バイク	bike, motorcycle
	volleyball		
パンク	a flat tire	バッジ	badge
	punk rock		
ピース	peace	ヒント	hint
	a name of cigarettes		
フック	hook, clasp	ブック	book
	hook, in boxing		
フロント	front desk	ボイコット	boycott
	front side		
ベース	base, used in baseball	ボート	boat
	base, as basis		
	bass		
マーク	mark, a sign, token	ポット	pot
	mark, for assessment or rating		
マガジン	magazine, periodical	ポップス	pops
	magazine rack		
マジック	magic trick	ラグビー	rugby
	permanent marker		
ライト	light	ランク	rank
	right		
ライン	line	ランプ	lamp
	the Rhine (river)		
リング	ring, a circular band	ロープ	rope
	ring, to sound		
ロック	rock	ワックス	wax
	lock		

(*Appendixes continue*)

Appendix B

Experimental Items and Their English Translations Used in Experiments 3–6

Ambiguous word

Medium NOS		Large NOS	
サイクル	cycle bicycle	セーブ	saving save in baseball
パンク	flat tire punk rock	マジック	magic trick permanent marker
スパイク	spike in volleyball spiked shoe	フロント	front desk front side
マーク	mark, a sign, token mark for assessment or rating	ピース	peace a name of cigarettes
バレー	ballet volleyball	ショート	short shortstop in baseball
マガジン	magazine, periodical magazine rack	オーバー	overcoat exaggeration
デート	date, time date, meeting someone	アウト	out in baseball outside
ポスト	post, pillar-box position	チップ	tip, money for a service chip as a small piece
マッチ	burning match matching up	バット	bat, a stick used in hitting a ball bat, a flying mammal
ベース	base used in baseball base as basis bass	ライン	line the Rhine (river)
パンチ	punch to strike punch, a kind of drink a specific form of a permanent wave	サークル	circle shape circle as an organization
クラウン	crown a name of a car	ライト	light right
コート	wearing coat court for sports	ケース	case for holding something case, what actually exists or happens
スタンド	table lamp gas station baseball stadium	ホーム	home home plate in baseball
トラック	track on field cargo truck	スーパー	super grocery store

Unambiguous word

Small NOS		Medium NOS	
バッジ	badge	ドーム	dome
ポット	pot	サンダル	sandals
ワックス	wax	フロア	floor
ラッパ	trumpet	ランク	rank
カクテル	cocktail	ワイヤー	wire
テント	tent	テニス	tennis
ヒント	hint	ピンク	pink
カード	card	アルミ	aluminum
ガソリン	gasoline	ボート	boat
ミシン	sewing machine	スープ	soup
スケート	skating	アイデア	idea
セーター	sweater	ウール	wool
ギター	guitar	チェーン	chain
ビール	beer	ボーナス	bonus
スタジオ	studio	シーズン	season

Note. NOS = number of synonyms.

Appendix C

Experimental Items and Their English Translations Used in Experiments 7–10

Small NOS		Large NOS	
煙突	a chimney	外車	an imported car
屋上	a rooftop deck	街灯	a street lamp
温室	a greenhouse	歌曲	a song
回路	a circuit	玩具	a toy
戸籍	a census register	教材	an educational material
新車	a new car	下宿	a lodging house
新書	a new book	原爆	an atomic bomb
水爆	a hydrogen bomb	原油	crude oil
聖書	Bible	詩集	a collection of poems
戦車	a tank	品物	goods
注射	an injection	車輪	a wheel
通帳	a passbook	賞金	a prize
荷物	a baggage	弾丸	a bullet
版画	a block print	電鉄	an electric railroad
表紙	a front cover	塗料	paint
封筒	an envelope	文具	stationery
踏切	a railway crossing	水着	a swimsuit
名刺	a business card	薬剤	a medicine
硫酸	a sulfuric acid	容器	a container
和室	a Japanese style room	和服	a Japanese dress

Note. NOS = number of synonyms.

Received February 27, 2001
Revision received January 26, 2002
Accepted January 26, 2002 ■