Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 1, 121–128

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.1.121

# Enemies and Friends in the Neighborhood: Orthographic Similarity Effects in Semantic Categorization

Diane Pecher
Erasmus University Rotterdam

René Zeelenberg
Indiana University and Erasmus University Rotterdam

Eric-Jan Wagenmakers
University of Amsterdam

Studies investigating orthographic similarity effects in semantic tasks have produced inconsistent results. The authors investigated orthographic similarity effects in animacy decision and in contrast with previous studies, they took semantic congruency into account. In Experiments 1 and 2, performance to a target (*cat*) was better if a previously studied neighbor (*rat*) was congruent (i.e., belonged to the same animate–inanimate category) than it was if it was incongruent (e.g., *mat*). In Experiments 3 and 4, performance was better for targets with more preexisting congruent neighbors than for targets with more preexisting incongruent neighbors. These results demonstrate that orthographic similarity effects in semantic categorization are conditional on semantic congruency. This strongly suggests that semantic information becomes available before orthographic processing has been completed.

The question of whether and how orthographic similarity affects the word-recognition process has been the focus of much research over the last 10 years. Two paradigms that have been used to investigate this issue are of interest to us in this study. The first paradigm is the long-term priming paradigm (e.g., Masson, 2002; Ratcliff & McKoon, 1995, 1996, 1997; Wagenmakers, Zeelenberg, & Raaijmakers, 2000; Zeelenberg, Wagenmakers, & Raaijmakers, 2002), which has been used to assess to what extent performance to a target stimulus (e.g., *lied*) is affected by prior study of an orthographically similar stimulus (e.g., *died*). The second paradigm that has been used to investigate the role of orthographic similarity in word identification examines the effect of the preexisting orthographic similarity structure in the lexicon (see Andrews, 1997, for a review; and Grainger & Jacobs, 1996, for a model). For example, performance for words (e.g., *bean*) that have many orthographic neighbors (i.e., words such as *lean* or *beat*, which differ from *bean* by one letter; Coltheart, Davelaar, Jonasson, & Besner, 1977; Landauer & Streeter, 1973) can be compared with performance for words (e.g., *clean*) with only a few orthographic neighbors.

In the long-term priming paradigm, several recent studies have provided evidence that orthographic similarity plays an important role in visual word recognition. One task in which orthographic similarity effects have been obtained is the forced-choice visual word identification task. In this task, a word (e.g., *lied*) is briefly flashed on a screen and subsequently masked. The mask is followed by the presentation of two alternatives (e.g., *lied* and *died*), and the participant's task is to choose which one of the two alternatives was flashed. In several experiments, Ratcliff and McKoon (1997; see also Wagenmakers et al., 2000) found that prior study of the target word (e.g., *lied*) increased target identification. However, prior study of an orthographically similar foil (e.g., *died*) decreased target identification (i.e., performance was worse when the foil was studied prior to being presented in the identification task compared with when neither alternative was studied). Another important finding is that this pattern of benefits and costs was obtained only when the alternatives were orthographically similar (e.g., *lied* vs. *died*). For orthographically dissimilar alternatives (e.g., *lied* vs. *sofa*), prior study had no effect (Masson, 2002; Ratcliff & McKoon, 1997; but see Bowers, 1999; and for a reply, see McKoon & Ratcliff, 2001).

Decrements in performance that were due to prior study of orthographically similar words have also been obtained in word-fragment completion (Ratcliff & McKoon, 1996; Zeelenberg et al., 2002) and word-stem completion (Ratcliff & McKoon, 1996). However, prior study of words that are orthographically similar to the later presented target word does not always result in a performance decrement. For instance, Bowers, Damian, and Havelka (2002) recently found facilitation in a lexical decision task for target words that were orthographically and phonologically similar to words that had been presented previously in the experiment (but

---

Diane Pecher, Department of Psychology, Erasmus University Rotterdam, Rotterdam, The Netherlands; René Zeelenberg, Department of Psychology, Indiana University and Department of Psychology, Erasmus University Rotterdam; Eric-Jan Wagenmakers, Department of Developmental Psychology, University of Amsterdam, Amsterdam, The Netherlands.

We thank Jordy Zandwijk for his help with Experiment 1 and Adam Sanborn for his help with Experiment 2. We also thank Rich Shiffrin and Jeroen Raaijmakers for useful discussions and Ken Forster, Jeff Bowers, and Jonathan Grainger for helpful comments on a draft of this article.

Correspondence concerning this article should be addressed to Diane Pecher, Erasmus University Rotterdam, Department of Psychology, Postbus 1738, 3000 DR Rotterdam, The Netherlands. E-mail: pecher@fsw.eur.nl

no effect for targets that were orthographically but not phonologically similar to previously presented words). This difference in results between studies might be due to characteristics of the lexical decision task. In particular, unlike the other tasks that have been used to investigate the effect of prior study of similar words, the lexical decision task does not require that the target word be uniquely identified among its orthographic competitors. Responses in the lexical decision task may instead be based on global activation or familiarity (e.g., Balota & Chumbley, 1984; Wagenmakers et al., 2004; Zeelenberg, Wagenmakers, & Shiffrin, 2004).

The task-dependent nature of the orthographic similarity effect is also observed in the paradigm that focuses on the preexperimental existing similarity structure in the lexicon. In studies based on this paradigm, researchers have investigated the effect of both neighborhood density (i.e., the total number of neighbors) and neighbor frequency (i.e., whether the target has a higher frequency neighbor) in a number of different word-recognition tasks. In general, these researchers have found that a large number of neighbors may facilitate performance in tasks such as lexical decision, in which a response can be made on the basis of global familiarity (Andrews, 1997; Grainger & Jacobs, 1996). However, the presence of a high-frequency neighbor often harms performance, particularly in tasks such as perceptual identification that require the unique identification of the presented stimulus (Andrews, 1997).

The researchers mentioned so far have investigated orthographic similarity effects in tasks that do not require access to the meaning of a word. Recently, a few researchers have investigated neighborhood effects in semantic-categorization tasks. The expectation was that neighborhood density and neighbor frequency would harm performance in semantic-categorization tasks because these tasks, like the perceptual-identification task, require the target word to be uniquely identified. An advantage of semantic-categorization tasks over perceptual-identification tasks in the study of orthographic similarity effects is that semantic categorization is relatively unlikely to be affected by guessing processes that may well affect the results obtained in perceptual identification (Andrews, 1997). Unfortunately, studies using semantic-categorization tasks have produced mixed results. Forster and Shen (1996; Experiment 4) manipulated four levels of neighborhood density and found no overall significant effect of neighborhood density. In a subsequent experiment, they found facilitatory effects of neighborhood density. Sears, Lupker, and Hino (1999) also observed facilitatory effects of neighborhood density (but no effect of neighbor frequency). Forster and Hector (2002), however, obtained no effect of neighborhood density for words. Finally, Carreiras, Perea, and Grainger (1997) obtained facilitatory effects of neighborhood density for words without a higher frequency neighbor but inhibitory effects for words with a higher frequency neighbor.

One possible factor that may have contributed to the inconsistent results in semantic-categorization tasks might be that existing studies failed to take semantic congruency into account. We define a semantically congruent orthographic neighbor as a word that is orthographically similar to the target word and belongs to the same semantic category. For instance, when deciding whether *cat* is an animal, *rat* is a semantically congruent orthographic neighbor, and *mat* is a semantically incongruent orthographic neighbor. We hypothesize that in semantic-categorization tasks, neighborhood density or neighbor frequency may not be the only neighborhood-related factor that affects performance. An important additional factor could be the balance between congruent and incongruent neighbors. A common assumption in previous studies was that (at least) two discrete stages of processing underlie performance in semantic-categorization tasks (Sears et al., 1999; see Forster & Hector, 2002, for discussion). That is, the target word first has to be uniquely identified among its orthographic competitors, and only after this identification process has been successfully completed can the required semantic knowledge associated with the target word be accessed. In such models, it is assumed that lexical identification and retrieval of semantic information are serial processes.

It is quite possible, however, that the activation or retrieval of semantic information takes place before the identification stage has been completed (i.e., before the word has been uniquely identified). Such an assumption is made in cascaded models of word recognition (e.g., McClelland, 1979; Plaut, McClelland, Seidenberg, & Patterson, 1996). If, indeed, semantic information becomes available before orthographic processing of the target word has been completed, the semantic status of words similar to the target may affect performance. To be specific, assume that quickly after presentation of a target word, the available perceptual information activates the orthographic representation of the target word and, albeit to a lesser extent, the orthographic representations of words that are orthographically similar to the target word. As processing progresses, the activated orthographic representations will retrieve or activate the semantic information with which they are associated. Thus, if many neighbors are semantically incongruent with the target word, the semantic information that is activated by the target word will be in conflict with the semantic information that is activated by its neighbors. Assuming that the semantic conflict will take time to resolve, the activation of semantically incongruent orthographic neighbors will result in relatively slow and inaccurate responses. If many neighbors are semantically congruent with the target word, the retrieved semantic information will be consistent, and this will lead to relatively fast and accurate responding. In sum, cascaded models predict that performance in a semantic-categorization task will be better for words with mostly semantically congruent neighbors than for words with mostly semantically incongruent neighbors. A serial model, in contrast, predicts that semantic congruency of orthographic neighbors will not affect semantic categorization, because by the time semantics are retrieved from memory, orthographic neighbors are no longer relevant.

In the present study, we investigated orthographic similarity effects in animacy decision, a semantic-categorization task. In contrast with previous studies, which have focused on neighborhood density or neighbor frequency, in our study, we manipulated semantic congruency. In Experiments 1 and 2, we used a long-term priming paradigm to study orthographic similarity effects and in Experiments 3 and 4, we studied the effect of the preexperimental existing neighborhood structure.

## Experiment 1

In Experiment 1, we investigated whether prior study of orthographic neighbors affects performance in animacy decision. A number of studies (e.g., Vriezen, Moscovitch, & Bellos, 1995;

Zeelenberg & Pecher, 2003) have shown long-term repetition priming effects in semantic-categorization tasks such as animacy decision. In these studies, prior study of the target itself resulted in faster and more accurate responses. To the best of our knowledge, however, no study has investigated the effect of prior study of a word that is orthographically similar to the target word in a semantic-categorization task. If prior study of a word increases the strength with which associated semantic information is activated, the cascaded model outlined above predicts that prior study of a semantically congruent neighbor will lead to better performance than will prior study of an incongruent neighbor. In the study phase of Experiment 1, participants responded to orthographic neighbors (henceforward referred to as primes) of targets that were later presented in the test phase. For targets in the congruent condition, the studied prime belonged to the same animate or inanimate category (e.g., target: *cat*; prime: *rat*). For targets in the incongruent condition, the studied prime belonged to the other category (e.g., target: *cat*; prime: *mat*). In Experiment 1, the complete list of prime words was presented first and was followed by the list of target words. In Experiment 1, we investigated whether performance to the target stimulus would be better if a congruent prime had been studied earlier in the experiment than if an incongruent prime had been studied.

## Method

*Participants.* Fifty-two University of Amsterdam students participated for course credit. All participants were native speakers of Dutch.

*Stimuli.* The experimental stimuli consisted of 56 Dutch word triplets. Each triplet consisted of a target word (e.g., *muur* [wall]), a congruent prime word from the same semantic category (e.g., *vuur* [fire]), and an incongruent prime word from a different category (e.g., *buur* [neighbor]). Each prime differed by only one letter from its target and by at least two letters from every other target on the list. There were 27 animate targets, for which the congruent prime was animate and the incongruent prime was inanimate, and 29 inanimate targets, for which the congruent prime was inanimate and the incongruent prime was animate. We obtained frequency counts from the CELEX norms (Baayen, Piepenbrock, & van Rijn, 1993). The mean frequency per million was 28.7 (range 1.1–281.5) for the target words, 81.7 (range 0.0–1,376.3) for the congruent prime words, and 80.1 (range 1.1–1,195.2) for the incongruent prime words.

We created two lists for counterbalancing purposes. On each list, 28 targets were preceded by a congruent prime, and the other 28 targets were preceded by an incongruent prime. Across the two lists, each target was presented once in the congruent condition and once in the incongruent condition. On one list, 13 animate targets and 15 inanimate targets were presented in the congruent condition, and 14 animate and 14 inanimate targets were presented in the incongruent condition. On the other list, 14 animate and 14 inanimate targets were presented in the congruent condition, and 13 animate and 15 inanimate targets were presented in the incongruent condition.

We selected an additional set of 14 animate and 14 inanimate words to serve as practice stimuli. None of the practice stimuli were neighbors of any of the experimental stimuli. No word appeared more than once during the entire experiment.

*Procedure.* Participants received spoken and written instructions to make an "animate" decision if the word represented something living (i.e., a human being, animal, or plant), or part of a living thing and to make an "inanimate" decision if the word represented something not living. Examples of animate and inanimate things were provided. Participants were instructed to respond as quickly as possible without making errors. The experiment consisted of a study block and a test block. In both blocks, words were presented in an animacy decision task. Participants were not informed about the relation between the study and test block. The experiment started with the 28 practice trials followed by a short break. In the study block, 56 primes were presented followed by another short break. Finally, in the test block, 56 targets were presented. Each set of stimuli was presented in a different random order for each participant.

Words were presented one at a time on a computer screen. Each trial started with the presentation of a fixation mark (* * * * *) for 500 ms. The fixation mark was followed immediately by the target stimulus, which remained on the screen until the participant had made a response by pressing the *?/* key for an "animate" response or the *Z* key for an "inanimate" response. If the participant made an error, the word *FOUT* [error] was presented for 2,000 ms. If the response was correct but slower than 1,800 ms, the words *TE LANGZAAM* [too slow] were presented for 2,000 ms. The next trial started 500 ms after the response or, in case of an erroneous or slow response, 500 ms after the presentation of feedback.

## Results and Discussion

We included only reaction times to correct responses that fell within three standard deviations of the participant's mean in the 2 (animacy status) × 2 (congruency) analysis of variance (ANOVA). This trimming procedure resulted in the removal of 2.4% of the reaction times to correct responses. Table 1 shows the mean reaction times and error percentages. Reaction times were faster for targets for which a congruent prime had been studied previously in the experiment than for targets for which an incongruent prime had been studied, $F(1, 51) = 8.56, p < .01$. Although the congruency effect was numerically larger for inanimate (24 ms) than for animate targets (8 ms), the Congruency × Animacy interaction was not significant, $F(1, 51) = 2.91$, $p = .094$. The error data showed no effect of congruency, $F(1, 51) < 1$, or interaction, $F(1, 51) < 1$.

Thus, in Experiment 1, we obtained a reliable congruency effect. In Experiment 2, we aimed at replicating and extending the results of Experiment 1. To this end, Experiment 2 differed from Experiment 1 in two aspects. First, we used a new set of materials and a different population of participants. Second, we included a neutral condition in Experiment 2. This condition was included to address the question of whether the difference between the congruent and incongruent conditions was due to facilitation, inhibition, or both. By comparing the incongruent and congruent conditions with a neutral condition, one can obtain measures of inhibition and facilitation. For targets in the neutral condition, no orthographic neighbor was studied. If the semantic-congruency effect is mainly due to facilitation, we would expect no difference between the neutral and the incongruent conditions. If the

Table 1

*Mean Reaction Times (in ms) and Error Rates in the Animacy Decision Task of Experiment 1*

| | RT | | ER | |
|---|---|---|---|---|
| Study condition | *M* | *SE* | *M* | *SE* |
| Congruent | 685 | 11.3 | 4.8 | 0.68 |
| Incongruent | 700 | 12.4 | 4.9 | 0.64 |
| Congruency effect | 15 | | 0.1 | |

*Note.* RT = reaction time; ER = error rate.

semantic-congruency effect is mainly due to inhibition, we would expect no difference between the neutral and the congruent conditions. If the effect is due to a combination of facilitation and inhibition, we would expect the neutral condition to be somewhere between the congruent and the incongruent conditions. In Experiment 2, the primes were presented twice instead of once in an attempt to increase the size of the congruency effect, which was small (but reliable) in Experiment 1.

## Experiment 2

### *Method*

*Participants.* Sixty-six Indiana University, Bloomington students participated for course credit. All participants were native speakers of English.

*Stimuli.* The experimental stimuli consisted of 60 English word triplets. The triplets were constructed in the same fashion as in Experiment 1. Each triplet consisted of a target, a congruent prime, and an incongruent prime. Each prime differed by only one letter from its target and by at least two letters from every other target on the list. There were 21 animate targets and 39 inanimate targets. We obtained frequency counts from the CELEX norms (Baayen et al., 1993). The mean frequency per million was 50.2 (range 1–1,306) for the target words, 51.2 (range 1–423) for the congruent prime words, and 81.3 (range 1–1,727) for the incongruent prime words.

We created three lists for counterbalancing purposes. On each list, the targets were equally divided among the congruent, incongruent, and neutral (i.e., nonprimed) conditions. Each condition contained 7 animate and 13 inanimate targets. Across the three lists, each target word was presented once in each condition.

We created two additional sets of stimuli: a set of practice stimuli consisting of 14 animate and 14 inanimate words and a set of filler stimuli consisting of 18 animate words. We added the filler stimuli to the test list in order to equate the number of animate and inanimate items.

*Procedure.* The procedure was the same as that used in Experiment 1, except that two study blocks were presented instead of one. In each block, all primes were presented once in random order. Different random orders were used in the two study blocks.

### *Results and Discussion*

We included only reaction times to correct responses that fell within three standard deviations of the participant's mean in a 2 (animacy status) × 3 (congruency) ANOVA. This trimming procedure resulted in the removal of 1.19% of the reaction times to correct responses. Table 2 shows the mean reaction times and error percentages. Reaction times and error percentages were lower for targets in the congruent condition than for targets in the incongruent condition. Responses to targets in the neutral condition fell somewhere between the other two conditions. The effect of prime condition was significant both for the reaction times, $F(2, 130) = 11.59$, $MSE = 4{,}260.56$, $p < .001$; and for the error percentages, $F(2, 130) = 22.36$, $MSE = 130.93$, $p < .001$. The Congruency × Animacy status interaction was not significant for reaction times, $F(2, 130) < 1$, or for the errors, $F(2, 130) = 1.11$. Planned comparisons showed that responses in the incongruent condition were slower than responses in the neutral condition, $t(65) = 2.42$, $p < .05$, and that responses in the neutral condition were slower than responses in the congruent condition, $t(65) = 2.61$, $p < .05$. More errors were made in the incongruent condition than in the neutral condition, $t(65) = 5.40$, $p < .001$, but the difference between the neutral condition and the congruent condition was not significant, $t(65) = 0.91$.

Table 2
*Mean Reaction Times (in ms) and Error Rates in the Animacy Decision Task of Experiment 2*

| | RT | | ER | |
|---|---|---|---|---|
| Study condition | *M* | *SE* | *M* | *SE* |
| Congruent | 782 | 11.0 | 10.3 | 0.88 |
| Neutral | 801 | 12.6 | 11.1 | 0.81 |
| Incongruent | 820 | 10.7 | 18.6 | 1.32 |
| Congruency effect | 38 | | 8.3 | |

*Note.* RT = reaction time; ER = error rate.

In sum, Experiment 2 replicated the congruency effect obtained in Experiment 1 with a new participant population and new set of materials in a different language.[1] In addition, Experiment 2 showed that compared with a neutral condition, responses are facilitated by studying a congruent prime and are inhibited by studying an incongruent prime. Thus, the congruency effect is the result of both facilitation and inhibition.

In Experiments 1 and 2, we investigated the effect of studying an orthographic neighbor of the target word. In Experiment 3, we investigated the role of the preexperimental neighborhood structure in animacy decision. If semantic congruency mediates the effect of prior study of an orthographic neighbor, it may also mediate the effect of the preexisting neighborhood structure. In past studies that have looked at neighborhood effects in semantic categorization, researchers have assumed that performance should be negatively correlated with neighborhood density and neighbor frequency because in semantic categorization, the target word has to be uniquely identified before a response can be given (Carreiras et al., 1997; Forster & Hector, 2002; Forster & Shen, 1996; Sears et al., 1999).

In identification tasks such as masked perceptual identification, performance relies on the unique identification of the target word. Therefore, activation of orthographic neighbors is expected to interfere with identification of the target and thereby decrease performance. In animacy decision, however, activation of a neighbor might enhance performance if the neighbor and the target are in the same semantic category and thus require the same response. Previous studies of neighborhood effects in semantic-categorization tasks did not take semantic congruency into account (with the exception of Forster & Hector, 2002, who compared

---

[1] As an aside, note that the effects of congruency were somewhat larger in Experiment 2 than they were in Experiment 1. Several factors may have contributed to this difference. One such factor was that in Experiment 1, the primes were presented only once, whereas in Experiment 2, they were presented twice. Two other plausible factors are the differences in materials and the differences in participants between the two experiments. Note that the overall reaction times in Experiment 2 were higher than those in Experiment 1, indicating that materials in Experiment 2 may have been more difficult or that the two participant groups may have been different in motivation or ability. Whatever the reasons for the larger effects in Experiment 2, the important point is that we again obtained a reliable semantic-congruency effect.

Table 3
*Mean Reaction Times (in ms) and Error Rates in the Animacy Decision Task of Experiment 3*

| | Animate targets | | | | Inanimate targets | | | |
|---|---|---|---|---|---|---|---|---|
| | RT | | ER | | RT | | ER | |
| Neighborhood condition | *M* | *SE* | *M* | *SE* | *M* | *SE* | *M* | *SE* |
| Mostly congruent | 732 | 9.2 | 9.6 | 0.55 | 803 | 10.9 | 7.6 | 0.47 |
| Neutral | 753 | 10.0 | 20.3 | 0.78 | 827 | 11.9 | 11.4 | 0.64 |
| Mostly incongruent | 769 | 10.3 | 21.9 | 0.89 | 827 | 11.9 | 18.9 | 0.69 |
| Congruency effect | 37 | | 12.3 | | 24 | | 11.3 | |

*Note.* RT = reaction time; ER = error rate.

performance in animal decision[2] to nonwords that were derived from animal names with that for nonwords derived from nonanimal names). Another complication is that in these studies, some of the presented words (targets and fillers) may have been neighbors of other target words. As we have seen in the present study, prior presentation of a neighbor can have a significant effect on target processing. Thus, the mixed results in the literature may be (partially) due to a mixture of neighborhood density, neighbor frequency, neighbor congruency, and neighbor priming effects. In Experiments 1 and 2, we showed that neighbor priming affects performance in animacy decision. To investigate the role of neighbor congruency on animacy decision performance, in Experiment 3, we compared performance in animacy decision for words that have mostly semantically congruent orthographic neighbors (e.g., *fridge*, with neighbors *bridge* and *fringe*) with performance for words that have mostly incongruent neighbors (e.g., *cheek*, with neighbors *check* and *creek*). In addition, Experiment 3 included a neutral or baseline condition in which words had approximately equal numbers of congruent and incongruent neighbors (e.g., *noodle*, with neighbors *poodle* and *doodle*). The neighborhood density was kept constant across conditions.

## Experiment 3

### Method

*Participants.* Seventy-four undergraduate students at Indiana University participated for course credit. All participants were native speakers of English. None of the participants had participated in Experiments 1 and 2.

*Stimuli.* We selected a set of 327 target words that consisted of three subsets of 109 words each for the congruent, neutral, and incongruent conditions. The target words in the three subsets were matched on word frequency, word length, and total number of neighbors. For each target, we tallied the number of congruent neighbors, incongruent neighbors, and neighbors that could not be categorized (such as verbs, adjectives, and abstract nouns). All target words were concrete nouns that could be classified on animacy. The semantically congruent condition consisted of targets that had more congruent neighbors than incongruent neighbors: on average, each target in the congruent condition had 1.54 congruent neighbors, 0.12 incongruent neighbors, and 0.84 unclassifiable neighbors. The neutral condition consisted of targets that had the same number of congruent and incongruent neighbors: On average, each target in the neutral condition had 0.49 congruent neighbors, 0.49 incongruent neighbors, and 1.51 unclassifiable neighbors. The incongruent condition consisted of targets that had more incongruent than congruent neighbors: On average, each target in the incongruent condition had 0.17 congruent neighbors, 1.56 incongruent neighbors, and 0.76 unclassifiable neighbors.

The mean word frequencies per million according to the CELEX norms (Baayen et al., 1993) were 21.5, 22.0, and 21.7 for the congruent, neutral, and incongruent conditions, respectively. The mean numbers of letters per word were 5.54, 5.51, and 5.54 for the congruent, neutral, and incongruent conditions, respectively. The total numbers of neighbors were 2.50, 2.49, and 2.50 for the congruent, neutral, and incongruent conditions, respectively. Each set consisted of 51 animate and 58 inanimate words. An additional set of 40 practice items (20 animate and 20 inanimate) and a set of 21 filler items (all animate) were selected. We added the filler stimuli to the test list in order to equate the number of animate and inanimate items. No word presented during the entire experiment was an orthographic neighbor of any other word.

*Procedure.* The procedure was similar to the animacy decision task in Experiments 1 and 2. The experiment started with the 40 practice items in random order. Then the 327 target words and the 21 filler words were presented in random order in blocks of 95 trials (the final block was 63 trials). Participants were instructed to decide as quickly as possible, without making errors, whether a stimulus referred to something animate or to something inanimate.

### Results and Discussion

We included only reaction times to correct responses that fell within three standard deviations of the participant's mean in the 2 (animacy status) × 3 (congruency) ANOVA. This trimming procedure resulted in the removal of 1.34% of the reaction times to correct responses. Performance was affected by the number of congruent and incongruent neighbors of the target word. Responses to words that had mostly semantically congruent neighbors were both faster, $F(2, 146) = 58.57$, $MSE = 638.31$, $p < .001$; and more accurate, $F(2, 146) = 217.34$, $MSE = 24.31$, $p < .001$, than responses for words that had mostly semantically incongruent neighbors. There was a significant Congruency × Animacy interaction in the reaction times, $F(2, 146) = 3.95$, $p < .05$; and in the errors, $F(2, 146) = 25.42$, $p < .001$. Because the interaction was significant, Table 3 shows the results for animate and inanimate targets separately. For the animate targets, responses in the congruent condition were both faster, $t(73) = 4.71$; and more accurate, $t(73) = 13.59$, than those in the neutral condition, and those in the neutral condition were faster, $t(73) = 3.63$, than those in the incongruent condition, but there was no significant difference in accuracy, $t(73) = 1.77$, between the latter two

---

[2] Note that the task that Forster and Hector (2002) used was animal decision, in which a "yes" response should be given only to animal names.

conditions. For the inanimate targets, responses in the congruent condition were both faster, $t(73) = 5.98$; and more accurate, $t(73) = 6.47$, than those in the neutral condition, and those in the neutral condition were more accurate, $t(73) = 9.45$, than those in the incongruent condition, but there was no significant difference in reaction times, $t(73) < 1$, between the latter two conditions. Thus, although not all comparisons were significant, responses to words with more congruent neighbors were overall faster and more accurate than responses to words with more incongruent neighbors.

One potential concern might be that we did not match the three sets of target words on every possible variable that might affect performance on semantic-categorization tasks. For example, the three different sets were not matched on neighbor frequency. Because neighbor frequency might affect performance, we calculated the proportion of targets with a higher frequency neighbor for each condition. These proportions were .16, .24, and .17 for the congruent, neutral, and incongruent conditions, respectively. The two conditions that differed most in the reaction times and errors, the congruent and the incongruent conditions, did not differ in the proportion of target words with a higher frequency neighbor. Thus, differences in neighbor frequency cannot account for our data. It is still conceivable, however, that the three sets of targets differed on some other variables and that these differences caused the congruency effect. In Experiment 4, we therefore presented the target words of Experiment 3 in a different semantic-categorization task, size decision (*is it smaller or larger than a television?*). In the size-decision task, the congruency in terms of animacy is irrelevant and hence should not affect performance. If, however, the results of Experiment 3 were not due to semantic congruency but to imperfect matching of stimuli, we should also observe a congruency effect in Experiment 4.

## Experiment 4

### Method

*Participants.* Forty undergraduate students at Indiana University participated for course credit. All participants were native speakers of English. None of the participants had participated in Experiments 1–3.

*Stimuli.* We selected a subset of 249 words from the target set in Experiment 3. From the original set of 327 words, we had to remove 51 because they could not be classified (e.g., things such as *locker, primate,* and *kite* can be both smaller and larger than a television). From the remaining set of 276 words, we removed another 27 words, so that the targets in the three conditions were matched on word frequency, word length, neighborhood size, and neighborhood frequency. We used the 27 removed words as fillers to keep the list as similar as possible to that of Experiment 3. In each condition, 38 words could be classified as smaller, and 45 words as larger, than a television. The mean log word frequencies were 2.07, 2.08, and 2.12 for the congruent, neutral, and incongruent conditions, respectively. The mean numbers of letters per word were 5.54, 5.55, and 5.58 for the congruent, neutral, and incongruent conditions, respectively. The mean total numbers of neighbors were 2.45, 2.45, and 2.39 for the congruent, neutral, and incongruent conditions, respectively. The mean total numbers of higher frequency neighbors were 0.30, 0.35, and 0.31 for the congruent, neutral, and incongruent conditions, respectively. We selected an additional set of 32 additional filler items (all smaller). We added the filler stimuli to the test list in order to equate the number of smaller and larger items. We used an additional set of 40 items (20 larger and 20 smaller) for practice. No word presented during the entire experiment was an orthographic neighbor of any other word.

*Procedure.* The procedure differed from that of Experiment 3 only in the instruction to the participants. Instead of deciding on the animacy of each target, they decided whether the target word referred to something that was smaller or larger than a regular television.

### Results and Discussion

We included only reaction times to correct responses that fell within three standard deviations of the participant's mean in the analysis. This trimming procedure resulted in the removal of 1.18% of the reaction times to correct responses. Table 4 shows the mean reaction times and error percentages. Performance was not affected by the number of congruent and incongruent neighbors of a word, $F(2, 78) = 1.39$, $p > .25$, $MSE = 349.32$ for reaction times; and $F(2, 78) < 1$, $MSE = 12.76$ for error percentages. Thus, performance in the size-decision task was not affected by the congruency in terms of animacy of a target's neighbors. This indicates that it is highly unlikely that the congruency effect obtained in the animacy decision task of Experiment 3 was due to imperfect matching of stimuli in the different conditions.

Because in Experiment 4 we used only a subset of the items that were used in Experiment 3, we reanalyzed the data of Experiment 3. In this reanalysis, we included only those target words that were also used in Experiment 4. For this stimulus set, we still obtained a reliable congruency effect in animacy decision both for reaction times, $F(2, 146) = 4.81$, $p < .01$, $MSE = 471.72$; and for error percentages, $F(2, 146) = 5.07$, $p < .01$, $MSE = 43.87$. Thus, the exact same set of stimuli that showed no effect in size decision did show a congruency effect in animacy decision. This conclusion was further confirmed by a combined ANOVA of the data of Experiments 3 and 4 with experiment as a between-subjects factor and congruency as a within-subjects factor. This analysis showed a significant Experiment × Congruency interaction both for reaction times, $F(2, 224) = 4.64$, $p < .05$, $MSE = 429.10$; and for error percentages, $F(2, 224) = 3.28$, $p < .05$, $MSE = 33.04$.

To summarize, the congruency effect obtained in Experiment 3 disappeared in a different task, size decision, in which animacy congruency was no longer relevant for the task. There is another type of congruency, however, that is relevant in the size-decision task, namely whether a target has neighbors that are congruent in size. That is, for a target that is larger than a television, all neighbors that are also larger than a television are congruent, and all neighbors that are smaller than a television are incongruent. Because of the same mechanism as in animacy decision, performance in size decision should be affected by the size congruency of the target's neighbors. Because we did not manipulate size congruency, we did not have matched sets of congruent, neutral,

Table 4
*Mean Reaction Times (in ms) and Error Rates in the Size Decision Task of Experiment 4*

| | RT | | ER | |
|---|---|---|---|---|
| Neighborhood condition | *M* | *SE* | *M* | *SE* |
| Mostly congruent | 798 | 15.7 | 17.6 | 1.04 |
| Neutral | 798 | 15.8 | 17.5 | 1.04 |
| Mostly incongruent | 792 | 14.4 | 16.9 | 1.12 |
| Congruency effect | −6 | | −0.7 | |

*Note.* Congruency is defined in terms of animacy. See text for explanation. RT = reaction time; ER = error rate.

and incongruent targets. Therefore, we calculated the difference between the number of congruent and incongruent neighbors for each target, so that a target with more congruent than incongruent neighbors had a positive number (e.g., 2 if the target had 3 congruent neighbors and 1 incongruent neighbor) and a target with fewer congruent than incongruent neighbors had a negative number. We correlated this congruency score with reaction time and error percentage for each target.[3] The correlation between the congruency score and reaction time was $-.225$ ($p < .001$), and the correlation between the congruency score and error percentage was $-.254$ ($p < .001$). The fact that these correlations were negative indicates that performance was better for targets with more size-congruent neighbors than it was for targets with more size-incongruent neighbors. Note that these results should be treated with caution, because the targets were not matched on other variables. Nevertheless, the correlation analysis provides some additional support for the assertion that semantic congruency affects performance in semantic-decision tasks.

## General Discussion

The present results clearly demonstrate that performance in semantic-categorization tasks is affected by the existence and availability of orthographically similar words. All experiments showed that the direction of this orthographic similarity effect is critically dependent on semantic congruency. In general, orthographic neighbors that belonged to the same semantic category (e.g., animate or inanimate in an animacy decision task) as the target word increased classification performance, and neighbors that belonged to a different semantic category than the target word decreased classification performance.

In Experiments 1 and 2, we experimentally increased the availability of orthographically similar words through prior presentation of these words in a study block. Performance was better for targets (e.g., *cat*) for which an orthographic neighbor from the same semantic category (e.g., *rat*) had been studied than it was for targets for which an orthographic neighbor from the different semantic category (e.g., *mat*) had been studied. To the best of our knowledge, in Experiments 1 and 2 of this study, we were the first to study orthographic similarity effects in semantic-categorization tasks using a long-term priming paradigm. In other studies, researchers have shown that prior study of orthographically similar words affects performance, but those results were obtained in tasks that do not require access to the word's meaning (e.g., Bowers et al., 2002; Ratcliff & McKoon, 1997; Wagenmakers et al., 2000). It is interesting to note that in a lexical decision task, Bowers et al. (2002) obtained facilitation for rhyming neighbors but no effect for nonrhyming neighbors. These results might also be considered a congruency effect, but one at the phonological level instead of the semantic level.

In Experiment 3 of this study, we showed that orthographically similar neighbors affect performance in animacy decision even without prior study in the experimental context. Performance for words that mostly have semantically congruent neighbors was better than performance for words that mostly have semantically incongruent neighbors. In previous studies, researchers have investigated neighborhood effects in semantic tasks, but none of these researchers have taken semantic congruency into account. The results of these studies have been inconsistent. The results of the present study show that one important factor in determining the effect of orthographic neighbors in semantic tasks is semantic congruency.[4] Experiment 4 showed that the specific type of congruency is task-dependent. Congruency in terms of animacy affected performance in animacy decision but not in size decision. The correlation analyses from Experiment 4 further indicated that when congruency was formulated in terms of size decision, we did find congruency effects in the size-decision task.

These novel findings strongly suggest that semantic information from the target and its neighbors becomes available before a word has been uniquely identified (but see Forster & Hector, 2002). This implies that the processes of word identification and activation of semantic information occur in cascade. In contrast, serial models in which the activation of semantics would proceed only after a word has been uniquely identified do not predict the observed effect of semantic congruency of orthographic neighbors. That is, serial models can predict neighborhood effects per se, because a high number of neighbors or a high-frequency neighbor may interfere with the identification of the target word. However, serial models do not predict an effect of semantic congruency, because according to these models, semantic information does not become available until after a unique word has been selected. Once a word has been selected, its neighbors cease to be relevant. Thus, in order to explain our results, a model of word recognition needs to have a mechanism that allows semantic information to become available before a unique word has been selected.

An important issue is how a semantic-category decision is made. Carreiras et al. (1997) and Forster and Hector (2002) have proposed that in animacy decision, a "yes" decision is made if there is a certain amount of semantic activation that represents animacy. A "no" decision is made if a deadline is reached without sufficient evidence for animacy. This is similar to the mechanism that explains lexical decision performance in the multiple read-out model (Grainger & Jacobs, 1996). However, this explanation predicts a congruency effect for animate targets but not for inanimate targets. For animate targets, evidence of animacy accumulates faster if there are congruent neighbors than it does if there are incongruent neighbors. For inanimate targets, however, the decision is based on the length of the deadline (not on the amount of evidence supporting an "inanimate" decision). An "inanimate" decision is made if the amount of evidence for an "animate" decision does not surpass the animate criterion within the deadline period. In contrast with this prediction, our data show congruency effects for both animate and inanimate targets.

A related finding was reported by Forster and Hector (2002). They compared performance for nonwords (e.g., *turple*) with an exemplar neighbor (e.g., *turtle*) with performance for nonwords (e.g., *tabric*) with a nonexemplar neighbor (e.g., *fabric*) and found a congruency effect; that is, nonwords derived from an exemplar were classified more slowly than nonwords derived from a nonexemplar (in their experiment, nonwords had to be classified as nonexemplars). Recently, Rodd (2004) replicated these findings using word targets.[5] As in the Forster and Hector study, the effect of neighbors was investigated for nonexamplars only. For words

---

[3] We thank Jeroen Raaijmakers for suggesting this analysis.

[4] Note that the results of Experiment 3 do not address the question of whether neighborhood density per se affects performance, because neighborhood density was matched among the three conditions.

[5] We thank Jeff Bowers for drawing our attention to this article.

(e.g., *leotard*) that had an animal neighbor (e.g., *leopard*), participants were slower in deciding that the word was not an animal name than they were for words (e.g., *toffee*) that had a nonanimal neighbor (e.g., *coffee*). These results are problematic for a model in which it is assumed that "no" decisions are based on a deadline criterion.

The finding of a congruency effect for inanimate targets is not problematic, however, for cascaded processing models in general if the assumption of a deadline criterion is dropped. For example, cascaded models can explain congruency effects for inanimate targets if one assumes that decisions are based on the amount of evidence for animacy and the amount of evidence for inanimacy. Forster and Hector (2002) dismissed the idea of evidence for a "no" decision on the grounds that it is implausible that words are coded on every semantic feature that they do not possess. But suppose that there are features that inanimate targets possess that provide evidence for a "no" decision. In animacy decision, for example, the feature *made of stone* or the feature *man-made object* provides evidence for a "no" decision. Thus, it may not be necessary to postulate a deadline for "no" decisions in semantic-categorization tasks.

In sum, this study showed that performance for target words in semantic-categorization tasks is affected by orthographically similar words. Performance is enhanced if a neighbor from the same category is primed or if a word has mostly neighbors from the same category. Performance is decreased if a neighbor from the opposite category is primed or if a word has mostly neighbors from the opposite category. These results pose problems for serial models of word processing. In order to explain our results, it should be assumed in these models that semantic information becomes available before processing of orthographic information has been completed.

## References

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4,* 439–461.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database* (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

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.

Bowers, J. S. (1999). Priming is not all bias: Commentary on Ratcliff and McKoon (1997). *Psychological Review, 106,* 582–596.

Bowers, J. S., Damian, M. F., & Havelka, J. (2002). Can distributed orthographic knowledge support word-specific long-term priming? Apparently so. *Journal of Memory and Language, 46,* 24–38.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 857–871.

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). Hillsdale, New Jersey: Erlbaum.

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

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 696–713.

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

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure or equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior, 12,* 119–131.

Masson, M. E. J. (2002). Bias in masked word identification: Unconscious influences of repetition priming. *Psychonomic Bulletin & Review, 10,* 773–779.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86,* 287–330.

McKoon, G., & Ratcliff, R. (2001). Counter model for word identification: Reply to Bowers (1999). *Psychological Review, 108,* 674–681.

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.

Ratcliff, R., & McKoon, G. (1995). Bias in the priming of object decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 754–767.

Ratcliff, R., & McKoon, G. (1996). Bias effects in implicit memory tasks. *Journal of Experimental Psychology: General, 125,* 403–421.

Ratcliff, R., & McKoon, G. (1997). A counter model for implicit priming in perceptual word identification. *Psychological Review, 104,* 319–343.

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

Sears, C. R., Lupker, S. J., & Hino, Y. (1999). Orthographic neighborhood effects in perceptual identification and semantic categorization tasks: A test of the multiple read-out model. *Perception & Psychophysics, 61,* 1537–1554.

Vriezen, E. R., Moscovitch, M., & Bellos, S. A. (1995). Priming effects in semantic classification tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 933–946.

Wagenmakers, E. M., Steyvers, M., Raaijmakers, J. G. W., Shiffrin, R. M., van Rijn, H., & Zeelenberg, R. (2004). A model for evidence accumulation in lexical decision. *Cognitive Psychology, 48,* 332–367.

Wagenmakers, E. M., Zeelenberg, R., & Raaijmakers, J. G. W. (2000). Testing the counter model for perceptual identification: Effects of repetition priming and word frequency. *Psychonomic Bulletin & Review, 7,* 662–667.

Zeelenberg, R., & Pecher, D. (2003). Evidence for long-term cross-language repetition priming in conceptual implicit memory tasks. *Journal of Memory and Language, 49,* 80–94.

Zeelenberg, R., Wagenmakers, E. M., & Raaijmakers, J. G. W. (2002). Repetition priming in implicit memory tasks: Prior study causes enhanced discriminability, not only bias. *Journal of Experimental Psychology: General, 131,* 38–47.

Zeelenberg, R., Wagenmakers, E. M., & Shiffrin, R. M. (2004). Nonword repetition priming in lexical decision reverses as a function of study task and speed stress. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 270–277.

Received April 11, 2003
Revision received June 4, 2004
Accepted June 25, 2004 ■