Journal of Experimental Psychology: General
2008, Vol. 137, No. 2, 303–323

0096-3445/08/$12.00 DOI: 10.1037/0096-3445.137.2.303

Dynamics of the Attentional Control of Word Retrieval: Analyses of Response Time Distributions

Ardi Roelofs
Nijmegen Institute for Cognition and Information and F. C. Donders Centre for Cognitive Neuroimaging

Since W. Wundt (1904) and H. J. Watt (1906), researchers have found no agreement on how goals direct word retrieval. A prevailing associative account (E. K. Miller & J. D. Cohen, 2001) holds that goals bias association strength, which determines retrieval latency and whether irrelevant words interfere. A symbolic account (A. Roelofs, 2003) holds that goals enable retrieval rules and predicts no strict dependence of interference on latency. Here, 3 chronometric experiments in which the role of relative retrieval latency was investigated through distributional analyses, following Watt, are reported. Participants verbally categorized picture–word pairs that were semantically related or unrelated, or they categorized single pictures or words. The pairs yielded semantic latency effects in both word and picture categorizing, although single words were categorized slower than single pictures. Semantic effects occurred in word categorizing even when postexposure of the pictures compensated for the difference in categorizing latency. Vincentile and ex-Gaussian analyses revealed that the semantic effects occurred throughout the latency distributions, excluding goal neglect as the cause of the effects. The results were interpreted as most consistent with the symbolic account, which was corroborated by computer simulations.

Keywords: attention, computational modeling, distributional analysis, memory retrieval, word production

In the early days of experimental psychology, Wundt (1904) criticized the now classic, associative Wernicke-Lichtheim model of word production by arguing that the retrieval of words from memory is an active goal-driven process rather than a passive associative process, as held by the model. According to Wundt, an attentional process centered in the frontal lobes of the human brain controls a word perception and production network located in perisylvian brain areas, described by the Wernicke-Lichtheim model. Whereas Wundt examined goal-driven processes in perception, Watt (1906) experimentally investigated goal-directed word retrieval. Successful realization of task-relevant associations, he demonstrated, cannot simply be the result of passive associations between stimuli and word responses. Instead, task goals play a vital role in determining the direction of the retrieval process. A century later, however, it is still a hotly debated issue how exactly goals direct word retrieval processes.

The work reported in the present article is intended to shed light on the role of one factor in the word retrieval process, namely, relative retrieval latency. I begin by briefly describing the classic study of Watt (1906) and the seminal theoretical accounts of Selz (1913) and Müller (1913). Next, I discuss current theories of the control of word retrieval, which are descendents of these early theoretical ideas. I then present three chronometric experiments on categorizing pictures and words that examined the relative merits of two theories, namely, the associative account of Cohen, Dunbar, and McClelland (1990) and Miller and Cohen (2001) and the symbolic account of Roelofs (1992, 2003). Next, the utility of the symbolic account is demonstrated through computer simulations of the key findings. Theoretical consequences of the results are discussed in a general discussion section.

Historical Background to the Problem

Since Aristotle, it has been widely believed that human memory is in some way organized in an associative fashion. Retrieval from memory proceeds via associations that lead from one memory element to another. The path through memory is determined by the strength of the associations between successive memory elements. This associative notion of memory has been a dominant view for over 2,000 years (e.g., Anderson & Bower, 1973; Boring, 1950; Humphrey, 1951; G. Mandler, 2007; J. M. Mandler & Mandler, 1964). A major contribution of Watt (1906) was to experimentally test this view and to demonstrate that it could not easily explain his findings on constrained association (Boring, 1950; Gibson, 1941; Humphrey, 1951; G. Mandler, 2007; J. M. Mandler & Mandler, 1964 Woodworth, 1938).

Instead of asking participants to produce any association to a printed stimulus word (e.g., Cattell & Bryant, 1889, Watt (1906) asked his German participants for word associations that satisfied a particular semantic relation or *Aufgabe* (task). For example, if the task was to produce a coordinate term and the stimulus word was HAMMER, the response could be "saw," but with the same

---

Ardi Roelofs, Nijmegen Institute for Cognition and Information, Nijmegen, the Netherlands, and F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, the Netherlands.

I wish to thank Herman Kolk for the *Aufgabe* and indispensable discussion and Sanne Verhoef for her help in running the experiments. The preparation of the article was supported by a VICI grant from the Netherlands Organisation for Scientific Research.

Correspondence concerning this article should be addressed to Ardi Roelofs, Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, Spinoza Building B.01.08, Montessorilaan 3, 6525 HR Nijmegen, the Netherlands. E-mail: A.Roelofs@nici.ru.nl

stimulus word and the task used to produce a category term, the response would be "tool." Other tasks were to name a part, a whole, another part of the same whole, or a subordinate. The time elapsing between the onset of a printed stimulus word and the onset of a vocal response to that stimulus, henceforth, the *response latency,* was measured using a voice key. Also, introspective reports were recorded and analyzed. The study of Watt had several limitations (e.g., a small number of participants, no proper statistical analysis), but it is important because of its instructive nature and lasting relevance.

Watt (1906) observed that the time of responding to the printed words was determined by the type of task (e.g., produce a coordinate term, produce a category term) on the one hand and the strength of the association between stimulus and verbal response on the other. He found that when verbal responses of the same intrinsic speed were grouped together, a variation of task had a similar effect across latency groups. He stated, "The influence of the task is *independent* of the rapidity of the tendency to reproduction itself" (Watt, 1906, p. 260), henceforth, *Watt's regularity.* Yet the verbal response is appropriate to the task at hand. Selz (1913, 1922, 1927) argued that the dependence of the type of word response on its logical relationship to the task goal could not easily be explained by an associative memory consisting of simple unlabeled links. If *saw* is more strongly connected to *hammer* than is the category *tool,* classic association theory predicts that "saw" is produced in response to HAMMER regardless of the task, unlike what is empirically observed. Moreover, the putatively strongest association of all, the response "hammer" itself, is also not produced. This raises the question of how participants are able to retrieve the task-relevant word from associative memory.

As a solution to the control problem, Selz (1913, 1922, 1927) hypothesized that the stimulus and the task goal together form the basis for the elicitation of a series of specific, reflexlike operations that produce the relevant verbal response. In more modern terms, the task-relevant response is recalled from associative memory by the application of a series of condition–action production rules (Simon, 1981). The condition of a production rule refers to the task goal as well as the stimulus word. For example, the task of retrieving a category term from memory would be accomplished by the production rule

(1) IF the goal is to produce a category term for word *x*
AND memory specifies that *x* is a *y*
THEN select *y*.

The task of producing a coordinate term may be achieved by a production rule that tests for a common superordinate. "Such rules, together with related Goal activations, we assume, are constitutive of intentional task control, or task-set" (Allport & Wylie, 1999, p. 277).

In response to the work of Watt (1906) and in defense of an associative view of the control of word retrieval, Müller (1913) argued that an instruction like "Produce a category term" acts as a directive representation, which selectively reinforces particular associations, such as *tool, furniture,* and *animal*. The activation from the directive representation endorses the task set. Among the associations put into heightened readiness by the task would also be the association for the particular stimulus word HAMMER. Consequently, among the several associations of HAMMER, the instruction "Produce a category term" reinforces the association with *tool,* which yields the correct response "tool."

Selz (1927) criticized this associative solution of Müller (1913), however, by arguing that directive representations like "Produce a category term" reinforce far too many associations to be effective. For example, "Produce a category term" should also reinforce the response "hammer," because *hammer* is a category itself (i.e., the category including *mallet, sledge hammer,* etc.). Moreover, within this associative framework, it is unclear how the goal of producing a coordinate term could be accomplished. An instruction like "Produce a coordinate term" would reinforce not only *saw* but also, again, *hammer* itself because *hammer* is a coordinate of *saw*. Thus, as a general solution to the control problem, an associative approach seems to fail. Still, it remains possible that memory contains directive representations favoring specific levels of abstraction, such as subordinate, basic, and superordinate levels of responding. This might solve the control problem in the specific case of categorization. That is, a directive representation saying "Produce a term at superordinate level" connected to *furniture, tool,* and *animal* might be effective in producing the category term "tool" in response to HAMMER. This solution is assumed in the present article.

Current Theoretical Approaches

Current theories of the control of word retrieval are in one way or another descendents of the theoretical ideas of Selz (1913, 1922, 1927) and Müller (1913). According to the prevailing associative account of the control problem in the literature (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001; Cohen et al., 1990; Melara & Algom, 2003; Miller & Cohen, 2001; Phaf, Van der Heijden, & Hudson, 1990), goal representations associatively bias the activation of one type of verbal response over another, following Müller. The present article concentrates on the computational account of Cohen et al. (1990) and Miller and Cohen (2001) because it is probably the most influential and most quoted, implemented associationistic theory of the control of word retrieval in the literature. The account is illustrated in Figure 1.

Figure 1 depicts a fragment of a network that links a pictured hammer and the printed word HAMMER via an intermediate level of nodes to the response words "hammer" and "tool." The asso-

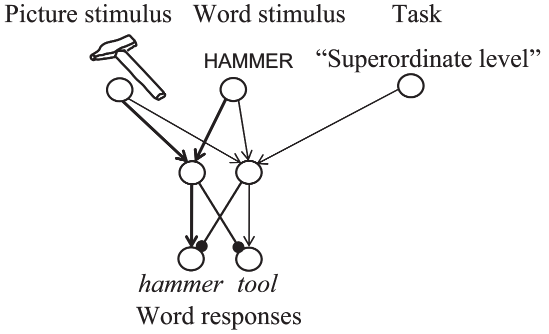

Figure 1. Illustration of the network structure of the associative account of the control of word retrieval advanced by Cohen et al. (1990) and Miller and Cohen (2001). The lines with filled circles indicate inhibitory connections.

ciations to *hammer* are assumed to be stronger than those to *tool,* as indicated by the thickness of the lines. Without task input that biases a particular response type, the responding of the network is determined by the strength of the associations between stimuli and responses. When presented with a pictured hammer or the word HAMMER, the network yields the response "hammer." However, if the directive representation "Produce a term at superordinate level" is activated, associations to superordinate-level responses are put into a heightened state of readiness. If now the pictured hammer or the word HAMMER is presented, the network generates the response "tool" rather than "hammer." This account is similar to the account that has been given for the control of responding in the classic color–word Stroop task (Botvinick et al., 2001; Cohen et al., 1990; Melara & Algom, 2003; Miller & Cohen, 2001; Phaf et al., 1990), which requires naming the ink color of incongruent color words (e.g., the word RED in blue ink) or reading the words aloud (Stroop, 1935).

Symbolic theories of the control of word retrieval, such as proposed in Roelofs (1992, 2003) and implemented in the WEAVER++ model (e.g., Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992, 2003, 2006), also assume that word information is represented in an associative network. However, it is acknowledged that a mere associative link between two nodes tells nothing about the relation between the entities represented. For example, *hammer* is associated with both *tool* and *saw,* but the relationship between *hammer* and *tool* is very different from the relationship between *hammer* and *saw*. Therefore, the relationship between memory elements is encoded by labeling the links, following Selz (1913, 1922, 1927). The network is illustrated in Figure 2. Condition–action rules determine what is done with the activated network information depending on the goal and the information on the links (cf. Anderson, 1983; Selz, 1927). When a goal symbol is placed in working memory, the retrieval process is controlled by those condition–action rules that include the goal among their conditions. For example, if the task is to produce a category term and the stimulus is the word HAMMER, a condition–action rule like Production Rule 1 fires, and "tool" is produced. In categorizing, selection is governed by links that are labeled IS-A. The goal and the labeled links work together in retrieving the appropriate word (cf. Woodworth, 1938). A similar Selzian account has been given for responding in the Stroop task (Roelofs, 2003).

The associative account of Cohen et al. (1990) and Miller and Cohen (2001) and the symbolic account of Roelofs (1992, 2003) make different predictions about the effect of irrelevant stimulus–response associations on the use of a target stimulus–response association. According to the associative account, whether irrelevant associations have an influence on responding depends on the relative strength of the associations involved. The inadvertent activation of strong stimulus–response associations interferes with the use of weaker stimulus–response associations for responding, but not the other way around (see Cohen et al., 1990, pp. 348–349). Association strength is reflected in the speed of responding. Cohen et al. (1990) stated, "The strength of a pathway determines speed of processing and whether one process will influence (interfere with or facilitate) another" (pp. 343–344). This also holds for similar models (Melara & Algom, 2003; Phaf et al., 1990). Assume, for example, that participants have to categorize the word HAMMER paired with a pictured dog. According to the associative account, if the association between the word HAMMER and the response "tool" is weaker than the association between the pictured dog and the response "animal" (and everything else is equal), the pictured dog should delay the production of "tool" in response to the word HAMMER. However, the word HAMMER should not delay saying "animal" in response to the pictured dog. In contrast, according to the symbolic account proposed by Roelofs (1992, 2003), both words and pictures activate categorical information and the associated word responses in memory, and therefore interference effects should be obtained in both tasks. That is, the word HAMMER should delay saying "animal" to a pictured dog, and the pictured dog should delay saying "tool" to HAMMER. Because pictures have priority access to concepts before lexical information, whereas the reverse holds for printed words (see Figure 2), the delay caused by distractor pictures in word categorizing should be larger than the delay caused by distractor words in picture categorizing (Roelofs, 1992).

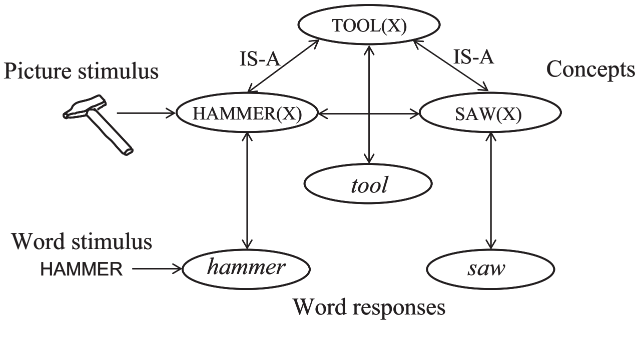

Figure 2. Illustration of the labeled network of the symbolic account of the control of word retrieval advanced by Roelofs (1992, 2003) and implemented in WEAVER++. Only IS-A links are shown.

Evidence on Categorizing Picture–Word Pairs

Although several researchers have examined the categorizing of single pictures or words, the literature on the categorizing of picture–word pairs is limited (see W. R. Glaser, 1992, for a review). I discuss the studies that are most relevant to the present article.

In a seminal study, Smith and Magee (1980) obtained evidence that at first sight would seem to be in line with neither the symbolic nor the associative account. Participants were presented with three types of stimulus cards: single cards containing only pictures, single cards containing only printed words, and combination cards containing picture–word pairs (i.e., words superimposed onto pictures). For incongruent combinations, the picture and word differed in both category and name (e.g., a pictured hammer combined with the word DOG). The participants were instructed to say "yes" if a picture or word was a member of a particular category prespecified for a card (e.g., *tools*) and "no" to all pictures and words falling outside the category. The time to complete each card was measured with a stopwatch. There was no significant difference in the speed of categorizing single words and pictures (the approximate mean latencies per word and picture were 563 ms and 579 ms, respectively). However, in the combination cards, word categorizing suffered considerable interference from incongruent pictures relative to the single cards, whereas picture categorizing was minimally affected by the presence of an incongruent word.

Thus, although the baseline speed of responding was matched between the pictures and words, asymmetrical interference was obtained, contradicting the predictions of the associative account (Cohen et al., 1990; Miller & Cohen, 2001). The account predicts no interference in either task if the baselines of responding are matched (see Cohen et al., 1990, pp. 348–349). Moreover, the symbolic account of Roelofs (1992, 2003) predicts that word categorizing should be slower than picture categorizing, but no such difference was obtained by Smith and Magee.

However, it is possible that there was no difference between picture and word categorizing latencies because of perceptual factors. In the study of Smith and Magee (1980), the physical size of the pictures was much greater than the size of the printed words. It is possible that in categorizing the cards with pictures only, participants suffered interference from seeing the adjacent pictures in their peripheral vision, whereas this was not or was less the case for the cards with words only. Consequently, the baseline latency of picture categorizing may have been overestimated. The difference in physical size between pictures and words may also have caused or contributed to the asymmetry in interference between the tasks performed on the combination cards (Melara & Algom, 2003). The physically small words may have yielded less interference in categorizing the physically large pictures than vice versa, yielding the asymmetry in interference effects. Thus, the evidence from Smith and Magee does not really challenge an associative account either.

W. R. Glaser and Düngelhoff (1984) reported evidence from an experiment using picture and word categorizing tasks that would seem to support the symbolic account and contradict the associative account. One group of participants categorized the pictures of picture–word pairs while trying to ignore superimposed printed words, and another group of participants categorized the words while trying to ignore the pictures. The picture–word stimuli were not presented on cards, but they were presented individually on a computer screen. Moreover, the participants did not say “yes” or “no,” but produced category terms, such as “tool,” in response to HAMMER. For each stimulus individually, the response latency was measured. Furthermore, the pictures and words were presented at various stimulus onset asynchronies (SOAs). The onset of the distractor could be 400, 300, 200, or 100 ms earlier than the onset of the target (henceforth, *distractor preexposure SOAs,* indicated by a minus sign), the onset of target and distractor coincided, or the onset of the presentation of the distractor was 100, 200, 300, or 400 ms later than the onset of the presentation of the target (henceforth, *distractor postexposure SOAs*). The picture–word pairs had varying types of relatedness between picture and word. For example, participants said “tool” in response to the word HAMMER, while trying to ignore a pictured saw (the semantic condition), a pictured dog (unrelated), a pictured hammer (identical), or an empty rectangle (the baseline condition).

W. R. Glaser and Düngelhoff (1984) observed that for both word and picture categorizing, responses were faster in the semantic than in the unrelated condition, whereas the latencies in the semantic, identical, and control conditions did not differ much. Hereafter, the difference in response time between the semantic and unrelated conditions is referred to as the *semantic effect,* whereas slower responding in any condition relative to baseline is referred to as *interference.* The semantic effects of words in picture categorizing occurred at SOAs ranging from −400 ms to −100 ms, whereas the effects of pictures in word categorizing occurred at SOAs ranging from −400 ms to 200 ms. However, across SOAs, picture categorizing was on average 100 ms faster than word categorizing in the baseline condition. At SOA = 0 ms, the difference was 126 ms. The semantic effect at SOA = 200 ms of pictures in word categorizing does not agree with the predictions of the associative account.

However, a problem with the experiment of W. R. Glaser and Düngelhoff (1984) is that the tasks were performed by separate groups of participants, different from Smith and Magee (1980). The use of separate groups of participants makes the assessment of the baseline speed of responding to the words and pictures somewhat problematic (e.g., one of the groups may have been unusually fast or slow), and the same holds for a comparison of the magnitudes of the distractor effects between tasks. Moreover, Glaser and Düngelhoff statistically tested for effects of the semantic, unrelated, and identical conditions relative to baseline but not specifically for semantic effects (i.e., differences between the semantic and unrelated conditions). Therefore, it remains unclear whether the 52-ms semantic effect at SOA = 200 ms of pictures in word categorizing was really statistically significant. Furthermore, although the absence of an effect of words in picture categorizing at SOA = 0 ms (the magnitude was numerically 0 ms) and at postexposure SOAs agrees with the predictions of Cohen et al. (1990), Costa, Mahon, Savova, and Caramazza (2003) did observe a significant semantic effect (of 29 ms) of word distractors in picture categorizing with simultaneous presentation of picture and word, different from Glaser and Düngelhoff. Word categorization latencies were not measured by Costa et al., however, which complicates an interpretation of their findings.

A final factor complicating the interpretation of the previous finding must be discussed. As Watt (1906) recognized, it is possible that differences in mean latency between conditions or experiments reflect differences in attentional engagement. To address this problem, Watt recommended examination of response latency distributions. He stated,

> If the conditions could be kept as constant as they are in the shortest possible reactions, the curve of distribution would be quite as regular for any set of conditions whatever. Peculiarities in the form of the curve of distribution would then be *symptomatic* of peculiarities in the reactions or in the factors which bring about these. (Watt, 1906, p. 263)

De Jong, Berendsen, and Cools (1999) argued that one such factor may be *goal neglect.* Goal neglect was defined by Duncan (1995) as the ignoring of task demands even if they have been understood. Goal neglect may result in distractor effects that would not have been present if processing had been tightly focused on the goal and task demands (De Jong et al., 1999; Kane & Engle, 2003; West & Alain, 2000). Goal neglect may result in no or small condition differences for relatively fast responses (when attention is sharply focused on the instructed task) and disproportionately large condition differences for relatively slow responses (De Jong et al., 1999). That the participant group performing word categorizing in the experiment of W. R. Glaser and Düngelhoff (1984) suffered from goal neglect, yielding the observed semantic effect of picture distractors at postexposure SOAs, cannot be excluded on the basis of the available evidence from condition means. If so, the findings of Glaser and Düngelhoff would not really challenge the associative account.

Plan of the Present Study

The present article reports three chronometric experiments intended to further investigate the dynamics of goal-driven word retrieval. One of the aims was to examine the merits of the associative account of Cohen et al. (1990) and Miller and Cohen (2001) and the symbolic account of Roelofs (1992, 2003) of the control of word retrieval in categorizing. The experiments addressed the concerns about previous studies and, in particular, followed Watt's (1906) recommendation of examining response latency distributions. In Experiment 1, participants categorized the word or picture of picture–word pairs that could be semantically related or unrelated, or they categorized single pictures and words in a baseline condition. Semantic, unrelated, and baseline trials were randomly intermixed. Picture and word were presented simultaneously (SOA = 0 ms). Three s before the presentation of the imperative stimulus on each trial, participants received a task cue ("Word," "Picture") indicating the target for that trial. The cues varied randomly from trial to trial. The aim of the experiment was to assess the baseline speed of responding to the pictures and words and to test whether semantic effects are obtained in picture categorizing when word categorizing is slower than picture categorizing. If semantic effects were obtained in both tasks even when word categorizing was slower than picture categorizing, this would challenge the associative account and support the symbolic account advanced by Roelofs (1992, 2003).

In Experiment 2, participants categorized the word of the picture–word pairs while the SOA between picture and word was manipulated. The picture was presented 150 ms before, simultaneously with, or 150 ms after word onset. Presenting the picture 150 ms after word onset compensated for the difference in speed between word and picture categorizing. On the basis of the within-participant design of Experiment 1, the difference was estimated to be 88 ms for the present stimuli on average. If semantic effects were obtained with picture postexposure, this would challenge the associative account and support the symbolic account.

To examine the roles of goal neglect and relative response latency within a task, latency distributions were examined in both experiments. If distractor effects arise because processing is insufficiently focused on the goal on some of the trials, the semantic effects should be present for only part of the latency distribution, namely, for the slow responses only, or the effects should be disproportionately large for the slow responses (De Jong et al. 1999). Alternatively, if the semantic effects are due to the structural factors illustrated in Figures 1 and 2, the effects are expected to be present through the entire latency range, and the magnitude of the effects should not increase disproportionately with response latency. Watt (1906) grouped together verbal responses of the same intrinsic speed and found that a variation of task had a similar effect across latency groups, although he did not quantify the effect. By comparing the latency distribution between word and picture categorizing, it was examined whether Watt's regularity holds for the present tasks, and the effect of variation of task was quantified using *quantile–quantile plots* (Thomas & Ross, 1980; Wilk & Gnanadesikan, 1968). A quantile–quantile plot is a standard technique for determining whether two distributions have the same shape (i.e., belong to the same distribution family). If they do, the plot should be linear, indicating that the distributions differ only by a scale or shift factor. Linearity in a distribution plot is by no means a necessity for latency distributions (e.g., De Jong, Liang, & Lauber, 1994; Zhang & Kornblum, 1997), so it is possible that Watt's regularity does not hold for the present verbal tasks. In addition, the data were examined at the level of distributional characteristics using ex-Gaussian analyses (cf. Heathcote, Popiel, & Mewhort, 1991; Luce, 1986). These analyses allow differences between conditions to be separated into shifts of the entire distribution and changes in the tail. Such analyses help to assess whether differences between condition means are due just to a longer tail in the slower condition.

Elsewhere (Roelofs, 1992, 2003, 2005, 2006), I argued that picture and word categorizing require access to the conceptual level (illustrated in Figure 2) but that words may be read aloud via a shallow route from word-form perception to word-form encoding, without accessing lemmas (the level of word responses in Figure 2) or concepts (cf. Wundt, 1904). Words may be read aloud via grapheme-to-phoneme correspondence rules or connections between morphemes. The processing levels assumed for word categorizing, picture categorizing, and word reading are illustrated in Figure 3. Word categorizing involves word-form perception (e.g., perceiving HAMMER), lemma retrieval (i.e., retrieving *hammer*), conceptualizing (i.e., retrieving TOOL[*X*]), lemma retrieval (i.e., retrieving *tool*), and word-form encoding (Route a in Figure 3). Picture categorizing involves object-form perception (e.g., a pictured hammer), conceptualizing (i.e., retrieving TOOL[*X*]), lemma retrieval (*tool*), and word-form encoding (Route b). Finally, word reading involves word-form perception (e.g., HAMMER) and word-form encoding. Under this account (Roelofs, 1992, 2003), pictures should influence word categorizing and vice versa, but pictures should not influence word naming regardless of relative response latency. This claim was tested in Experiment 3 by comparing word reading and word categorizing and by performing distributional analyses. Also, it was examined whether Watt's regularity holds for word reading and word categorizing.

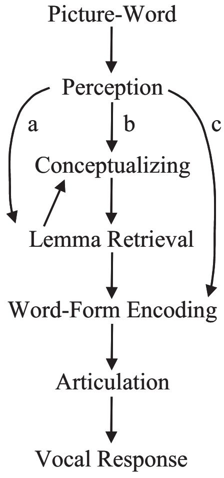

Figure 3. Illustration of the levels of processing involved in word categorizing (Route a), picture categorizing (Route b), and word reading (Route c), according to Roelofs (1992, 2003) and implemented in WEAVER++.

Experiment 1

In Experiment 1, the participants' task was to verbally categorize the pictured item while ignoring the superimposed word or to categorize the word while ignoring the pictured item. The task was cued well before presentation of the imperative stimulus, allowing participants to prepare for the task (cf. Watt, 1906). The latencies of the picture and word categorizing responses were measured. The pictures and words were either semantically related (e.g., the word HAMMER printed on a picture of a saw) or unrelated (e.g., HAMMER printed on a picture of a dog), or else a string of *X*s was superimposed on a picture (the baseline condition for picture categorizing) or the word appeared in an empty rectangle (the baseline condition for word categorizing). The aim of the experiment was to assess the baseline speed of responding to the pictures and words and to test whether semantic effects are obtained in picture categorizing when word categorizing is slower than picture categorizing. If semantic effects were obtained in both tasks even when word categorizing was slower than picture categorizing, this would challenge the associative account advanced by Cohen et al. (1990) and Miller and Cohen (2001), and it would support the symbolic account advanced by Roelofs (1992, 2003). Latency distributions were examined for baseline word and picture categorizing, for picture categorizing in the semantic and unrelated word contexts, and for word categorizing in the semantic and unrelated picture contexts.

Method

Participants. The experiment was carried out with a group of 16 paid participants, who were students at Radboud University Nijmegen, Nijmegen, the Netherlands. All participants were young adults and native speakers of Dutch. They were paid €4 ($6 U.S.) for their participation.

Materials and design. From a picture gallery, 32 pictured objects from eight different semantic categories were selected together with their basic-level names in Dutch. The Appendix lists the materials. The pictures were white line drawing on a black background. They were digitized and scaled to fit into a virtual frame of 10 cm × 10 cm. The printed words and the printed task cues "Word" and "Picture" in Dutch were presented in white color in 36-point lowercase Arial font.

The first independent variable was context. Each picture was combined with a printed word from the same semantic category (the semantic condition) or with a word from another semantic category (the unrelated condition), or else a string of *X*s was superimposed on a picture (the baseline condition for picture categorizing) or the word appeared in an empty rectangle (the baseline condition for word categorizing). The semantic and unrelated conditions were created by recombining pictures and words. All target pictures and words occurred equally often in all conditions. The Appendix gives the picture–word pairings. There were 32 stimuli in each of the three distractor conditions, yielding 96 different stimuli in total. The order of presenting the stimuli across trials was random, except that repetitions of pictures and words on successive trials were not permitted. The second independent variable was target. On each trial, participants had to categorize the picture or word, depending on the task cue.

Procedure and apparatus. The participants were tested individually. They were seated in front of a computer monitor (NEC Multisync) and a Sennheiser microphone connected to an electronic voice key. The distance between participant and screen was approximately 50 cm. The participants were asked to categorize the word or picture depending on the cue.

After a participant had read the instructions, a block of 64 practice trials was administered. During the practice session, all pictures and printed words were presented in isolation, and they were categorized once. After this, testing began. The structure of a test trial was as follows. A trial started with the presentation of the visual cue "Word" or "Picture" in Dutch for 800 ms. The screen was then blanked for 2,200 ms, followed by the presentation of the picture–word stimulus for 2,000 ms. The vocal response latency was measured to the nearest millisecond from target stimulus onset. The participants were told that the response had to be initiated before the picture–word pair disappeared from the screen (i.e., within 2.0 s after picture–word presentation onset). Before the start of the next trial, there was a blank interval of 1.0 s. Thus, the total duration of a trial was 6.0 s. An IBM-compatible computer controlled the stimulus presentation and data collection, including the voice key.

Analysis. After each trial, the experimenter coded the response for errors. Five types of incorrect responses were distinguished: wrong response word, wrong pronunciation of the word, disfluency, triggering of the voice key by a nonspeech sound, and failure to respond within 2.0 s after target presentation onset. Incorrect responses were discarded from the statistical analyses of the response latencies. The latencies and errors were submitted to analyses of variance with the crossed variables target and context. Both variables were tested within participants. The analyses were performed both by participants, denoted by F_1 and t_1, and by items, denoted by F_2 and t_2 (Clark, 1973).

To obtain the latency distributions, the rank-ordered latencies for each participant were divided into deciles (10% quantiles), and mean latencies were computed for each decile, separately for the baseline word and picture categorizing responses, for the picture categorizing responses in the semantic and unrelated picture contexts, and for the word categorizing responses in the semantic and unrelated word contexts. By averaging these decile means across participants for each condition, Vincentized cumulative distribution curves were obtained (Ratcliff, 1979). For response latencies, Vincentizing the data across individual participants provides a way of averaging data to obtain group distributions while preserving the shapes of the individual participant distributions (cf. Woodworth & Schlosberg, 1954). Distributions were also obtained on the basis of the rank-ordered latencies for the items. The latencies for the items were ordered by target response rather than by individual word or picture to have a sufficient number of data points per decile. All distribution plots in the present article show response latency on the horizontal axis and cumulative relative frequency (estimating cumulative probability) on the vertical axis (cf. Hohle, 1965; Luce, 1986, p. 101). The cumulative distribution functions were examined by conducting analyses of variance (cf. Bub, Masson, & Lalonde, 2006; De Jong et al., 1999; Ridderinkhof, 2002).

In addition, the data were examined at the level of distributional characteristics using ex-Gaussian analyses. Whereas Vincentizing does not depend on prior distributional assumptions and examines the raw data directly, ex-Gaussian analyses characterize a response latency distribution by assuming an explicit function for the shape of the distribution. Ex-Gaussian functions provide good fits of empirical response time distributions and have been widely adopted (e.g., Heathcote et al., 1991; Hohle, 1965; Luce, 1986; Spieler, Balota, & Faust, 1996; Yap & Balota, 2007). The ex-Gaussian function consists of a convolution of a Gaussian and an exponential distribution, and it has three parameters. These are μ and σ, the mean and standard deviation of the Gaussian distribution, and τ, reflecting the mean and standard deviation of the exponential distribution. Theoretically, the sum of μ and τ is equal to the mean of the overall distribution (Hohle, 1965; Luce, 1986). Ex-Gaussian analyses allow differences between conditions to be separated into distributional shifting, reflected in μ, and distributional skewing, reflected in τ. By performing the analyses, it may be assessed whether experimental manipulations shift the entire distribution or affect the tail only, complementing the Vincentile analyses. The distributional parameters of the present data were estimated using the continuous maximum-likelihood method proposed by Van Zandt (2000). The ex-Gaussian parameters (μ, σ, and τ) were obtained for each participant individually using the program of Brown and Heathcote (2003) with continuous maximum-likelihood fitting. The parameters were then examined by conducting analyses of variance (e.g., Heathcote et al., 1991; Spieler et al., 1996; Yap & Balota, 2007).

Results and Discussion

Figure 4 gives the mean categorizing latencies and the mean error percentages for target and context. There were virtually no technical errors; most errors involved wrong response words. Therefore, the types of errors were not separated. Figure 4 shows that semantic effects were obtained for both word and picture targets. Categorizing was slower in the unrelated than in the semantic context. The semantic effect was larger for word categorizing than for picture categorizing. Word categorizing was slower than picture categorizing in the baseline condition.

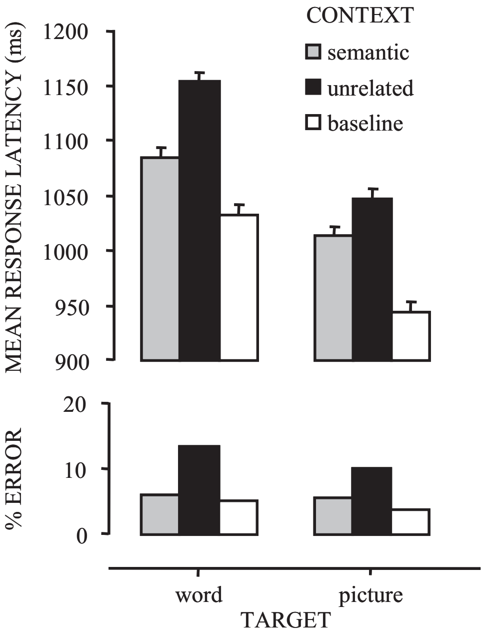

Figure 4. Mean response latencies and error percentages per context and target in Experiment 1. The error bars indicate the standard error of the mean.

The statistical analysis of the vocal categorizing latencies yielded main effects of target, $F_1(1, 15) = 56.68$, $p < .001$, $F_2(1, 31) = 80.18$, $p < .001$, and context, $F_1(2, 30) = 47.72$, $p < .001$, $F_2(2, 62) = 74.25$, $p < .001$. Target and context did not interact overall, $F_1(2, 30) = 1.45$, $p = .25$, $F_2(2, 62) = 1.71$, $p = .19$. There was a main effect of semantic relatedness, $F_1(1, 15) = 27.79$, $p < .001$, $F_2(1, 31) = 26.46$, $p < .001$. Planned comparisons showed that the semantic effect was larger for the words than for the pictures, $t_1(15) = 1.80$, $p < .05$, $t_2(31) = 1.83$, $p < .04$. The semantic effect was present for both the picture targets, $t_1(15) = 3.12$, $p < .003$, $t_2(31) = 2.36$, $p < .02$, and the word targets, $t_1(15) = 4.56$, $p < .001$, $t_2(31) = 5.60$, $p < .001$. Picture categorizing was faster than word categorizing in the baseline condition, $t_1(15) = 4.83$, $p < .001$, $t_2(31) = 7.03$, $p < .001$.

Figure 4 shows that more errors were made in the unrelated condition than in the other conditions, which did not differ much. The error percentages were slightly higher for word categorizing than for picture categorizing. The statistical analysis of the errors yielded main effects of target, $F_1(1, 15) = 13.29$, $p < .002$, $F_2(1, 31) = 10.67$, $p < .003$, and context, $F_1(2, 30) = 14.10$, $p < .001$, $F_2(2, 62) = 65.44$, $p < .001$. Target and context did not interact, $F_1(2, 30) = 1.57$, $p = .23$, $F_2(2, 62) = 1.63$, $p = .204$. Given that most errors were made in the task and condition with the slowest responses (i.e., the unrelated condition in word categorizing), there is no evidence for a speed–accuracy tradeoff in the data.

Figure 5 shows the participant-based latency distributions for baseline word and picture categorizing (left panel), picture categorizing in the semantic and unrelated word contexts (middle panel), and word categorizing in the semantic and unrelated picture contexts (right panel). The item-based latency distributions were equivalent. The left panel of Figure 5 shows that baseline picture categorizing was faster than baseline word categorizing throughout the latency distribution. Statistical analysis yielded a main effect of target, $F_1(1, 15) = 22.74$, $p < .001$, $F_2(1, 7) = 38.69$, $p < .001$, which depended on decile, $F_1(9, 135) = 3.64$, $p < .001$, $F_2(9, 63) = 5.11$, $p < .001$. However, although the magnitude of the difference in baseline speed increased with the absolute latency, picture categorizing was faster than word categorizing for the first decile, $t_1(15) = 4.26$, $p = .001$, $t_2(7) = 6.64$, $p = .001$, as well as the 10th decile, $t_1(15) = 2.89$, $p = .01$, $t_2(7) = 4.65$, $p = .002$. The middle panel of Figure 5 shows that picture categorizing was faster in the semantic than in the unrelated contexts through almost the entire latency range. Statistical analysis yielded a main effect of context by participants, $F_1(1, 15) = 11.01$, $p = .005$, $F_2(1, 7) = 4.54$, $p = .07$,[1] which was independent of decile, $F_1(9, 135) = 0.36$, $p = .95$, $F_2(9, 63) = 0.25$, $p = .98$. Thus, the magnitude of the semantic effect did not increase overproportionately with response latency. Instead, the constancy of the semantic effect indicates that it varied less than proportionally with latency. The right panel of Figure 5 shows that word categorizing was also faster in

---

[1] The by-item analysis here was based on the nine different target responses, which had less statistical power than the by-item analysis based on the 32 different pictures/words reported earlier. This explains why the semantic effect was only marginally significant in the analysis here but fully significant in the analysis based on the individual pictures/words.

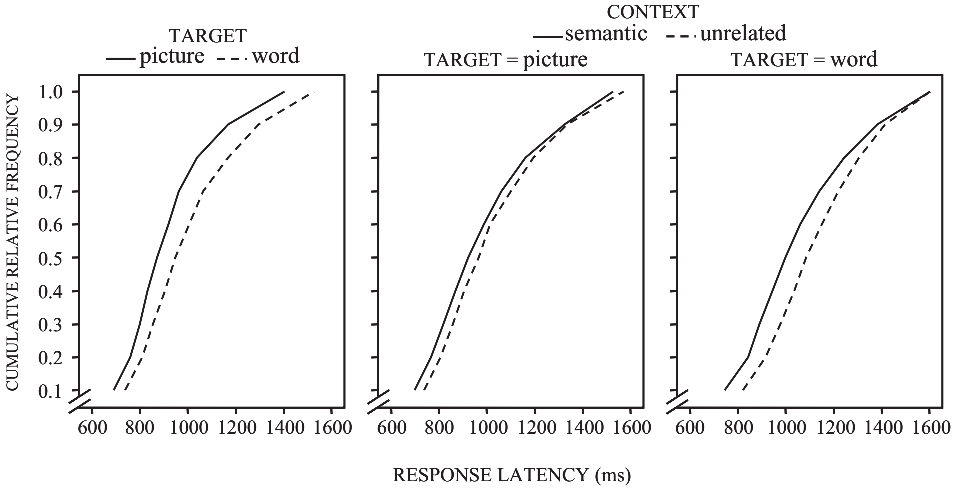

Figure 5. Vincentized cumulative distribution curves for the baseline word and picture categorizing latencies (left panel), for the picture categorizing latencies in the semantic and unrelated word contexts (middle panel), and for the word categorizing latencies in the semantic and unrelated picture contexts (right panel).

the semantic than in the unrelated contexts through almost the entire latency range. Statistical analysis yielded a main effect of context, $F_1(1, 15) = 21.81$, $p < .001$, $F_2(1, 7) = 110.96$, $p < .001$, which depended on decile, $F_1(9, 135) = 4.05$, $p < .001$, $F_2(9, 63) = 5.60$, $p < .001$. The semantic effect was present in the first decile, $t_1(15) = 5.72$, $p < .001$, $t_2(7) = 7.06$, $p < .001$, but not in the 10th decile, $t_1(15) = 0.00$, $p = .95$, $t_2(7) = 0.17$, $p = .86$. Thus, again, the magnitude of the semantic effect did not increase overproportionately with response latency.

Although the magnitude of the difference in baseline speed between word and picture categorizing increased with the absolute latency, picture categorizing was faster than word categorizing over the entire latency distribution. Moreover, further analyses revealed that the increase with latency was not disproportionate. The word and picture categorizing latencies were linearly related, as illustrated by the quantile–quantile plots in Figure 6. The figure shows the relationship between word and picture categorizing latencies based on a grouping of the latencies by participants (see Figure 6A) and by target responses (see Figure 6B). This finding of a linear relationship between picture and word categorizing latencies is in line with the empirical regularity observed by Watt (1906). The influence of the task in Experiment 1 was also independent of the relative speed of responding.

The ex-Gaussian analyses revealed that baseline picture and word categorizing differed in μ (761 ms vs. 817 ms, respectively), $F(1, 15) = 6.03$, $p = .027$; but not in σ (76 ms vs. 92 ms), $F(1, 15) = 1.26$, $p = .28$; and also not in τ (185 ms vs. 217 ms), $F(1, 15) = 2.72$, $p = .12$. Thus, relative to picture categorizing, word categorizing shifted the latency distribution rightward without affecting the standard deviation and tail. In picture categorizing, the semantic effect was reflected in μ (798 ms and 848 ms for the semantic and unrelated distractors, respectively), $F(1, 15) = 6.84$, $p = .019$; but not in σ (105 ms vs. 118 ms), $F(1, 15) < 1$, $p = .54$; and also not in τ (216 ms vs. 202 ms), $F(1, 15) = 1.12$, $p = .31$. Thus, relative to the unrelated distractors, the semantic distractors shifted the latency distribution leftward without affecting the standard deviation and tail. Finally, in word categorizing, the semantic effect was reflected in μ (869 ms vs. 992 ms), $F(1, 15) = 10.90$, $p = .005$; but not in σ (124 ms vs. 133 ms), $F(1, 15) < 1$, $p = .70$; and only marginally in τ (219 ms vs. 164 ms), $F(1, 15) = 4.07$, $p = .06$. Thus, relative to the unrelated distractors, the semantic distractors shifted the latency distribution leftward and increased its tail somewhat while leaving the standard deviation unaffected (cf. Heathcote et al., 1991; Spieler et al., 1996). To conclude, the ex-Gaussian analyses of baseline word and picture categorizing and the semantic effects in both tasks show that the differences between tasks and conditions are consistently reflected in distributional shifting (μ) but not consistently in distributional skewing (τ). The results from the analyses of distribution parameters show that the effects are not restricted to the tail of the distributions, which converges with the results of the Vincentile analyses.

In summary, semantic effects were obtained for both word and picture targets even though word categorizing was much slower than picture categorizing. The impact of distractors in both tasks disagrees with the observations of Smith and Magee (1980), and it replicates the findings of W. R. Glaser and Düngelhoff (1984). However, whereas Glaser and Düngelhoff obtained no semantic effect in picture categorizing with a simultaneous presentation of picture and word, such an effect was obtained in Experiment 1. This result replicates Costa et al. (2003). Importantly, the present data replicate the semantic effect in word categorizing observed by Glaser and Düngelhoff and the semantic effect in picture categorizing observed by Costa et al. using a single group of participants. Baseline picture categorizing was faster than baseline word categorizing through the entire latency range. The present findings on semantic effects and relative baseline speeds are in line with the predictions of Roelofs (1992, 2003), as I demonstrate later in the present article by the results of WEAVER++ simulations. Distributional analyses revealed that the semantic effects in picture and word categorizing were somewhat depended on the absolute speed of responding. However, if anything, the magnitude of the effects was smallest for the slowest responses, contrary to what goal neglect would predict (De Jong et al., 1999). The associative account of Cohen et al. (1990) and Miller and Cohen (2001) is

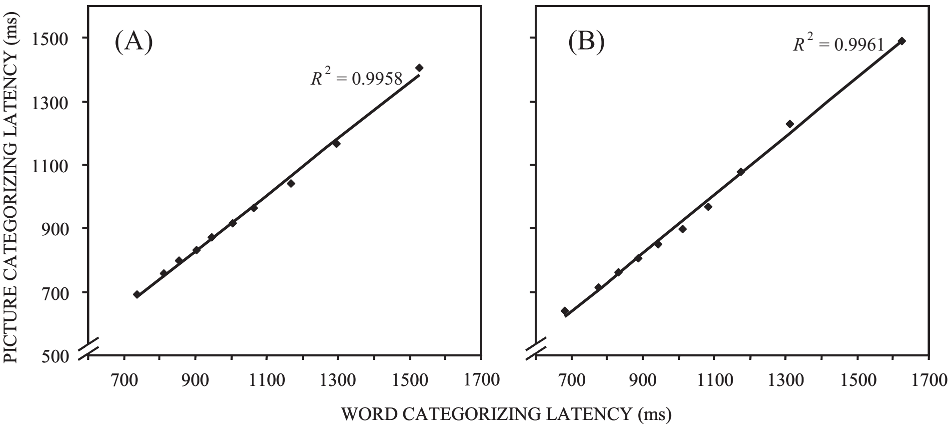

Figure 6. Quantile–quantile plots showing the relationship between word and picture categorizing latencies based on a grouping of the latencies by participants (Panel A) and by target responses (Panel B).

challenged by the observation of semantic effects for both word and picture targets.

Experiment 2

In Experiment 2, the participants' task was to verbally categorize the printed words while ignoring the pictured items. Participants did not have the task of categorizing the pictures, different from Experiment 1. The latencies of word categorizing were measured. The target words and context pictures were semantically related (e.g., the word HAMMER printed on a picture of a saw) or unrelated (e.g., HAMMER printed on a picture of a dog), or else they denoted the identical object (e.g., HAMMER printed on a pictured hammer). In the baseline condition, the word appeared in an empty rectangle. The SOA between word and picture was manipulated. The picture was presented 150 ms before, simultaneously with, or 150 ms after word onset. Presenting the picture 150 ms after word onset should compensate for the difference in speed between word and picture categorizing (which was 88 ms for the present stimuli, according to Experiment 1). An SOA of 150 ms was chosen rather than a larger one so as not to miss the semantic effect if it were present and still have an SOA larger than the difference in baseline latency between picture and word categorizing. Both preexposure and postexposure SOAs were included in the experiment so as not to bias the experiment toward postexposure SOAs. The identical condition was included to fully replicate the conditions of W. R. Glaser and Düngelhoff (1984). If semantic effects were obtained with picture postexposure, this would challenge the associative account and support the symbolic account. To examine the role of goal neglect, latency distributions were examined for each SOA.

Method

Participants. The experiment was carried out with a group of 12 paid (€4 [$6 U.S.]) participants, who were students at Radboud University Nijmegen. All participants were young adults and were native speakers of Dutch. None of them had participated in Experiment 1.

Materials and design. The materials were the same as in Experiment 1 (see the Appendix). The first independent variable was context. Each printed target word was combined with a picture from the same semantic category (the semantic condition), with a picture from another semantic category (the unrelated condition), with an identical picture (the identical condition), or an empty rectangle (the baseline condition). A participant received 32 stimuli in each of the four distractor conditions, yielding 128 different stimuli in total. The order of presenting the stimuli across trials was random, except that repetitions of pictures and words on successive trials were not permitted. The second independent variable was SOA. The picture was presented 150 ms before, simultaneously with, or 150 ms after word onset. Trials were blocked by SOA. The order of SOA blocks was counterbalanced across participants.

Procedure, apparatus, and analysis. The apparatus was the same as in Experiment 1. On each trial, the word and picture were presented with the appropriate SOA. The word was presented for 1,500 ms, and the response latency was measured from word onset. The participants were told that the response had to be initiated before the picture–word pair disappeared from the screen (i.e., within 1.5 s after picture–word presentation onset). Before the start of the next trial, there was a blank interval of 1.0 s. Thus, the total duration of a trial was 2.5 s. The production latencies and errors were submitted to analyses of variance with the crossed variables context and SOA. Both variables were tested within participants. To obtain the cumulative latency distributions, the rank-ordered latencies for each participant were again divided into deciles, and mean latencies were computed for each decile, separately for the responses in the semantic and unrelated contexts at each SOA. Also, item-based cumulative latency distributions were computed. Ex-Gaussian analyses were performed as in Experiment 1.

Results and Discussion

Figure 7 gives the mean word categorizing latencies and the mean error percentages for context and SOA. There were again virtually no technical errors; most errors involved wrong responses. Therefore, the types of errors were not separated. Figure 7

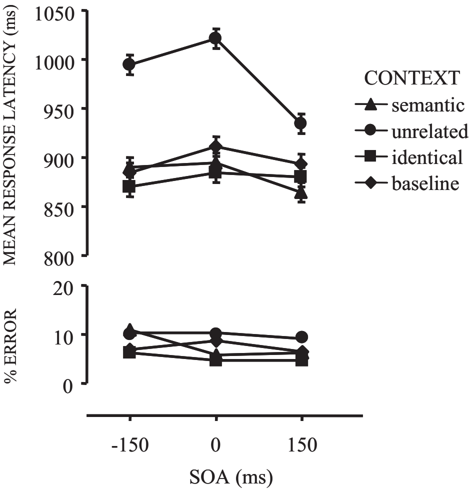

Figure 7. Mean response latencies and error percentages per context and SOA in Experiment 2. The error bars indicate the standard error of the mean. SOA = stimulus onset asynchrony.

shows that semantic effects were obtained at each SOA. The semantic effect was smaller at the postexposure SOA than at the other SOAs.

The statistical analysis of the categorizing latencies yielded main effects of context, $F_1(3, 33) = 38.64$, $p < .001$, $F_2(3, 93) = 66.70$, $p < .001$, and SOA in the by-item analysis, $F_1(2, 22) = 2.79$, $p = .08$, $F_2(2, 62) = 15.66$, $p < .001$. Context and SOA interacted, $F_1(6, 66) = 5.29$, $p < .001$, $F_2(6, 186) = 5.10$, $p < .001$. There was a main effect of semantic relatedness, $F_1(1, 11) = 66.87$, $p < .001$, $F_2(1, 31) = 88.54$, $p < .001$, which varied with SOA, $F_1(2, 22) = 4.47$, $p = .024$, $F_2(2, 62) = 4.65$, $p = .013$. Importantly, the semantic effect was present at all three SOAs: for SOA = −150 ms, $t_1(11) = 4.90$, $p < .001$, $t_2(31) = 6.46$, $p < .001$; for SOA = 0 ms, $t_1(11) = 8.35$, $p < .001$, $t_2(31) = 8.26$, $p < .001$; and for SOA = 150 ms, $t_1(11) = 5.50$, $p < .001$, $t_2(31) = 5.66$, $p < .001$. There was no difference in effect between semantic and identical contexts, $F_1(1, 11) < 1$, $p = .44$, $F_2(1, 31) < 1$, $p = .10$, and also no difference depending on the SOA, $F_1(2, 22) = 2.21$, $p = .13$, $F_2(2, 62) = 1.97$, $p = .15$. Finally, there was no difference in baseline categorizing speed between SOA blocks, $F_1(2, 22) = 1.15$, $p = .34$, $F_2(2, 62) = 2.37$, $p = .10$.

Figure 7 shows that more errors were made in the unrelated condition than in the other conditions, which did not differ much. The error percentages were similar for the three SOAs. The statistical analysis of the errors yielded a main effect of context, $F_1(3, 33) = 3.82$, $p < .02$, $F_2(3, 93) = 5.82$, $p < .001$, but not of SOA, $F_1(2, 22) = 1.11$, $p = .35$, $F_2(2, 62) = 2.31$, $p < .11$. Context and SOA did not interact, $F_1(6, 66) = 1.52$, $p = .19$, $F_2(6, 186) = 1.79$, $p = .10$. Given that most errors were made in the unrelated condition, which also had the slowest responses, there is no evidence for a speed–accuracy tradeoff.

Figure 8 shows the participant-based latency distributions for word categorizing in the semantic and unrelated picture contexts at each SOA. The item-based latency distributions were equivalent. At SOA = −150 ms, word categorizing was faster in the semantic than in the unrelated picture contexts through the entire latency range. Statistical analysis yielded a main effect of context, $F_1(1, 11) = 23.69$, $p < .001$, $F_2(1, 7) = 109.66$, $p < .001$, which did not depend on decile, $F_1(9, 99) = 1.11$, $p = .37$, $F_2(9, 63) < 1$, $p = .80$. Word categorizing was also faster in the semantic than in the unrelated picture contexts through the entire latency range at SOA = 0 ms. Statistical analysis yielded a main effect of context, $F_1(1, 11) = 72.26$, $p < .001$, $F_2(1, 7) = 55.83$, $p < .001$, which did not depend on decile, $F_1(9, 99) = 1.30$, $p = .25$, $F_2(9, 63) = 2.05$, $p = .05$. Finally, at SOA = 150 ms, word categorizing was faster in the semantic than in the unrelated picture contexts through almost the entire latency range. Statistical analysis yielded a main effect of context, $F_1(1, 11) = 33.72$, $p < .001$, $F_2(1, 7) = 43.88$, $p < .001$, which depended on decile, $F_1(9, 99) = 3.71$, $p < .001$, $F_2(9, 63) = 6.33$, $p < .001$. There appeared to be no semantic effect for the first and 10th deciles, $t_1(11) = 1.63$, $p = .13$, $t_2(7) < 1$, $p = .52$, and $t_1(11) = 1.17$, $p = .27$, $t_2(7) = 2.49$, $p = .04$,

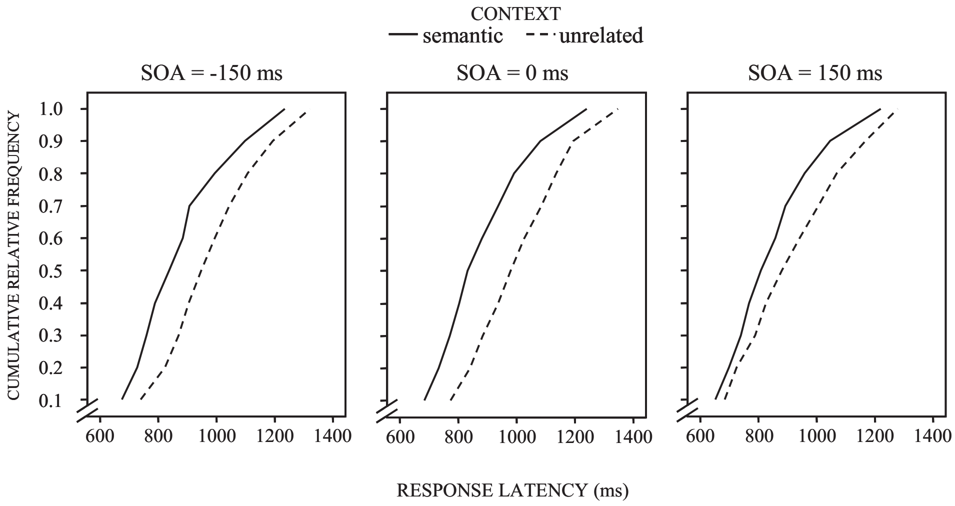

Figure 8. Vincentized cumulative distribution curves for the word categorizing latencies in the semantic and unrelated picture contexts per SOA. SOA = stimulus onset asynchrony.

respectively. However, the semantic effect was present in the second and ninth deciles, $t_1(11) = 2.57$, $p = .03$, $t_2(7) = 4.26$, $p = .004$, and $t_1(11) = 3.98$, $p = .002$, $t_2(7) = 3.44$, $p = .01$, respectively, and the deciles in-between. Again, the magnitude of the semantic effect did not increase overproportionately with response latency.

The ex-Gaussian analyses revealed that at SOA = −150 ms, the semantic effect was reflected in μ (746 ms and 876 ms for the semantic and unrelated distractors, respectively), $F(1, 11) = 25.81$, $p < .001$; but not in σ (78 ms vs. 100 ms), $F(1, 11) = 1.20$, $p = .30$; and also not in τ (147 ms vs. 117 ms), $F(1, 11) = 2.95$, $p = .11$. Thus, relative to the unrelated distractors, the semantic distractors shifted the latency distribution leftward without affecting the standard deviation and tail. At SOA = 0 ms, the semantic effect was reflected in μ (755 ms vs. 927 ms), $F(1, 11) = 53.96$, $p < .001$; in σ (75 ms vs. 126 ms), $F(1, 11) = 9.00$, $p = .012$; and in τ (141 ms vs. 94 ms), $F(1, 11) = 6.62$, $p = .03$. Thus, relative to the unrelated distractors, the semantic distractors shifted the latency distribution leftward, decreased the standard deviation, and increased the tail. Finally, at SOA = 150 ms, the semantic effect was reflected in μ (722 ms vs. 834 ms), $F(1, 11) = 23.82$, $p < .001$; in σ (78 ms vs. 134 ms), $F(1, 11) = 11.48$, $p = .006$; but not reliably in τ (145 ms vs. 99 ms), $F(1, 11) = 3.19$, $p = .10$. Thus, relative to the unrelated distractors, the semantic distractors shifted the latency distribution leftward, decreased the standard deviation, but had no effect on the tail. To conclude, the ex-Gaussian analyses of response latencies at the three SOAs show that the differences between the semantic and unrelated conditions are consistently reflected in distributional shifting (μ) but not consistently in distributional skewing (τ), showing that the effects are not restricted to the tail of the distributions. The results from the analyses of distribution parameters converge with the results of the Vincentile analyses.

In summary, semantic effects were obtained in word categorizing even when postexposure of the pictures compensated for the difference in categorization speed between pictures and words. Distributional analyses revealed that the semantic effects occurred through almost the entire latency range. These findings are in line with the predictions of Roelofs (1992, 2003). Although the distributional analyses revealed that the semantic effects were somewhat dependent on the absolute speed of responding, if anything, the magnitude of the effects was smallest for the fastest and slowest responses, contrary to what goal neglect would predict (De Jong et al., 1999). The associative account of Cohen et al. (1990) and Miller and Cohen (2001) is challenged by the observation of semantic effects with postexposure of the pictures.

Influence of Target Switching

The results of Experiments 1 and 2 reveal that distractors may yield semantic effects even when they make memory information available slower than the targets. Moreover, in both experiments, the semantic effects were largely independent of the absolute response latencies. However, there is also a difference between experiments. Whereas responding in the semantic condition was slower than baseline in Experiment 1, $F_1(1, 15) = 25.22$, $p < .001$, $F_2(1, 31) = 38.75$, $p < .001$, no such difference was obtained in Experiment 2, $F_1(1, 11) = 3.23$, $p = .10$, $F_2(1, 31) = 2.50$, $p = .12$. If anything, responding was faster rather than slower in the semantic condition compared with baseline in Experiment 2. The interaction of context (semantic vs. baseline) and experiment was significant, $F_1(1, 26) = 8.02$, $p < .009$, $F_2(1, 62) = 24.38$, $p < .001$.

It may be that this difference between experiments is due to the fact that participants had to perform both picture and word categorizing in Experiment 1 but only word categorizing in Experiment 2. Task switching may yield a kind of proactive interference (Allport, Styles, & Hsieh, 1994; Waszak, Hommel, & Allport, 2003). For example, if a participant categorizes the word on one trial and categorizes the picture on the next, the task set of the previous trial (word categorizing) may still be active or may be reactivated on the current trial (picture categorizing) and hamper responding. The phenomenon of proactive interference in memory retrieval was discovered by Müller and Schumann (1894) and is readily explained by Müller's (1913) associative account. "If *a* is already connected with *b,* then it is difficult to connect it with *k, b* gets in the way" (Müller & Schumann, 1894, quoted by Stroop, 1935, p. 644). Research has shown that responding is slower when participants switch between tasks than when the task is repeated (e.g., Allport et al., 1994; Jersild, 1927, reviewed by Woodworth, 1938). Rogers and Monsell (1995) observed that in task-switching experiments, responding in congruent conditions (like the semantic condition in Experiment 1) may be slower than baseline, as observed in Experiment 1. Task-switch costs are typically obtained for stimuli that allow both tasks in the experiment, like the semantically related and unrelated pairs in the present experiments. This raises the question to what extent the semantic effects in Experiment 1 are due to task switching.

As concerns responding to the semantically related and unrelated pairs in Experiment 1, statistical analysis showed that response latencies were slightly longer on switch than on repeat trials (1,082 ms vs. 1,063 ms), $t_1(15) = 2.54$, $p < .01$, $t_2(31) = 1.84$, $p < .04$. However, trial type (repeat vs. switch) did not interact with task, $F_1(1, 15) < 1$, $p = .96$, $F_2(1, 31) < 1$, $p = .59$. Also, the magnitude of the semantic effect did not depend on trial type, $F_1(1, 15) < 1$, $p = .93$, $F_2(1, 31) < 1$, $p = .79$. Thus, there was a task-switch cost, but it did not affect the magnitude of the semantic effect. Given that the semantic effects occurred both on switch and repeat trials, it is unlikely that the effects arose because of task switching.

Experiment 3

The aims of Experiment 3 were (a) to test the prediction of Roelofs (1992, 2003) that pictures should influence word categorizing but should not influence word reading regardless of response latency, (b) to obtain better estimates of the right tail of the response time distributions by not limiting the trial duration, and (c) to examine whether Watt's regularity holds for word reading and word categorizing even when the tasks are performed by different groups of participants. To these ends, one group of participants had to verbally categorize the printed words while ignoring the pictures, and another group had to vocally read the printed words while ignoring the pictures. The latencies of word reading and categorizing were measured. The target words and context pictures were semantically related (e.g., the word HAMMER printed on a picture of a saw) or unrelated (e.g., HAMMER printed on a picture of a dog), or else they denoted the identical object (e.g., HAMMER printed on a pictured hammer). In the baseline condition, the word appeared in an empty rectangle. If word categorizing requires access to the conceptual and lemma levels and word reading does not (see Figure 3), pictures should influence word categorizing but should not influence word reading, regardless of response latency. This claim was tested by comparing the distractor effects in word reading and word categorizing and by examining latency distributions. If the response time distribution is shifted leftward while the tail is increased or if the distribution is shifted rightward while the tail is decreased, it may be that mean latencies do not differ between conditions even though an effect is present (e.g., Heathcote et al., 1991; Spieler et al., 1996). This possibility was assessed by examining the shape of the distributions for word reading rather than just the condition means.

In Experiments 1 and 2, the semantic effect was absent for the slowest responses in some of the distribution curves. The right tail of empirical response time densities is usually fairly long (Luce, 1986). It is therefore possible that the absence of semantic effects reflects a ceiling effect because of the fixed limited duration of the trials in the experiments. In Experiment 3, the duration of a trial was therefore determined by response onset to prevent a ceiling effect and to obtain a better estimate of the magnitude of the semantic effect for the slowest responses. The distributions were again used to test for Watt's regularity.

To summarize, the aims of Experiment 3 were to test whether pictures influenced word categorizing but not word reading regardless of response latency, to obtain a better estimate of the right tail of the response time distribution, and to test for Watt's regularity.

Method

Participants. The experiment was carried out with 24 paid (€4 [$6 U.S.]) participants, who were students at Radboud University Nijmegen. All participants were young adults and were native speakers of Dutch. None of them had participated in Experiment 1 or 2.

Materials and design. The materials were the same as in Experiment 2 (see the Appendix). The first independent variable was context. Each printed target word was combined with a picture from the same semantic category (the semantic condition), with a picture from another semantic category (the unrelated condition), with an identical picture (the identical condition), or an empty rectangle (the baseline condition). A participant received 32 stimuli in each of the four distractor conditions, yielding 128 different stimuli in total. The order of presenting the stimuli across trials was random, except that repetitions of pictures and words on successive trials were not permitted. The second independent variable was response type (reading, categorizing), which was tested between participants. Half the participants read the words aloud, and the other half categorized the words.

Procedure, apparatus, and analysis. The apparatus was the same as in Experiments 1 and 2. On each trial, the word and picture were presented simultaneously. The response latency was measured from word onset. There was now no fixed trial duration. Instead, the triggering of the voice key completed a trial. The participants were told that a picture–word pair remained on the screen till articulation onset. Between the triggering of the voice key and the start of the next trial, there was a blank interval of 1.0 s. The production latencies and errors were submitted to analyses of variance with the crossed variables context and response type. Context was tested within participants and response type between participants. To obtain the cumulative latency distributions, the rank-ordered latencies for each participant were again divided into deciles and mean latencies were computed for each decile, separately for the responses in the semantic and unrelated contexts for each response type. Also, item-based cumulative latency distributions were computed. Ex-Gaussian analyses were conducted as in the previous experiments.

Results and Discussion

Figure 9 gives the mean word reading and categorizing latencies and the mean error percentages for context and response type. There were again virtually no technical errors; most errors involved wrong responses. Therefore, the types of errors were not separated. Figure 9 shows that a semantic effect was obtained in word categorizing, but there was no such effect at all in word reading. In word reading, there were no differences among context conditions.

The statistical analysis of the response latencies yielded main effects of context, $F_1(3, 66) = 16.99$, $p < .001$, $F_2(3, 186) = 10.96$, $p < .001$, and response type, $F_1(1, 22) = 121.86$, $p < .001$, $F_2(1, 62) = 426.08$, $p < .001$. Context and response type interacted, $F_1(3, 66) = 15.38$, $p < .001$, $F_2(3, 186) = 9.85$, $p < .001$. There was a main effect of semantic relatedness, $F_1(1, 22) = 31.91$, $p < .001$, $F_2(1, 62) = 19.75$, $p < .001$, which varied with response type, $F_1(1, 22) = 30.22$, $p < .001$, $F_2(1, 62) = 18.21$, $p < .001$. The semantic effect was present in word categorizing, $t_1(11) = 5.90$, $p < .001$, $t_2(31) = 4.50$, $p < .001$, but not in word reading, $t_1(11) = 0.22$, $p = .82$, $t_2(31) = 0.31$, $p = .76$. In word

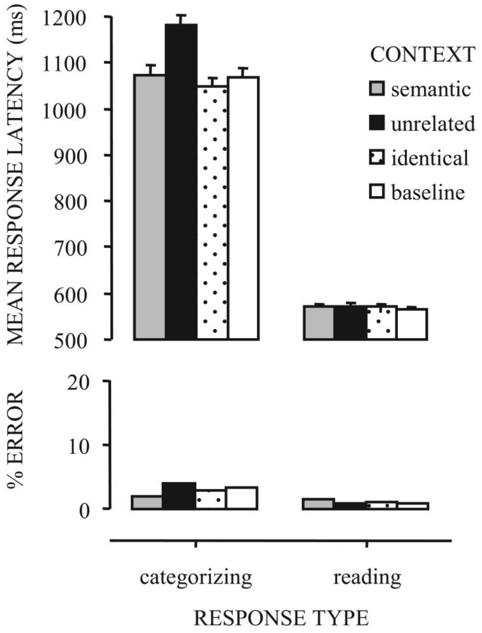

Figure 9. Mean response latencies and error percentages per context and response type in Experiment 3. The error bars indicate the standard error of the mean.

categorizing, responding in the unrelated condition was slower than baseline, $t_1(11) = 5.58$, $p < .001$, $t_2(31) = 4.44$, $p < .001$, whereas responding in the semantic and identical conditions did not differ, $t_1(11) = 1.90$, $p = .08$, $t_2(31) = 1.04$, $p = .30$. Also, the latencies of responding in the semantic condition did not differ from baseline, $t_1(11) = 0.66$, $p = .52$, $t_2(31) = 0.30$, $p = .77$. In word reading, there was no effect of context at all, $F_1(3, 33) < 1$, $p = .81$, $F_2(3, 93) < 1$, $p = .80$.

Figure 9 shows that slightly more errors were made in word categorizing than in word reading. The statistical analysis of the errors yielded no effect of context, $F_1(3, 66) < 1$, $p = .82$, $F_2(3, 186) < 1$, $p = .88$, but there was an effect of response type in the by-item analysis, $F_1(1, 22) = 2.20$, $p = .15$, $F_2(1, 62) = 20.21$, $p < .001$. Context and response type did not interact, $F_1(3, 66) = 1.87$, $p = .14$, $F_2(3, 186) = 1.41$, $p = .24$. Given that most errors were made in word categorizing, which was also slowest, there is no evidence for a speed–accuracy tradeoff. Whereas the error rates in Experiment 3 were below 5%, they were occasionally over 10% in the previous experiments. In Experiments 1–2, the trial duration was limited, whereas this was not the case in Experiment 3. Perhaps the participants in Experiment 3 felt less pressed than those in Experiments 1–2, reducing the number of errors.

Figure 10 shows the participant-based latency distributions for word reading and word categorizing in the baseline condition (left panel) and in the semantic and unrelated picture contexts (right panel). The item-based latency distributions were equivalent. Word reading was faster than word categorizing across the entire latency range (there is no need for a statistical test here). Word categorizing was faster in the semantic than in the unrelated picture contexts across the entire latency range. Statistical analysis yielded a main effect of context, $F_1(1, 11) = 33.34$, $p < .001$, $F_2(1, 7) = 20.88$, $p < .001$, which did not depend on decile, $F_1(9, 99) < 1$, $p = .77$, $F_2(9, 63) = 1.19$, $p = .32$. Thus, the semantic effect was constant across the latency distribution in Experiment 3. This suggests that the absence of the effect for the slowest responses for some of the curves in Experiments 1 and 2 may be a ceiling effect due to the fixed limited duration of the trials. Still, the constancy of effects across the latency distribution excludes goal neglect as the cause of the semantic effect because goal neglect would predict an overproportionate increase of the effect with latency. Word reading latencies were the same in the semantic and the unrelated picture contexts across the entire latency range. Statistical analysis yielded no effect of context, $F_1(1, 11) < 1$, $p = .40$, $F_2(1, 7) = 1.85$, $p = .22$, and no dependence on decile, $F_1(9, 99) < 1$, $p = .97$, $F_2(9, 63) < 1$, $p = .98$.

The left panel of Figure 10 shows that the latency range was very small for word reading compared with word categorizing, which confirms that word reading is a very different process from word categorizing (e.g., Roelofs, 2003). Still, Watt's regularity appeared to hold for the two tasks, as illustrated by the quantile–quantile plots in Figure 11. The figure shows the relationship between word reading and word categorizing latencies based on a grouping of the latencies by participants (see Figure 11A) and by semantic domain (see Figure 11B). The item-based grouping was based on semantic domain for both tasks because the actual responses differed between tasks (i.e., basic-level names for reading and category names for categorizing). Word reading and word categorizing latencies were linearly related. Thus, again, the influence of the task was independent of the relative speed of responding itself. Watt's regularity appears to hold even when the tasks are performed by different groups of participants.

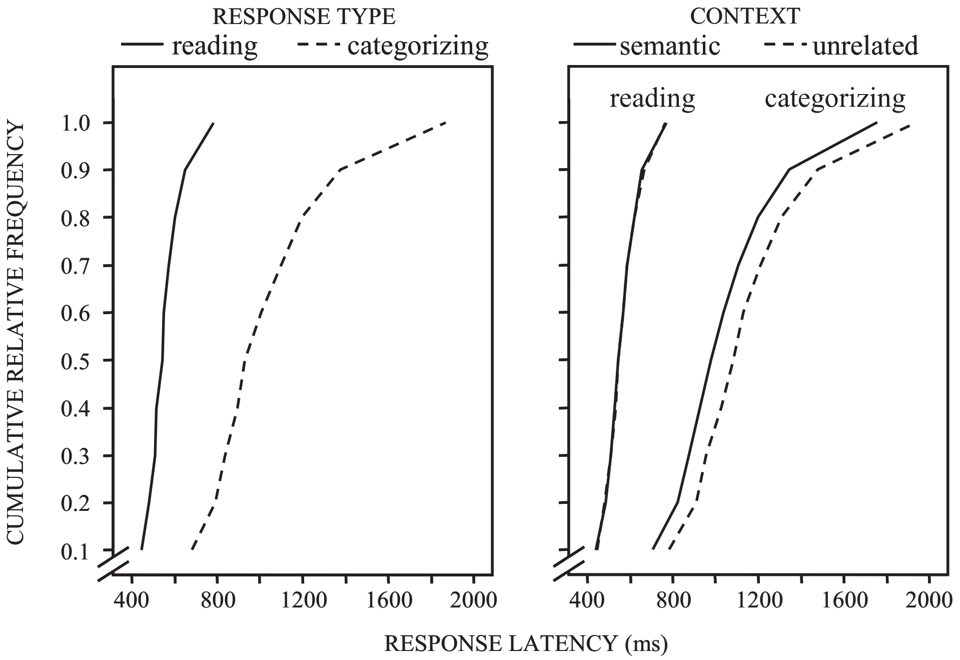

Figure 10. Vincentized cumulative distribution curves for word categorizing and word reading in the baseline condition (left panel) and in the semantic and unrelated picture contexts (right panel).

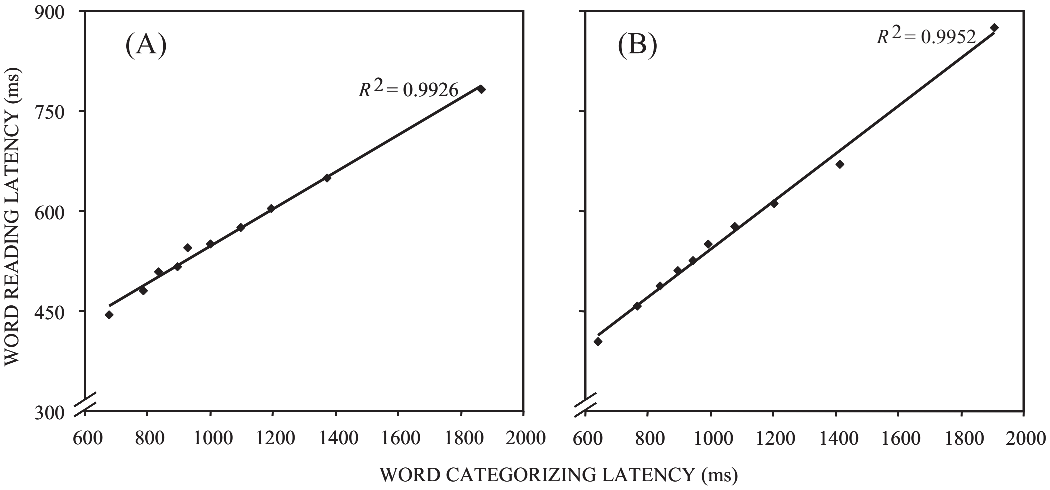

Figure 11. Quantile–quantile plots showing the relationship between word reading and word categorizing latencies based on a grouping of the latencies by participants (Panel A) and by semantic domain (Panel B).

The ex-Gaussian analyses revealed that baseline word reading and word categorizing differed in μ (490 ms vs. 766 ms, respectively), $F(1, 11) = 48.16$, $p < .001$; in σ (36 ms vs. 102 ms), $F(1, 11) = 16.79$, $p = .002$; and also in τ (74 ms vs. 296 ms), $F(1, 11) = 84.99$, $p < .001$. Thus, relative to word reading, word categorizing shifted the latency distribution rightward, increased the standard deviation, and increased the tail. In word reading, there was no semantic effect in any of the distribution parameters, μ (494 ms and 499 ms for the semantic and unrelated distractors, respectively), $F(1, 11) < 1$, $p = .45$; σ (43 ms vs. 47 ms), $F(1, 11) < 1$, $p = .50$; and τ (76 ms vs. 72 ms), $F(1, 11) < 1$, $p = .62$. Finally, in word categorizing, the semantic effect was reflected in μ (820 ms vs. 909 ms), $F(1, 11) = 5.98$, $p = .033$; but not in σ (110 ms vs. 118 ms), $F(1, 11) < 1$, $p = .84$; and also not in τ (255 ms vs. 274 ms), $F(1, 11) < 1$, $p = .62$. Thus, relative to the unrelated distractors, the semantic distractors shifted the latency distribution leftward, without affecting the standard deviation and tail. To conclude, the ex-Gaussian analyses of baseline word reading and word categorizing and the semantic effects in both tasks show that the effects observed in the means are consistently reflected in distributional shifting (μ) but not consistently in distributional skewing (τ). Again, the results from the analyses of distribution parameters show that the effects are not restricted to the tail of the distributions, which converges with the results of the Vincentile analyses.

In summary, the experiment showed that Watt's regularity holds for verbal tasks that are very different, namely, word reading and word categorizing. Semantic effects were obtained in word categorizing but not in word reading. Distributional analyses revealed that the semantic effect occurred through the entire latency range in word categorizing, whereas the semantic effect was absent across the entire latency range in word reading. These findings are in line with the predictions of Roelofs (1992, 2003). The observation that the semantic effect in word categorizing was constant across the latency distribution suggests that the absence of the effect for the slowest responses for some of the curves in Experiments 1 and 2 may have been a ceiling effect.

Categorizing Versus Naming

In word and picture categorizing in Experiments 1–3, participants were faster in categorizing the semantically related than the unrelated picture–word pairs. In contrast, in naming the picture of picture–word pairs, the naming response is typically slower for the semantically related than the unrelated picture–word pairs (e.g., W. R. Glaser & Düngelhoff, 1984; Roelofs, 2006; Smith & Magee, 1980). Like word and picture categorizing, picture naming requires access to the conceptual level (see Figure 2). In terms of the levels of processing illustrated in Figure 3, picture naming involves object-form perception (e.g., the form of a pictured hammer), conceptualizing (i.e., retrieving HAMMER[*X*]), lemma retrieval (i.e., retrieving *hammer*), and word-form encoding. This raises the question why the semantic effect in categorizing (descriptively: semantic facilitation) is the reverse of the effect in picture naming (descriptively: semantic interference).

According to Selz (1927), a stimulus may evoke multiple word responses, leading to competition in response selection. However, this competition "moves within the narrow range of task-relevant reproductions" (Selz, 1927, translated by J. M. Mandler & Mandler, 1964, p. 231). Category terms compete with category terms but not with coordinate terms. Elsewhere (Roelofs, 1992, 2003, 2006), I argued that this explains why distractor word SAW interferes in naming a pictured hammer (say "hammer") compared with distractor word DOG, whereas SAW facilitates saying "tool" compared with DOG. In picture naming, the distractors SAW and DOG denote task-relevant responses (e.g., they are at the appropriate level of abstraction). Consequently, *saw* and *dog* compete with the target *hammer*. Because *saw* and *hammer* are connected at the conceptual level (see Figure 2), the pictured hammer inadvertently activates the word response *saw,* whereas the response *dog* is not activated. The response activation makes *saw* a stronger competitor than *dog* in selecting *hammer,* which yields semantic interference. In contrast, in a categorizing task, only superordinate terms are task-relevant responses. Consequently, in picture categorizing, the distractor DOG activates the category concept ANI-

MAL(X) and the corresponding category name *animal,* which leads to competition with the target *tool* in response selection. In contrast, the word SAW does not activate a competing category name. As a consequence, slower responses are obtained in the unrelated than in the semantically related condition: semantic facilitation. Elsewhere (Roelofs, 1992), I demonstrated the utility of this account by quantitative WEAVER++ fits of the data on the time course of semantic effects in word and picture categorizing obtained by W. R. Glaser and Düngelhoff (1984). It should be acknowledged, however, that under what conditions a response set gets established, if at all, is a much debated issue (e.g., Caramazza & Costa, 2000; Costa et al., 2003; Roelofs, 2001, 2003, 2006).

WEAVER++ Simulations

In this section, the results of new computer simulations, which show the utility of the symbolic account embodied by the WEAVER++ model (Levelt et al., 1999; Roelofs, 1992, 2003) in accounting for the key findings on categorizing from Experiments 1–3, are reported. WEAVER++ is described extensively in other articles (Levelt et al., 1999; Roelofs, 1992, 2003, 2006, 2007), and I refer the reader to these publications for details on the model and the simulation protocol.

Two semantic domains were used in the present simulations, namely, tools (e.g., tool, hammer, saw) and animals (e.g., animal, dog, cat). The structure of a semantic domain is illustrated in Figure 2. In the simulations, a picture activated the corresponding concept node and a printed word activated the corresponding word node. Whereas the target provided input activation to the network until the selection of a word node as response, the distractor provided activation input for a limited period of time, the *distractor duration.* The model proceeded through time in discrete time steps. On each time step, activation spread through the network following a linear activation function with a decay factor. Activation of nodes in the network triggered the application of production rules, like Rule 1 given earlier, depending on the task goal (word categorizing, picture categorizing). A word node was selected as response when its level of activation exceeded that of the other nodes by some critical amount, the *selection threshold.* In the simulations of word and picture categorizing reported by Roelofs (1992), the distractor duration and the selection threshold parameters were given task-dependent values to optimize the fit of the model to the time-course data of W. R. Glaser and Düngelhoff (1984). In the present simulations, the values of these two parameters were identical for picture and word categorizing. The distractor duration was set to 75 ms and the selection threshold to 5.0 for both picture categorizing and word categorizing. All other parameter values were fixed and identical to those of Roelofs (1992, 2003).

The simulations revealed that there was a difference in baseline latency between tasks in the model. Word categorizing was 127 ms slower than picture categorizing. W. R. Glaser and Düngelhoff (1984) obtained a difference of 126 ms. The difference in baseline observed in Experiment 1 was 88 ms, so the model overestimated the observed difference. However, despite the sizeable difference in baseline latency between categorizing pictures and words in the model, there were semantic effects in both tasks, as illustrated in Figure 12. The model overestimated the semantic effect in word categorizing in Experiment 1, although the magnitude of the effect in the model (107 ms) was close to the magnitude of the effect observed in Experiment 2 (126 ms at SOA = 0 ms) and Experiment 3 (108 ms). The semantic effect was larger in word categorizing than in picture categorizing, in agreement with the empirical observations in Experiment 1.

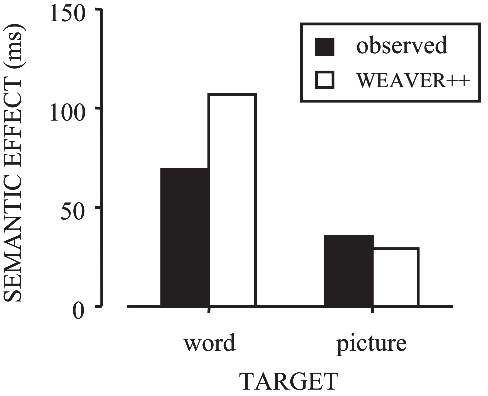

Figure 12. Semantic effects in picture and word categorizing: Observed data from Experiment 1 and WEAVER++ simulations.

Figure 13 shows the simulated cumulative distribution curves for picture categorizing in the semantic and unrelated word contexts. Although picture categorizing was faster than word categorizing in the model, the semantic effect of words on picture categorizing was present across the whole latency range, as empirically observed (see the middle panel of Figure 5). The latency range is smaller in the model than in the real data because the simulations concerned word retrieval only. Word and picture perception and word-form encoding (see Figure 3) were not included. These processes presumably have their own inherent stochastic variability, which adds to the variability of the eventual response and the range of the response latency distribution. The baseline latencies for picture and word categorizing were linearly related in the simulations (R^2 = .9931), as empirically observed.

To conclude, the computer simulations show that the characteristics of performance of WEAVER++ on word and picture categorizing agree with the key empirical observations on word and picture categorizing in the present article. The simulation results demonstrate the utility of the symbolic account.

General Discussion

Watt's (1906) dissertation work at Würzburg University directed the attention of psychologists to the problem of how humans are capable of retrieving information from associative memory in a flexible manner to meet task demands. How is it possible that a task-relevant weaker association is realized instead of one that normally is stronger? In the present article, three chronometric experiments intended to further investigate this problem have been reported. One of the aims of the experiments was to examine the relative merits of the prevailing associative account of Cohen et al. (1990) and Miller and Cohen (2001) and the symbolic account of Roelofs (1992, 2003) of goal-driven memory retrieval. The accounts were tested in experiments on picture–word categorizing. The experiments examined the influence of strong irrelevant associations on the use of weaker target associations. The associative

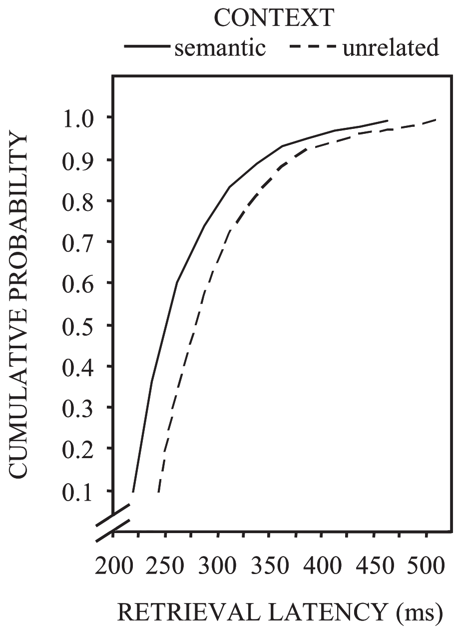

Figure 13. Cumulative distribution curves of word retrieval latencies for picture categorizing in the semantic and unrelated word contexts in WEAVER++ simulations.

account (Cohen et al., 1990; Miller & Cohen, 2001) holds that goals selectively reinforce particular associations, and it predicts that strong irrelevant associations influence weaker target ones, but not vice versa. The symbolic account (Roelofs, 2003) holds that goals enable particular retrieval rules, which would predict no such dependence on association strength.

In Experiment 1, participants categorized the word or picture of picture–word pairs, which could be semantically related or unrelated, or they categorized single pictures and words in a baseline condition. The aim of the experiment was to assess the baseline speed of responding to the pictures and words and to test whether semantic effects are obtained in picture categorizing when word categorizing is slower than picture categorizing. If semantic effects were obtained in both tasks even when word categorizing was slower than picture categorizing, this would challenge the associative account and support the symbolic account. In the experiment, semantic effects were obtained for both word and picture targets even though words were categorized slower than pictures. Distributional analyses revealed that the difference in response latency between pictures and words was present through the entire latency range. The semantic effects occurred for most of the latency deciles for both pictures and words, excluding goal neglect as the cause of the semantic effects. I have argued that although proactive interference effects may be present in a task-switching situation, the semantic effects observed in Experiment 1 cannot be due to task switching per se. This is because the semantic effects occurred on both switch and repeat trials.

In Experiment 2, participants had to categorize the word of the picture–word pairs, while the SOA between picture and word was manipulated. The picture was presented 150 ms before, simultaneously with, or 150 ms after word onset. Presenting the picture 150 ms after word onset compensated for the difference in speed between word and picture categorizing, as assessed by Experiment 1. If semantic effects were obtained with picture postexposure, this would challenge the associative account and support the symbolic account. In the experiment, semantic effects were obtained in word categorizing even when postexposure of the pictures compensated for the difference in categorization speed between pictures and words. Distributional analyses revealed that the semantic effects occurred for most of the latency deciles at each SOA, and the effect did not increase disproportionately with latency, excluding goal neglect as the cause of the semantic effects.

In Experiment 3, participants had to categorize the words (one group of participants) or read aloud the words (another group) and ignore the pictures. Semantic effects were obtained in word categorizing but not in word reading. Distributional analyses revealed that the semantic effect occurred through the entire latency range in word categorizing, whereas the semantic effect was absent across the entire latency range in word reading. These findings are in line with the claim of Roelofs (1992, 2003) that word categorizing requires access to the conceptual level, whereas word reading does not, as illustrated in Figure 3. Moreover, the observation that the semantic effect in word categorizing was constant across the latency distribution suggests that the absence of the effect for the slowest responses for some of the curves in Experiments 1 and 2 may have been a ceiling effect. Still, the constancy excludes goal neglect as the cause of the semantic effects.

Watt (1906) grouped together verbal responses of the same intrinsic speed and found that a variation of task had a similar effect across latency groups. This regularity also appeared to hold for the tasks used in the present experiments. In Experiment 1, word categorizing and picture categorizing latencies were almost perfectly linearly related, and the same held for the word reading and word categorizing latencies in Experiment 3. The linear relation between the latencies in the tasks demonstrates that the influence of the task was independent of the relative speed of responding itself. The linear relationship between task latencies appears to hold even when the tasks are performed by different groups of participants, as was the case in Experiment 3.

Associative Versus Symbolic Accounts

I have argued that the results of the present experiments are most consistent with the symbolic as compared with the associative account. The effect of relative associative strength in the model of Cohen et al. (1990) is most evident from their simulations of interference as a function of training. MacLeod and Dunbar (1988) had participants learn to call each of four different arbitrary shapes by a different color name (“green,” “pink,” “orange,” “blue”). Participants received 20 practice sessions on shape naming, on 20 separate days. After 1, 5, and 20 sessions of practicing shape naming (respectively, 72, 504, and 2,520 trials per stimulus), performance was tested for interference from incongruent colors on shape naming (e.g., saying “blue” to the polygon named “blue” in green color) and from incongruent shapes on color naming (e.g., saying “green” to the color green of the polygon named “blue”) relative to a baseline condition. On Day 1, colors interfered with shape naming, but shapes did not interfere with color naming. On Day 5, there was interference in both tasks. Finally, on Day 20, shapes interfered with color naming, but colors no longer interfered with shape naming. Thus, training on task associations may reverse the patterns of interference between tasks. In Roelofs (2003), I presented successful WEAVER++ simulations of these findings.

In the simulations of Cohen et al. (1990), color naming was initially some 110 ms faster than shape naming (my estimate from their graphs). Relative to a baseline condition, incongruent colors interfered with shape naming (about 70 ms interference), but incongruent shapes did not interfere with color naming (<10 ms). At the intermediate point in training, the latencies of color and shape naming were the same, and there was no interference in either of the tasks (unlike the real data). After 2,520 training epochs, shape naming was some 50 ms faster than color naming. Now, incongruent shapes interfered with color naming (about 40 ms interference), but incongruent colors no longer interfered with shape naming (<10 ms).

In Experiment 1, picture categorizing was about 88 ms faster than word categorizing, corresponding to the initial difference of 110 ms between color and shape naming in the simulations of Cohen et al. (1990). However, in Experiment 1, semantic effects were obtained in both tasks (34 ms and 69 ms for picture and word categorizing, respectively). Relative to the baseline condition, the effects of incongruent unrelated distractors (e.g., the word HAMMER superimposed onto a pictured dog) were over 100 ms in both tasks. These empirical results contradict the model of Cohen et al. In their model, incongruent colors interfere with shape naming (about 70 ms), but incongruent shapes do not interfere with color naming (<10 ms). In Experiment 2, pictures yielded a semantic effect (of 70 ms) on word categorizing at SOA = 150 ms, even though the SOA compensated for the difference in speed between word and picture processing. In contrast, at the intermediate point in training in the simulations of Cohen et al., the latencies of color and shape naming were the same, but there was no interference in either of the tasks (<10 ms). The empirical findings of Experiment 2 disagree with these simulation results.

To conclude, the simulations reported by Cohen et al. (1990) reveal that when one task is faster than the other, the stimuli for the faster task affect the performance of the slower task but not the other way around. Moreover, if tasks are of comparable speed, virtually no interference is obtained in either task. These characteristics of the model do not correspond to the present data. There seems no immediately obvious way for accommodating the present findings within the framework of the model. The same conclusion holds for similar models (Melara & Algom, 2003; Phaf et al., 1990).

One may argue that the semantic effect of distractor words on picture categorizing in Experiment 1 reflects an underlying probability mixture of trials with different relative processing speeds for pictures and words. Although word categorizing is, on average, slower than picture categorizing, distractor words may have been categorized faster than target pictures on some trials because of random variability in latencies. Perhaps the semantic effect from words in picture categorizing arose from the subset of trials at which word processing was actually faster than picture processing. If this were true, the semantic effect of word distractors should be largest for the relatively long picture categorizing latencies, when word categorizing had the highest chance to be faster than picture categorizing. However, this does not correspond to what was observed in Experiment 1. If anything, the semantic effect of word distractors was smaller rather than larger for the long picture categorizing latencies. Similarly, the semantic effect of picture distractors on word categorizing should be smallest for the relatively short word categorizing latencies in Experiments 1–3. For these short word categorizing latencies, picture categorizing had the highest chance to be slower than word categorizing. Again, this prediction differs from what the real data show. To conclude, there is no evidence that the presence of semantic effects reflects an underlying probability mixture of trials with different relative processing speeds for picture and word.

Computer simulations demonstrated that a model embodying the symbolic account, WEAVER++, is able to explain the presence of semantic effects in both picture and word categorizing even when words are categorized slower than pictures. Selz (1927) distinguished associative (Müller, 1913), specific response (Selz, 1913), and superimposition accounts of the control of memory retrieval. According to a superimposition account, human memory employs both an associative activation process and specific responses, as implemented in WEAVER++. Selz (1927) rejected this possibility by arguing that if a system of specific responses exists, associative activations are superfluous. Elsewhere (Roelofs, 2003), I argued, however, that associative spreading of activation may be important because it solves the problem of how to prevent all production rules from always having to test for their conditions even when they are entirely irrelevant. Spreading activation may restrict production rule testing to part of memory, for example, to the semantic domain of tools only. One may speculate that spreading activation is metabolically cheap and easy to implement by the human brain, whereas symbolic computations (like those involved in production rule application) are relatively difficult and metabolically expensive (Anderson, 1983). Deacon (1997) argued that symbolic computation is so difficult to implement by primate brains that it has emerged only once in evolution, namely, in human brains only. He maintained that the frontal lobes of the human brain play a critical role in symbolic computation.

The constrained association task used by Watt (1906) and Selz (1913, 1922) has seen a revival with the advent of modern neuroimaging techniques. In seminal functional imaging experiments, Petersen, Fox, Posner, Mintun, and Raichle (1988) asked participants to produce a use for a word, for example, say “hit” to HAMMER. Several later imaging studies have followed this lead (e.g., Thompson-Schill, D’Esposito, Aguirre, & Farah, 1997). In these studies, it is typically observed that generating a use for a word increases blood flow in frontal and temporal areas of the human brain relative to reading the words aloud (e.g., say “hammer” to HAMMER). The increased metabolic activity in frontal areas disappears after repeated generation of the same use to a word (Petersen, van Mier, Fiez, & Raichle, 1998). The evidence suggests that frontal areas play a role in selection among competing response alternatives (Thompson-Schill et al., 1997), the control of memory retrieval, or both (Badre, Poldrack, Paré-Blagoev, Insler, & Wagner, 2005). Such functions are in agreement with the evidence for the implication of frontal areas in attentional control, as suggested by Wundt (1904). For example, frontal areas are involved in Stroop task performance (e.g., Miller & Cohen, 2001). Evidence for the critical involvement of frontal areas in the kind of categorization tasks used in the present article comes from impaired performance. Semantic retrieval problems due to lesions of temporal areas of the human brain typically preserve the ability to generate category terms. Humphreys and Forde (2005) reported evidence on a patient with combined frontal-temporal damage who had, instead, a specific impairment of generating category terms. According to Humphreys and Forde, the unusual impairment resulted because categorizing requires the attentional control provided by the frontal lobes. Similarly, frontal lesions impair Stroop task performance (Stuss, Floden, Alexander, Levine, & Katz, 2001).

The symbolic (Roelofs, 1992, 2003) and associative models (Cohen et al., 1990; Miller & Cohen, 2001) discussed in the present article characterize goal-directed word retrieval at a psychological level of description. The models have not been designed to account for specific features of real neural networks, but they specify the type of processing that the brain is capable of performing. To be physically realized, both symbolic and local associative mechanisms must ultimately be implemented via distributed neurophysiological systems. A critical distinction between the associative and symbolic models concerns the role of association strength (see Roelofs, 2003, for a discussion of other distinctions). The findings from Experiments 1–3 suggest that whether irrelevant associations interfere in word retrieval is not strictly dependent on association strength and retrieval latency, which is most consistent with the symbolic account. This conclusion concerning the type of psychological processes regulating word retrieval (i.e., symbolic rather than associative) may help identify the actual brain mechanisms. As Watt (1906) stated, “a more exact definition of psychological factors and their sphere of operation can only be welcome to physiology” (p. 265). Likewise, psychological models may be constrained by physiological evidence, such as the finding that neural circuits in prefrontal cortex encode abstract rules (e.g., Bunge, 2004; Bunge, Kahn, Wallis, Miller, & Wagner, 2003; Wallis, Anderson, & Miller, 2001). Roelofs and Hagoort (2002) evaluated the symbolic and associative models using physiological evidence on Stroop task performance.

Relation to the Color–Word Stroop Task

The associative account of Cohen et al. (1990) and Miller and Cohen (2001) and the symbolic account of Roelofs (1992, 2003) have also previously been tested on chronometric findings from the color–word Stroop task (Cohen et al., 1990; Roelofs, 2003, 2005). Although the categorization problem and the problem posed by the Stroop task seem related, the empirical patterns of interference are different between picture–word categorizing and color–word naming. When the tasks of color naming and word reading are performed by different groups of participants, participants are usually much slower in naming the presentation color of an incongruent color word than in naming the color of a colored series of *X*s in the baseline condition, whereas there is no interference from incongruent colors relative to baseline in word reading (M. O. Glaser & Glaser, 1982; MacLeod, 1991). The absence of an effect in word reading is obtained even when color patches are preexposed at an SOA (e.g., −400 or −300 ms) that compensates for the difference in speed between responding to colors and words (i.e., word reading is typically some 100–200 ms faster than color naming). A similar asymmetry in effects has been obtained with the naming of picture–word stimuli (e.g., W. R. Glaser & Düngelhoff, 1984; Roelofs, 2006, 2007; and Experiment 3). Incongruent words interfere with picture naming (e.g., say “hammer” to a pictured hammer), but incongruent pictures do not influence word reading (e.g., say “hammer” to the word HAMMER; see Experiment 3). However, when the tasks of picture and word categorizing are performed by different groups of participants, semantic effects are obtained in both tasks (W. R. Glaser & Düngelhoff, 1984). Thus, the patterns of effects obtained with the color–word Stroop task and the picture–word categorizing task are different.

Elsewhere (Roelofs, 1992, 2003, 2005, 2006, 2007), I argued that the asymmetry in effects between Stroop-like picture and color naming on the one hand and word reading on the other arises because words are read via a shallow route from word-form perception to word-form encoding (see Figure 3), without even accessing the levels of word responses or concepts illustrated in Figure 2. This is different from picture categorizing and word categorizing, which both require access to the conceptual level. Consequently, pictures influence word categorizing and vice versa (e.g., W. R. Glaser & Düngelhoff, 1984, and Experiments 1–3), but pictures do not influence word naming (Experiment 3) even though words affect picture naming (e.g., W. R. Glaser & Düngelhoff, 1984; Roelofs, 2003, 2005, 2006). Unlike picture/color naming and reading, the interference effects in word and picture categorizing should be obtained regardless of whether the tasks are performed by a single group of participants or by different groups, as empirically observed (different groups: W. R. Glaser & Düngelhoff, 1984; same group: Experiment 1).

Summary and Conclusions

In Experiment 1, semantic effects were obtained for both word and picture categorizing even though words were categorized slower than pictures. Moreover, in Experiment 2, semantic effects were obtained in word categorizing even when postexposure of the pictures compensated for the difference in categorization speed between pictures and words. In Experiment 3, pictures yielded a semantic effect in word categorizing but not in word reading. In all three experiments, the semantic manipulations resulted in distributional shifting, except for word reading, which exhibited no semantic effect. Distributional analyses revealed that, when present, the semantic effects occurred for most of the latency deciles in all experiments. Moreover, the effects did not increase disproportionately with latency, excluding goal neglect as the cause of the semantic effects. The experiments confirmed Watt’s regularity. The influence of the task was independent of the relative speed of responding. This held not only for word and picture categorizing but also for word categorizing and reading. The results have been interpreted as most consistent with the symbolic account advanced by Roelofs (1992, 2003) as compared with the associative account advanced by Cohen et al. (1990) and Miller and Cohen (2001). WEAVER++ simulations showed that the symbolic account agrees with the key findings on word and picture categorizing.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.

Allport, A., & Wylie, G. (1999). Task-switching: Positive and negative priming of task-set. In G. W. Humphreys, J. Duncan, & A. Treisman

(Eds.), *Attention, space, and action: Studies in cognitive neuroscience* (pp. 273–296). Oxford, England: Oxford University Press.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. R., & Bower, G. (1973). *Human associative memory.* Washington, DC: Winston.

Badre, D., Poldrack, R. A., Paré-Blagoev, E., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47,* 907–918.

Boring, E. G. (1950). *A history of experimental psychology.* New York: Appleton-Century-Crofts.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Brown, S., & Heathcote, A. (2003). QMLE: Fast, robust, and efficient estimation of distribution functions based on quantiles. *Behavior Research Methods, Instruments, & Computers, 35,* 485–492.

Bub, D. N., Masson, M. E. J., & Lalonde, C. E. (2006). Cognitive control in children: Stroop interference and suppression of word reading. *Psychological Science, 17,* 351–357.

Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 564–579.

Bunge, S. A., Kahn, I., Wallis, J. D., Miller, E. K., & Wagner, A. D. (2003). Neural circuits subserving the retrieval and maintenance of abstract rules. *Journal of Neurophysiology, 90,* 3419–3428.

Caramazza, A., & Costa, A. (2000). The semantic interference effect in the picture–word interference paradigm: Does the response set matter? *Cognition, 75,* B51–B64.

Cattell, J. M., & Bryant, S. (1889). Mental association investigated by experiment. *Mind, 14,* 230–250.

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., Dunbar, K., & McClelland, J. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332–361.

Costa, A., Mahon, B., Savova, V., & Caramazza, A. (2003). Level of categorization effect: A novel effect in the picture–word interference paradigm. *Language and Cognitive Processes, 18,* 205–233.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain.* New York: Norton.

De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica, 101,* 379–394.

De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus–response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 731–750.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 721–733). Cambridge, MA: MIT Press.

Gibson, J. J. (1941). A critical review of the concept of set in contemporary experimental psychology. *Psychological Bulletin, 38,* 781–817.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 875–894.

Glaser, W. R. (1992). Picture naming. *Cognition, 42,* 61–105.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640–654.

Heathcote, A., Popiel, J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109,* 340–347.

Hohle, R. H. (1965). Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology, 69,* 382–386.

Humphrey, G. (1951). *Thinking: An introduction to its experimental psychology.* London: Methuen.

Humphreys, G. W., & Forde, E. M. E. (2005). Naming a giraffe but not an animal: Basic-level but not superordinate naming in a patient with impaired semantics. *Cognitive Neuropsychology, 22,* 539–558.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology, 14*(Whole No. 89).

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47–70.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–38.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 126–135.

Mandler, G. (2007). *A history of modern experimental psychology: From James and Wundt to cognitive science.* Cambridge, MA: MIT Press.

Mandler, J. M., & Mandler, G. (Eds.). (1964). *Thinking: From association to gestalt.* Wiley: New York.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110,* 422–471.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202.

Müller, G. E. (1913). Zur Analyse der Gedächtnistätigkeit und des Vorstellungsverlaufs [On the analysis of memory and the flow of images]. *Zeitschrift für Psychologie, 8,* 475–489.

Müller, G. E., & Schumann, F. (1894). Experimentelle Beiträge zur Untersuchung des Gedächtnisses [Experimental contributions on the investigation of memory]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane, 6,* 81–190, 257–339.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988, February 18). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331,* 585–589.

Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences, USA, 95,* 853–860.

Phaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology, 22,* 273–341.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446–461.

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz & B. Hommel (Eds.), *Attention and Performance XIX: Common mechanisms in perception and action* (pp. 494–519). Oxford, England: Oxford University Press.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.

Roelofs, A. (2001). Set size and repetition matter: Comment on Caramazza and Costa (2000). *Cognition, 80,* 283–290.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110,* 88–125.

Roelofs, A. (2005). The visual–auditory color–word Stroop asymmetry and its time course. *Memory & Cognition, 33,* 1325–1336.

Roelofs, A. (2006). Context effects of pictures and words in naming objects, reading words, and generating simple phrases. *Quarterly Journal of Experimental Psychology, 59,* 1764–1784.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language, 57,* 232–251.

Roelofs, A., & Hagoort, P. (2002). Control of language use: Cognitive modeling of the hemodynamics of Stroop task performance. *Cognitive Brain Research, 15,* 85–97.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.

Selz, O. (1913). *Über die Gesetze des geordneten Denkverlaufs* [On the laws of ordered thinking]. Stuttgart, Germany: Spemann.

Selz, O. (1922). *Zur Psychologie des produktiven Denkens und des Irrtums* [On the psychology of productive thinking and error]. Bonn, Germany: Friedrich Cohen.

Selz, O. (1927). Die Umgestaltung der Grundanschauungen vom intellectuellen Geschehen [The revision of the fundamental conceptions of the intellectual processes]. *Kantstudien, 32,* 273–289.

Simon, H. A. (1981). Otto Selz and information-processing psychology. In N. H. Frijda & A. de Groot (Eds.), *Otto Selz: His contribution to psychology* (pp. 147–163). The Hague, the Netherlands: Mouton.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General, 109,* 373–392.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 461–479.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001). Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. *Neuropsychologia, 39,* 771–786.

Thomas, E. A. C., & Ross, B. H. (1980). On appropriate procedures for combining probability distributions within the same family. *Journal of Mathematical Psychology, 21,* 136–152.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, USA, 94,* 14792–14797.

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review, 7,* 424–465.

Wallis, J. D., Anderson, K. C., & Miller, E. (2001, June 21). Single neurons in prefrontal cortex encode abstract rules. *Nature, 411,* 953–956.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive Psychology, 46,* 361–413.

Watt, H. J. (1906). Experimental contribution to a theory of thinking. *Journal of Anatomy and Physiology, 40,* 257–266.

West, R., & Alain, C. (2000). Evidence for the transient nature of a neural system supporting goal-directed action. *Cerebral Cortex, 8,* 748–752.

Wilk, M. B., & Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. *Biometrika, 55,* 1–17.

Woodworth, R. S. (1938). *Experimental psychology.* New York: Holt.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology* (rev. ed.). New York: Holt.

Wundt, W. (1904). *Principles of physiological psychology.* Leipzig, Germany: Engelmann.

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 274–296.

Zhang, J., & Kornblum, S. (1997). Distributional analysis and De Jong, Liang, and Lauber's (1994). dual-process model of the Simon effect. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1543–1551.

Appendix

Materials of Experiments 1–3

Target		Context		
Category name	Word/picture	Semantic	Unrelated	Identical
dier (animal)	zwaan (swan)	schildpad	rok	zwaan
	schildpad (tortoise)	zwaan	beker	schildpad
	konijn (rabbit)	hert	paleis	konijn
	hert (deer)	konijn	bureau	hert
kleding (clothing)	trui (sweater)	rok	dolk	trui
	rok (skirt)	trui	zwaan	rok
	hemd (shirt)	vest	oor	hemd
	vest (vest)	hemd	kasteel	vest
vervoer (transportation)	fiets (bike)	trein	kast	fiets
	trein (train)	fiets	arm	trein
	schip (ship)	vliegtuig	been	schip
	vliegtuig (plane)	schip	glas	vliegtuig
gebouw (building)	molen (windmill)	fabriek	kom	molen
	fabriek (factory)	molen	neus	fabriek
	kasteel (castle)	paleis	vest	kasteel
	paleis (palace)	kasteel	konijn	paleis
wapen (weapon)	dolk (dagger)	speer	trui	dolk
	speer (spear)	dolk	tafel	speer
	kanon (cannon)	pistool	bord	kanon
	pistool (pistol)	kanon	bed	pistool
servies (service)	beker (cup)	kom	schildpad	beker
	kom (bowl)	beker	molen	kom
	glas (glass)	bord	vliegtuig	glas
	bord (plate)	glas	kanon	bord
meubel (furniture)	tafel (table)	kast	speer	tafel
	kast (cupboard)	tafel	fiets	kast
	bed (bed)	bureau	pistool	bed
	bureau (desk)	bed	hert	bureau
lichaamsdeel (body part)	arm (arm)	neus	trein	arm
	neus (nose)	arm	fabriek	neus
	been (leg)	oor	schip	been
	oor (ear)	been	hemd	oor

Note. English translations of the Dutch targets are given in parentheses.

Received April 27, 2007
Revision received October 12, 2007
Accepted October 12, 2007 ■