

NOW YOU SEE IT, NOW YOU DON'T: ON TURNING SEMANTIC INTERFERENCE INTO FACILITATION IN A STROOP-LIKE TASK

Matthew Finkbeiner and Alfonso Caramazza

(Department of Psychology, Harvard University, Cambridge, MA, USA)

ABSTRACT

We use a masked priming procedure to test two accounts of the picture-word interference (PWI) effect: the lexical selection by competition account (Levelt et al., 1999; Roelofs, 1992) and the response selection account (Lupker, 1979; Miozzo and Caramazza, 2003). In the visible (standard) condition, we replicated the often-observed semantic interference effect. In the masked condition, we observed semantic *facilitation*. We take the polarity shift as a function of masking to mean that the semantic interference and semantic facilitation in the PWI task should be attributed to two qualitatively different processes. We argue that this conclusion follows naturally from the response selection account, but only with great difficulty from the lexical selection by competition account.

Key words: lexical access, picture word interference, masked priming, response selection

The process of translating one's intention to communicate into articulated speech is known as *lexical access* and although there are various points of agreement among researchers on how lexical access processes proceed, there are also disagreements. It is generally well accepted that when trying to name a picture of an object, say a cat, the intended semantic representation (CAT) and closely related ones (such as DOG, PURR, TAIL, FUR) all become active (Caramazza, 1997; Dell, 1986; Levelt et al., 1999). It is also well accepted that activated conceptual representations send activation down onto the lexical layer, thereby activating several semantically related lexical representations (but for an alternative proposal see Bloem and La Heij, 2003 and Bloem et al., 2004). Because of this assumption, a lexical selection mechanism is required to "decide" which of the activated lexical representations should be chosen for further processing. Although there is still widespread disagreement over the precise way in which this mechanism works, a prominent assumption is that lexical selection is a competitive process (Levelt et al., 1999; Starreveld and La Heij, 1996; but see Caramazza and Hillis, 1990; Dell and O'Seaghdha, 1992). According to this "selection by competition" assumption, the ease with which a target lexical node is selected depends not only on its own activation level, but on the activation level of competing lexical nodes as well, such that there is an inverse relationship between the time required to select a target lexical node and the relative activation levels of competing lexical nodes. The higher the relative activation level of a competitor's lexical node, the longer it takes to select the target lexical node.

The picture-word naming task is, perhaps, the task that has been used most often to investigate

lexical selection processes. In the picture-word naming task, which is a version of the Stroop task, participants are asked to name pictures as quickly as they can while ignoring a distractor word that is superimposed on the picture. A well-established finding with this task is that participants take longer to name a target picture when a distractor word is present (*vs.* no distractor word) and longer yet when the to-be-ignored distractor word is semantically (categorically) related to the picture (e.g., Caramazza and Costa, 2000; Glaser and Glaser, 1989; Lupker, 1979). Recently, this finding has been interpreted in terms of competitive lexical selection processes (Roelofs, 1992; Starreveld and La Heij, 1996; Levelt et al., 1999). Essentially, a distractor word that is related to the picture will be more highly activated than an unrelated word because of semantic priming from the picture, and, because of its increased activation, will compete more fiercely for selection. Here we would like to suggest an alternative account of the picture-word interference (PWI) effect, one that does not require invoking competition at the stage of lexical selection.

At the heart of this alternative account, which we will refer to as the response selection account, are the following assumptions. First, in a naming task such as the picture-word naming task, we assume that individuals, upon detecting the presence of the word distractor, cannot help but to formulate a (covert) verbal response to the distractor word and that this response is available for production before the picture-naming response. This is no different from the long-held assumption that subjects are unable to avoid reading the word distractors in the Stroop task. The second assumption follows from the limitation imposed by the architecture of the speech production system, which is that there is only one output channel and

that only one response may be produced at a time over this channel. In the case of Stroop-like tasks where an individual cannot help but to generate two responses for a single stimulus (see first assumption), we assume that individuals must “decide” between responses and remove (‘block’) the inappropriate response from the output buffer so that the appropriate response may be produced over the output channel. There are two critical aspects to this response selection process that we make explicit here. First, we assume that response selection processes operate over phonologically well-formed responses in an output buffer, not lexical nodes at a more abstract level of representation. Second, following Lupker (1979), we assume that the speed with which the decision-level response selection process may be completed is modulated by whatever relevance the non-target response may have *vis-à-vis* the task at hand; in the case of the PWI task, this is producing a name for the picture.

Support for the response selection proposal comes from a series of recent studies in our laboratory. For example, Miozzo and Caramazza (2003) found that low frequency distractor words interfere more than high frequency words in the PWI task. This finding, which is very robust, constitutes a challenge to the lexical selection by competition account (Levelt et al., 1999; Roelofs, 1992; Starreveld and La Heij, 1996). This is because high frequency distractors, by virtue of having relatively high levels of activation, should produce more (or, at the very least, *not less*) interference than low frequency distractors. Another challenge for the selection by competition account comes from recent findings which reveal that although co-ordinate picture-word pairs (e.g., car – truck) lead to interference, “has-a” picture-word pairs (car – bumper) facilitate responses (Costa et al., 2005). The selection by competition account predicts that all types of semantically related words should cause interference. Both findings fall naturally out of the response selection proposal. In the case of the distractor frequency effect, the sooner a distractor can be blocked from production, the sooner the picture name can be selected (c.f., Miozzo and Caramazza, 2003). Because high frequency words are able to trigger a response and engage the response-blocking mechanism sooner than low frequency words, it takes less time to block high frequency words. In the case of interference for co-ordinate picture-word pairs and facilitation for “has-a” picture-word pairs, we suggest that the response selection process is more difficult in the case of co-ordinate distractors because the decision mechanism is unable to reject the response triggered by the distractor word on the basis of categorical information alone. In the case of “has-a” distractors, though, these words (e.g., bumper) have little relevance *vis-à-vis* the task of naming whole

objects (e.g., car). Hence they should not compete. Why, though, did “has-a” distractors *facilitate* picture naming? Here we suggest that in the absence of any response relevance, the cost attributed to response selection processes is minimal and, crucially, is less than the benefit conferred by semantic priming.

Put simply, we assume that two independent factors play a role in the PWI effect. The first has to do with how quickly the non-target response is made available to the output buffer. Factors such as lexical frequency may play a role here. The second has to do with how quickly the “decision” to reject the non-target response may be executed. Factors which conspire to determine a distractor word’s “response relevance” (such as category membership) play a role here.

Following from this response selection account of the PWI effect is the interesting possibility that masked distractor words may produce a shift in the polarity of the semantic interference effect in this task. That is, insofar as the masking procedure is capable of preventing a phonologically well-formed response from becoming available for production, there is no reason to think that the processes needed to block the production of the distractor word would be engaged¹. And, given the finding that masked stimuli have been found to produce semantic priming (Carr et al., 1982; Finkbeiner et al., 2004; Grainger and Frenck-Mestre, 1998), it is reasonable to consider the possibility that masked words may produce a *facilitatory* effect of semantic relatedness in picture naming. Alternatively, if, as the lexical selection by competition account holds, the distractor word competes for selection at the level of lexical representations, then we may expect that the masking procedure should simply modulate the magnitude of the interference effect: depending upon the strength of the masking procedure, the semantic interference effect size could range from zero to the same effect size obtained with unmasked stimuli. Importantly, the lexical selection by competition account does not predict a shift in the polarity of the semantic interference effect as a function of masking, whereas the response selection account allows for this possibility. The purpose of the present article is to investigate these two possibilities further.

To anticipate our findings briefly, we found that semantically related prime words presented under masked conditions significantly facilitated picture-naming processes; the same word stimuli presented in unmasked conditions significantly interfered with picture-naming processes, even when the prime’s display duration was held constant between the masked and unmasked conditions (Experiment 2).

¹Note that this does not imply that we are presuming that the masking procedure is effective in preventing *any* activation at the phonological level; rather, we are simply speculating that the masking procedure may be effective in preventing the formulation of a phonologically well-formed response capable of engaging response selection processes.

EXPERIMENT 1

The purpose of this first experiment was to investigate the possibility that masked distractor words in the picture-word naming task may produce a qualitatively different effect from the standard interference effect found with non-masked distractors. Assuming that the masking procedure is effective in preventing the formulation of a phonologically well-formed response, masked distractor words should not engage response selection processes and, hence, we should not observe interference in this condition. In contrast, we reason that masked distractors may produce facilitation by virtue of engaging word-recognition processes without engaging response selection processes. We expect to replicate the often-observed PWI effect when the distractor word is clearly visible because, according to our account, these stimuli are capable of engaging response selection processes.

Participants

Eighteen undergraduate students at Harvard University were recruited and paid for their participation. All participants were native speakers of English and were between 18 and 25 years old.

Materials and Design

Forty-six pictures were selected from Snodgrass and Vanderwart (1980), from Art Explosion, or were constructed in our lab. Each picture was paired with two different prime words: One was a semantically related prime (e.g., *lion* – *TIGER*) and one was a semantically unrelated control prime (*boot* – *TIGER*). An effort was made to select experimental primes that were semantically very close to the targets. Each control prime was selected to match the experimental prime it was paired with on several factors, including length (letters), frequency (CELEX), familiarity, concreteness, and imageability. Two lists were constructed such that no item appeared more than once on each list. The target pictures were the same on both lists, but the primes differed. Target pictures that were preceded by semantically related primes on List A were preceded by control primes on List B, and vice versa. In this way, materials were counterbalanced across the prime factor. In addition to the 46 experimental items, there were 10 practice items that were presented prior to the critical items. Participants were assigned to and tested on either List A or B, depending upon their order of participation.

Procedure

Participants named each picture twice: once in the masked condition and once in the visible condition. Items were blocked according to presentation type, with the masked condition always

preceding the visible condition. The reason for having the masked condition precede the visible condition comes from recent findings which suggest that stimulus-response mappings learned with visible stimuli may be invoked automatically on subsequent trials, even when those stimuli are masked on the subsequent trials (Damian, 2001). There is little reason, on the other hand, to be concerned that words presented previously as masked primes would affect later performance as the masked priming effect is very short lived (Forster and Davis, 1984).

The experiment was controlled by a Pentium PC, using DMDX (Forster and Forster, 2003), which synchronizes the display with the monitor's raster. In the masked condition, each trial began with a forward mask (#####), which was presented for approximately 500 msec. The forward mask was followed immediately by the prime word. The prime word was presented in lowercase letters for a duration of 53 msec (4 refresh cycles at 75 kHz). The prime word was immediately followed by a backward mask, which was a different randomly generated consonant string on each trial (e.g., MZWXNMSCZP). The use of a consonant string as a backward mask was motivated by recent findings which reveal that this type of mask is more effective than hash marks or ampersands in eliminating phonological priming effects with 53 msec prime durations (Grainger et al., 2003). The backward mask and the target picture appeared simultaneously, with the backward mask superimposed on the picture. On any given trial, the forward mask, prime and backward mask appeared immediately one after the other in the same location; on each successive trial, this location randomly varied from 0 to 3 degrees of visual angle along both the X and Y axes from the center of the screen. The pictures, subtending approximately 8.5 degrees, appeared in the center of the screen on each trial for 2 seconds or until the participants' response triggered the voice key. After the voice key was triggered, two seconds elapsed before the onset of the next trial. Items were presented in a different fully randomized order for each participant with the only constraint that the experimental items followed the practice items.

In the visible condition, each trial began with a central fixation point (+) presented for approximately 500 msec. The fixation point was replaced immediately with the distractor word presented in uppercase letters for 53 msec, which was then immediately replaced with a simultaneous display of the target picture and the distractor word. Following the same procedure used in the masked condition, the position of the distractor word varied on a trial by trial basis.

After the experiment proper, participants were debriefed and shown the series of events comprising each trial (forward mask – prime – simultaneous onset of picture and backward mask). After being shown how items in the experiment

had been constructed, participants were asked if they had been able to notice if any words had been presented prior to the onset of the pictures. In the present experiment, one participant reported being able to see some letters of words before the onset of the backward mask and picture in the masked condition. This participant was replaced.

Results and Discussion

Three types of responses were scored as errors (a) production of names that differed from those designated by the experimenter; (b) verbal disfluencies (stuttering, utterance repairs, production of nonverbal sounds which triggered the voice key); and (c) recording failures. In the masked condition, 7.3% of the responses were coded as errors and eliminated from analyses. For each comparison of interest, two analyses of variance were performed, one treating subjects as a random effect (F1), the other treating items as a random effect (F2). The factors were Groups (subject groups in the subject analysis, item groups in the item analysis), and Prime type (semantically related vs. unrelated). The Groups factor was included to remove variance due to the counterbalancing procedure, and was a non-repeated factor in both analyses. The factor of Prime type was a repeated measures factor in both analyses. The mean naming latencies are presented in Table I. A significant effect of masked semantic priming was obtained, such that pictures were named faster when preceded by a masked presentation of a semantically related word [$F1(1, 16) = 5.82, p = .02$; $F2(1, 44) = 6.15, p = .01$]. Analysis of the error rates revealed no significant differences between the related and unrelated conditions (all p 's $> .1$).

In the visible condition, 6.3% of the responses were coded as errors. Once again, an analysis of these error rates revealed no significant differences between the related and unrelated conditions (all p 's $> .1$). Looking at Table I, it is clear that the masking procedure reversed the polarity of the semantic effect on picture naming. Namely, changing the presentation so that the word was clearly visible had the effect of turning the masked priming effects into reliable interference effects [$F1(1, 16) = 5.74, p = .03$; $F2(1, 44) = 5.89, p = .02$]. As far as we are aware, these results

represent the first time that the same word-picture pairs have produced both facilitation and interference effects at the same stimulus onset asynchrony (SOA). More importantly, these findings provide support for the hypothesis that it is the appreciation of the distractor word as a possible response that triggers the response selection processes responsible for the PWI effect.

One possible concern in interpreting these findings lies in the possibility that the observed dissociation has more to do with distractor duration than it has to do with the masking procedure. The masked prime was presented for 53 msec, while the visible distractor was presented alone for 53 msec and then superimposed on the target until the voice key was triggered (825 msec on average). Hence, the exposure durations of the masked prime and the visible distractor were very different and the concern is that this difference may have been the source of the dissociation. Experiment 2 addresses this issue by holding the SOA and the duration of the masked prime and the visible distractors constant.

EXPERIMENT 2

The purpose of Experiment 2 was twofold: one, to replicate the juxtaposed interference and facilitation effects reported in Experiment 1, and, two, to investigate the possible effect of distractor word duration on the polarity shift. Twenty participants were recruited from the same student population as in Experiment 1. To ensure that the pattern of effects reported in Experiment 1 generalize to other items, thirty-eight new pictures were selected or created and paired with distractor words according to the same criteria used in Experiment 1. The procedure and design were identical to those of Experiment 1 except for the interpolation of a brief period (13.3 msec) between the offset of the prime and onset of the target picture. This resulted in an SOA of -66 msec, and a prime duration of 53 msec. In the masked condition, the backward mask was presented during the refresh cycle preceding the onset of the target picture and was then superimposed on the picture. In the visible condition, no backward mask was used, which left the screen blank during the interpolated period and allowed the picture to be presented unobstructed by a mask or distractor word. The purpose of interpolating the brief period between the offset of the distractor and the onset of the target picture was to prevent the picture from acting as an immediate backward mask for the distractor word in the visible condition².

TABLE I
Naming latencies (msec), standard deviations (in parenthesis) and error rates by relatedness and presentation conditions in Experiment 1

	Prime type	Naming latency	Error %
Masked primes	Related	764 (84)	8.2
	Unrelated	796 (116)	6.3
	Priming effect	32	
Visible distractors	Related	841 (76)	6.9
	Unrelated	809 (96)	5.6
	Interference effect	-32	

²Although this procedure equates the duration of the prime in the masked and visible conditions, participants were aware of the primes in the visible condition but not in the masked condition. This difference is presumably due to competition between prime and backward mask in an early capacity-limited stage of pattern recognition (Di Lollo et al., 2000).

1 and 2. In the case of the distractor frequency effect, we appeal to the same hypothesis proposed by Miozzo and Caramazza (2003), which holds that the sooner a distractor can be blocked from production, the sooner the picture name can be selected. Because high frequency words are available for production sooner, they can be blocked earlier, allowing articulation of the picture-naming response to commence sooner. In the case of Bloem et al.'s manipulation of SOA, we propose that semantic facilitation was observed in the -400 msec SOA condition because the blocking processes responsible for the semantic interference effect had already run their course by the time the target was presented. In the case of the present study, we propose that semantic facilitation was observed because the masking procedure prevented the formulation of a phonologically well-formed response for the distractor word, which prevented the response selection processes from becoming engaged. We conclude that the response selection account provides for the most straightforward way to account for the polarity shift reported in the present study as well as the polarity shift reported by Bloem et al. (2004).

There are several questions raised by the response selection proposal that we have not addressed in detail here. First we should emphasize that we do not assume that response selection processes are engaged in normal speech production; rather, it is only in Stroop-like tasks, where each stimulus affords two possible responses, that a decision-level response-selection mechanism is necessary to invoke. Another possible question has to do with our claim that it is a phonologically realized response for the distractor stimulus that engages the response selection processes (which are responsible for the interference effect). If this is the case, how is it that participants do not simply say the first response available for production? That is, how does a decision mechanism operating over phonologically-realized responses "know" which response is correct and which is incorrect? Here we speculate that it is the participants' "conscious" detection of the distractor stimulus that leads to the formulation of a phonologically well-formed response and subsequently allows the participant (decision mechanism) to "know" the provenance of the response (word distractor, not picture) and, hence, the type of decision that must be made (reject or articulate). That is, we confer a special function upon the conscious detection of a distractor stimulus in Stroop-like tasks, such as the PWI paradigm.

Of course this position may be too strong. Perhaps the triggering of an alternative naming response, though strongly correlated with the phenomenological experience of having detected a distractor stimulus, is not caused by the conscious detection of the distractor. That is, there may be

two independent thresholds: one that triggers the to-be-blocked naming response and one that leads to the phenomenological experience of detecting a stimulus. Though this is logically possible and very difficult to refute, it lacks parsimony and it ignores the building consensus on the role of awareness. For example, Crick and Koch (1998) propose that the role of awareness is to make information available "... to the parts of the brain that contemplate and plan voluntary motor output" (such as a naming response; cited in Jack and Shallice, 2001). Similarly, Dehaene and Naccache (2001) argue that information becomes conscious when it is represented by long-distance "workspace" neurons, which, by virtue of their connectivity, make the information available to a variety of processes, including memorization, evaluation and intentional action (such as the formulation and subsequent blocking of a naming response). We suggest that the present findings may be taken to provide some initial support for these general hypotheses on the role of awareness by showing how awareness of a distractor word appears to "trigger a phonological response", or, stated more generally, serves to make specific information available to other processes for subsequent action.

Acknowledgments. The research reported here was supported in part by National Institutes of Health (NIH) grant DC 04542-04. We gratefully acknowledge Bradford Mahon, Robin Peterson and Kay Negishi for their help in developing materials and testing participants.

REFERENCES

- BLOEM I and LA HEIJ W. Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, 48: 468-488, 2003.
- BLOEM I, VAN DEN BOOGAARD S and LA HEIJ W. Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language*, 51: 307-323, 2004.
- CARAMAZZA A. How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14: 177-208, 1997.
- CARAMAZZA A and COSTA A. The semantic interference effect in the picture-word interference paradigm: Does the response set matter? *Cognition*, 75: B51-B64, 2000.
- CARAMAZZA A and HILLIS A. Where do semantic errors come from? *Cortex*, 26: 95-122, 1990.
- CARR T, MCCAULEY C, SPERBER R and PARMELEE C. Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, 8: 757-777, 1982.
- COSTA A, ALARIO F-X and CARAMAZZA A. On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulletin and Review*, 12: 125-131, 2005.
- CRICK F and KOCH C. Consciousness and neuroscience. *Cerebral Cortex*, 8: 97-107, 1998.
- DAMIAN MF. Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, 271: 154-165, 2001.
- DEHAENE S and NACCACHE L. Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, 79: 1-37, 2001.
- DELL GS. A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 3: 283-321, 1986.

- DELL GS and O'SEAGHDHA PG. Stages of lexical access in language production. *Cognition*, 42: 287-314, 1992.
- DI LOLLO V, ENNS JT and RENSINK RA. Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, 129: 481-507, 2000.
- FINKBEINER M, FORSTER KI, NICOL J and NAKAMURA K. The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, 51: 1-22, 2004.
- FORSTER KI and DAVIS C. Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 104: 680-698, 1984.
- FORSTER KI and FORSTER JC. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers*, 35: 116-124, 2003.
- GLASER WR and GLASER MO. Context effects on Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118: 13-42, 1989.
- GRAINGER J, DIEPENDAELE K, SPINELLI E, FERRAND L and FARIOLI F. Masked repetition and phonological priming within and across modalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29: 1256-1269, 2003.
- GRAINGER J and FRENCK-MESTRE C. Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes*, 13: 601-623, 1998.
- JACK A and SHALLICE T. Introspective physicalism as an approach to the science of consciousness. *Cognition*, 79: 161-196, 2001.
- LEVELT WJM, ROELOFS A and MEYER AS. A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22: 1-38, 1999.
- LUPKER SJ. The semantic nature of response competition in the picture-word interference task. *Memory and Cognition*, 7: 485-495, 1979.
- MIOZZO M and CARAMAZZA A. When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology: General*, 132: 228-252, 2003.
- ROELOFS A. A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42: 107-142, 1992.
- SNODGRASS JG and VANDERWART M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Experimental Psychology*, 6: 174-215, 1980.
- STARREVELD PA and LA HEIJ W. Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22: 896-918, 1996.

Matthew Finkbeiner, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA. e-mail: msf@wjh.harvard.edu

(Received 11 March 2005; accepted 30 May 2005)