Psychonomic Bulletin & Review
1999, 6 (2), 347-351

Notes and Comment

Only the Shadower knows: Comment on Hamburger and Slowiaczek (1996)

STEPHEN D. GOLDINGER
Arizona State University, Tempe, Arizona

The phonological priming effect may reflect basic processes in spoken word perception and has thus been a central topic of recent research. In this journal, Hamburger and Slowiaczek (1996) reported phonological priming data collected in a shadowing task. They replicated a prior study (Slowiaczek & Hamburger, 1992), but added new procedures to minimize bias. After observing inhibitory priming in a "low-expectancy" condition, they concluded that facilitatory priming reflects perceptual/response bias, but that inhibitory priming reflects automatic processes of lexical access. This commentary critiques Hamburger and Slowiaczek's method and presents new data that demonstrate persistent biases in primed shadowing. I suggest that such biases reflect natural, context-sensitive listening strategies.

The *phonological priming effect* (Slowiaczek, Nusbaum, & Pisoni, 1987) is an important empirical finding, potentially revealing the dynamics of spoken word perception. Such dynamics are often modeled as processes of activation and competition, with each process guided by physical stimulus dimensions (see, e.g., Marslen-Wilson, 1990; McClelland & Elman, 1986). Thus, phonological priming effects may link theories to data. However, phonological priming has not inspired great confidence—it is unstable across tasks and materials (Radeau, Morais, & Dewier, 1989; Radeau, Morais, & Segui, 1995; Slowiaczek & Pisoni, 1986), and it shows associated perceptual/response biases (Goldinger, Luce, Pisoni, & Marcario, 1992). Indeed, the brief history of phonological priming has primarily entailed a quest for "pure," reliable priming. The effect arises in degraded-word identification and (sometimes) lexical decision, but concomitant biases also arise (Goldinger et al., 1992). Presumably, if the veneer of bias could be removed, true priming effects would remain. Toward this end, researchers have advocated "shallow" response time (RT) tasks, such as primed shadowing (Radeau et al., 1989; Slowiaczek & Hamburger, 1992).

This article concerns a recent primed shadowing study by Hamburger and Slowiaczek (1996), who extended their own prior study (Slowiaczek & Hamburger, 1992). In the earlier study, primes and targets shared varying degrees of overlap, ranging from zero (control) to four (repetition) shared phonemes. Listeners heard clear primes and targets, shadowing the targets as quickly as possible. The RTs showed facilitatory priming for low-overlap pairs and inhibitory priming for high-overlap pairs, but no repetition effects. In their later study, Hamburger and Slowiaczek examined primed shadowing in a standard condition and in a condition that reduced the likelihood of perceptual/response biases. Goldinger et al. (1992) showed that biases produce facilitatory priming between words sharing initial phonemes (in lexical decision and perceptual identification). However, Goldinger et al. could not address the priming that occurs with greater phonological overlap, nor could they address shadowing data. Thus, Hamburger and Slowiaczek sought to examine the role of biases in primed shadowing.

The high-expectancy condition featured a long (500-msec) interstimulus interval (ISI) and a high (75%) proportion of related trials. The low-expectancy condition featured a short (50-msec) ISI and a low (21%) proportion of related trials. The high-expectancy condition replicated the earlier data (Slowiaczek & Hamburger, 1992), showing facilitation for pairs with a one- or two-phoneme overlap, and "interference" for pairs with a three-phoneme overlap. In the low-expectancy condition, the facilitation effects vanished, but a three-phoneme inhibition effect grew robust. As Hamburger and Slowiaczek (1996) wrote,

> Of greater interest in the present study, high-similarity interference was observed under both high- and low-expectancy conditions, indicating that this effect is not strategic. . . . Thus, this phonological priming effect most likely reflects the operations of the auditory word recognition system. (p. 524)

The Experimental Method: See No Evil?

Despite its intuitive appeal, I suggest that Hamburger and Slowiaczek's (1996) conclusion is unwarranted, due primarily to a faulty assumption. Their reasoning can easily be summarized: (1) For obvious reasons, the high-expectancy data should replicate the earlier results. (2) If high-expectancy priming is bias driven, it should vanish if expectancies are removed. (3) The three-phoneme interference effect does *not* vanish in the low-expectancy condition. Therefore, (4) that result is not due to bias. This logic critically depends on the removal of bias, which was never assessed. Without assessment, we cannot determine whether the low-expectancy data were truly bias free, or whether the manipulations were simply too weak to eliminate bias. The latter possibility seems plausible. For example, prime–target pairs with high overlap (three out of

Support was provided by NIDCD Grant R29-DC02629-03. Comments were kindly provided by Stephen Lupker, Tamiko Azuma, Paul Luce, and Mark Pitt. Correspondence should be addressed to S. D. Goldinger, Department of Psychology, Arizona State University, Box 871104, Tempe, AZ 85287-1104 (e-mail: goldinger@asu.edu).

four phonemes) should be quite salient, relative to pairs with less overlap. As reported later, virtually all participants noticed that prime phonemes were occasionally repeated. However, primes were only *partly* repeated, to unpredictable degrees across trials. To perform efficiently, participants should avail themselves of prime information, but must not "garden-path" too far into the prime while repeating the target. As suggested later, this dilemma may create a pattern of small positive and negative priming effects with partial overlap, as well as null repetition effects (in Slowiaczek & Hamburger, 1992).[1]

In essence, I suggest that Hamburger and Slowiaczek (1996) were not sufficiently vigilant in their study. Rather than assessing bias, they presumed its absence in the low-expectancy condition. Thus, interpretation requires faith that the low-expectancy parameters (21% related trials, 50-msec ISI) truly eliminate biases. If we agree to "see no evil," the priming effects seem relevant to models of spoken word perception. If not, the pattern is impossible to interpret. This raises an obvious question: Do the low-expectancy methods truly eliminate biased processing? In this regard, I conducted two experiments, here described together for brevity and clarity. Experiment 1 was designed to replicate Hamburger and Slowiaczek, using their procedures (with new stimuli). Presuming a successful replication, the true goal of Experiment 1 was to conduct *time-ordered* analyses on the control trials. If participants develop biases during the experiment, their control-trial performance should be systematically affected. Specifically, a bias intended to maximize performance in related trials may invoke a "cost" in control trials (Goldinger et al., 1992; Posner & Snyder, 1975).

Hamburger and Slowiazcek's (1996) main assumption was that biases would be negligible in the low-expectancy condition. The present Experiment 2 was a simple test of this assumption, again focusing on control trials. Whereas the low-expectancy condition of Experiment 1 contained 79% control trials, Experiment 2 contained 100% control trials. If the former condition is truly bias free, its control RTs should be equivalent to those of Experiment 2. However, if some bias remains, its control trials should be slow and may show a different time-ordered distribution relative to those of Experiment 2.

EXPERIMENTS 1 AND 2

Method

Participants. A total of 136 Arizona State University undergraduates participated for course credit. Experiment 1 included 40 and 56 students in the high- and low-expectancy conditions, respectively. Another 20 students participated in Experiment 2. All were native English speakers with no (self-reported) speech or hearing disorders.

Stimulus materials. The stimulus materials were 500 four-phoneme, monosyllabic words. One hundred target words were paired with primes sharing zero, one, two, or three phonemes (counting from initial position). All words were recorded (spoken in isolation) on digital audiotape in a sound-attenuated booth by a male speaker. The words were low-pass filtered at 4.8 kHz and digitized at a sampling rate of 10 kHz, using a 12-bit analog-to-digital converter. All words were excised from the recorded list using a speech waveform editor (CSRE) and were stored as digital files.

Procedure. In both experiments, participants were tested individually in a sound-attenuated room, seated before a Gateway 2000 computer with Beyer-Dynamic T100 headphones, a Beyer-Dynamic microphone, and a Gerbrands voice key. In each trial, a 500-Hz tone played for 100 msec, followed by a 250-msec silent interval. Afterward, a clear spoken prime was presented binaurally at approximately 70 dB. Upon prime offset, either a 50- or 500-msec ISI elapsed (depending on condition), followed by a clear spoken target. A timer began at target onset, and RTs were recorded when the voice key detected a sound above threshold. After the RT was recorded, a 2-sec interval elapsed, followed by the next trial. An experimenter manually recorded any errors.

In the high-expectancy condition of Experiment 1, every participant received 25 trials at each level of prime–target overlap (zero, one, two, and three phonemes). Across participants, all prime–target pairs were used equally. In the low-expectancy condition, every participant received 79 control trials and 21 related trials (7 per overlap level). Because this condition required odd numbers of related trials per cell, complete counterbalancing of 100 targets was not possible. To rectify this, 2 targets were selected as "permanent" control stimuli. The remaining 98 targets were organized into 14 counterbalanced lists, used equally across participants. In Experiment 2, only control pairs were used.

Results

Experiment 1. Before analysis, trials were deleted if they contained errors or RTs above 3 sec (such trials accounted for $<$ 3.5% of the data). The mean correct RTs are shown in Figure 1. For comparison, RTs from Hamburger and Slowiaczek (1996) are also shown. As shown, the replication was successful, although the present RTs were slightly faster. These data were first analyzed with a 2 (expectancy level) $\times$ 4 (overlap level) mixed-model analysis of variance (ANOVA). The most salient aspect of Figure 1 is the large (96-msec) RT difference between the low- and high-expectancy conditions [$F(1,94) = 203.50$, $MS_e = 39{,}021$]. (All tests assumed a $p < .05$ criterion.) An interaction of expectancy level $\times$ overlap level was also observed [$F(3,250) = 3.85$, $MS_e = 6{,}077$], as in Hamburger and Slowiaczek's (1996) data.

Within conditions, priming effects were assessed via t tests, comparing each overlap level (one, two, or three phonemes) to control (zero phonemes). In the high-expectancy condition, reliable facilitation (26 msec) was observed in the one-phoneme trials [$t(39) = 2.91$], and marginal facilitation (25 msec) was observed in the two-phoneme trials. An unreliable inhibitory trend (6 msec) was observed in the three-phoneme trials. In the low-expectancy condition, unreliable facilitatory trends were observed in the one- and two-phoneme trials (6 and 9 msec, respectively), but reliable inhibition (25 msec) emerged in the three-phoneme trials [$t(55) = 3.04$].

The key analyses involved control trials, which were grouped into five "time windows" of 20 trials each, shown in Figure 2. The mean time window RTs were analyzed in a 2 $\times$ 5 (expectancy level $\times$ time window) ANOVA,

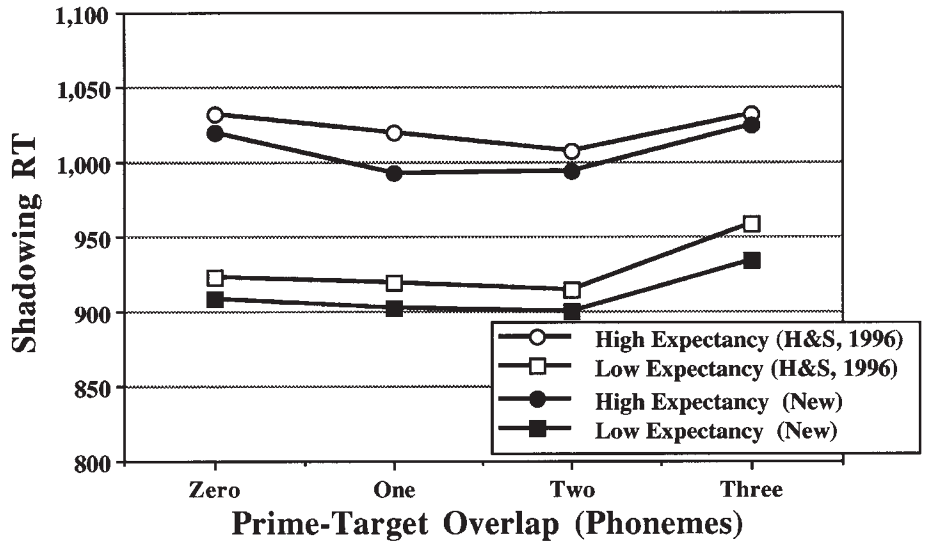

Figure 1. Mean shadowing response times (RTs) from all conditions of Experiment 1. For comparison, mean RTs from Hamburger and Slowiaczek (1996) are also shown.

followed by separate tests on each expectancy condition. In the overall analysis, robust effects of expectancy level [$F(1,94) = 121.04$, $MS_e = 28{,}011$] and time window [$F(4,376) = 68.50$, $MS_e = 26{,}911$] reflected the separate, rising RT functions. The high-expectancy condition produced a steeper rise across time windows than did the low-expectancy condition (comparison of first and last windows showed that these conditions produced 96- and 54-msec increases, respectively). Thus, the high-expectancy condition apparently created a stronger bias. However, the expectancy level × time window interaction was null [$F(4,376) = 1.71$], possibly reflecting the noisy high-expectancy data (a by-product of having fewer control trials). In separate tests, reliable time window effects were observed in both the high-expectancy [$F(4,156) = 22.09$, $MS_e = 36{,}088$] and low-expectancy [$F(4,220) = 49.61$, $MS_e = 21{,}007$] conditions, suggesting that biases affected each group.

Experiment 2. In Experiment 1, the mean RT to low-expectancy control trials was 909 msec. As noted earlier, Experiment 2 contained only control trials (and a 50-msec ISI), yielding a mean RT of 851 msec. This robust

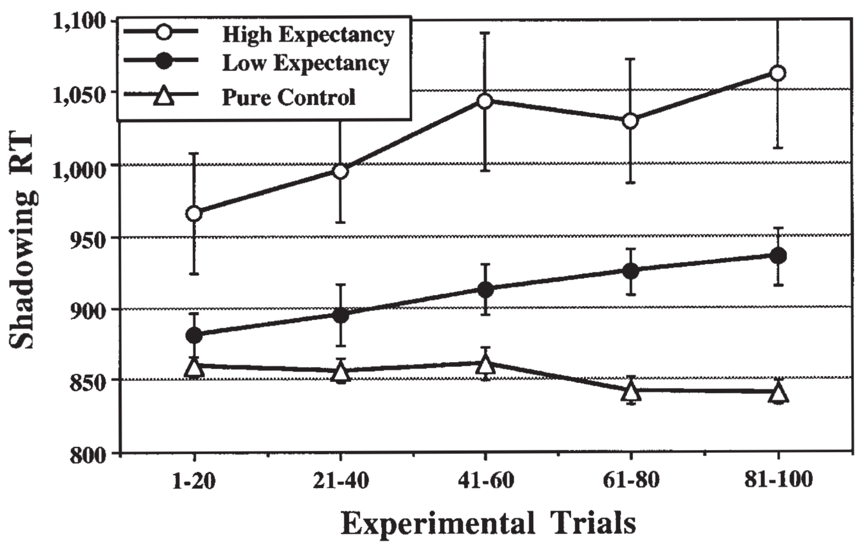

Figure 2. Mean control-trial response times (RTs) (and standard errors) from Experiments 1 and 2, shown as a function of 20-trial "time windows."

58-msec difference [$F(1,75) = 23.99$, $MS_e = 18{,}512$] suggests a lingering bias in the low-expectancy condition of Experiment 1. Beyond a difference in central tendency, the Experiment 2 trials also remained relatively constant over time windows, as shown by the "pure control" function in Figure 2. These data showed no time window effect [$F(4,76) = 0.51$, $MS_e = 4{,}011$, n.s.]. Moreover, when analyzed in tandem with the low-expectancy control trials from Experiment 1, these data produced a significant time window × experiment interaction [$F(4,284) = 20.65$, $MS_e = 9{,}908$]. The interaction seems to verify that the rising control-trial functions in Experiment 1 were not due to simple fatigue or habituation.

Discussion

Experiment 1 replicated Hamburger and Slowiaczek's (1996) results, but also revealed a "cost" in unrelated priming. This well-known bias signature (Neely, 1977; Posner & Snyder, 1975) was observed in both the high- and the low-expectancy conditions. To test this further, low-expectancy control trials from Experiment 1 were compared with pure control trials (Experiment 2), again providing evidence for a bias. Clearly, the low-expectancy procedures did not eliminate bias. As an epilogue to Experiment 1, all participants were interviewed by a research assistant. In answer to direct questions, virtually all students (40/40 in the high-expectancy condition; 55/56 in the low-expectancy condition) acknowledged noticing the shared phonemes between primes and targets. Interestingly, many (32/40 in the high-expectancy condition; 39/56 in the low-expectancy condition) apparently tried to *ignore* the primes, which were variously characterized as "distracting," "annoying," and "confusing." These complaints may reflect the unpredictable nature of the task—across trials, primes and targets are unrelated, slightly related, or closely related. Given such inconsistency, participants may struggle to avoid "false positives" (i.e., anticipating more prime phonemes than targets truly contain). As a result, related trials create small, bidirectional priming effects, and unrelated trials become slower over time.

The present participants' comments imply that Hamburger and Slowiaczek's (1996) method evokes an inefficient bias, which may explain a puzzling aspect of the data—why are responses dramatically *slower* in the high-expectancy condition? Given that 75% of trials follow the expected relationship, the high-expectancy condition should generate *faster* responses. The erratic nature of the experiment (across trials) seems responsible. In another test, I repeated Hamburger and Slowiaczek's experiment, using consistent levels of phonological overlap within blocks of trials. With this simple change, participants became far more efficient—all responses became faster, all priming effects became facilitatory, and control trials (added to each block) showed powerful "costs" that were commensurate with levels of phonological overlap per block.

Priming and bias. In phonological priming, researchers have abandoned "strategic" tasks (e.g., perceptual identification) in favor of shadowing. However, even shadowing involves top-down processing (Marslen-Wilson, 1985) and requires cautious interpretation. Indeed, nearly all priming tasks entail top-down processes, thus confounding the automatic and controlled components of behavior (den Heyer, Briand, & Dannenbring, 1983; Neely, 1977). This is especially true when facilitatory priming is expected—the hypothesized lexical processes and the human participant are both expected to maximize performance. As a result, net priming effects are difficult to partition into theoretically relevant components. To fix this problem, participants must be unaware of the priming relationship, or their "learned" and "natural" responses can be placed in mutual opposition. With respect to the first approach, making participants unaware of priming relationships is quite challenging, as demonstrated here. Thus, some investigators have resorted to subliminal presentation of visual primes (Balota, 1983; Fischler, 1977; Forster & Davis, 1991; Fowler, Wolford, Slade, & Tassinary, 1981; Marcel, 1983). Of course, this method also requires precautions (Holender, 1986), and it may not generalize to spoken primes.

The second approach—setting components to cross purposes—may hold more promise in studies of spoken word perception. In a semantic priming study, Neely (1977) used natural associates (e.g., BIRD–*robin*) as primes and targets, separated by varying ISIs. However, for some category names (e.g., BODY), participants were advised that targets would belong to a different category (e.g., building parts). Because this usually held true, participants seeing the prime BODY learned to expect "building" words (e.g., *door*) as targets. But, contrary to instructions, natural associates (e.g., *arm*) were occasionally shown. The data showed separate priming and bias effects: At short ISIs, participants showed facilitated processing of natural prime–target pairs (e.g., BODY–*arm*). But at longer ISIs, this "normal" priming was reversed—RTs followed the instruction-derived bias, instead of real conceptual relations (see Jacoby, 1991, for a similar approach).

Beyond priming, copious research shows that biases pervade speech communication, such as listeners' use of sentence context to resolve lexical ambiguity. Accordingly, biases may be considered a natural aspect of perceptual processing (Ratcliff & McKoon, 1997). However, in priming research, biases are often considered a nuisance—participants try to please the experimenter, obscuring the "automatic processes" of true interest. Hamburger and Slowiaczek (1996) tried to eliminate such biases, but most likely failed. Given its obvious difficulty, perhaps the quest for "pure" priming should be reconsidered. Priming data testify to the adaptive nature of perceptual processes across changing contexts. If the ultimate goal is to understand lexical processes, it seems that biases merit direct investigation.

REFERENCES

Balota, D. A. (1983). Automatic semantic activation and episodic memory encoding. *Journal of Verbal Learning & Verbal Behavior*, **22**, 88-104.

den Heyer, K., Briand, K., & Dannenbring, G. L. (1983). Strategic factors in a lexical-decision task: Evidence for automatic and attention-driven processes. *Memory & Cognition*, **11**, 374-381.

Fischler, I. (1977). Associative facilitation without expectancy in a lexical decision task. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 18-26.

Forster, K. I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory & Language*, **30**, 1-25.

Fowler, C. A., Wolford, G., Slade, R., & Tassinary, L. (1981). Lexical access with and without awareness. *Journal of Experimental Psychology: General*, **110**, 341-362.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1211-1238.

Hamburger, M., & Slowiaczek, L. M. (1996). Phonological priming reflects lexical competition. *Psychonomic Bulletin & Review*, **3**, 520-525.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral & Brain Sciences*, **9**, 1-23.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology*, **15**, 197-237.

Marslen-Wilson, W. D. (1985). Speech shadowing and speech comprehension. *Speech Communication*, **4**, 55-73.

Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In G. Altmann (Ed.), *Cognitive models of speech processing* (pp. 148-172). Cambridge, MA: MIT Press.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, **106**, 226-254.

Posner, M. I., & Snyder, C. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V* (pp. 669-681). New York: Academic Press.

Radeau, M., Morais, J., & Dewier, A. (1989). Phonological priming in spoken word recognition: Task effects. *Memory & Cognition*, **17**, 525-535.

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 1297-1311.

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

Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1239-1250.

Slowiaczek, L. M., Nusbaum, H. C., & Pisoni, D. B. (1987). Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 64-75.

Slowiaczek, L. M., & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory & Cognition*, **14**, 230-237.

NOTE

1. This commentary addresses the Hamburger and Slowiaczek (1996) study, which did not include repetition trials. However, when repetition trials were included by Slowiaczek and Hamburger (1992), they produced no priming effect. Because repetition priming is usually a very robust phenomenon, this failure may indicate a flaw in their design. Moreover, the authors described a (nonimplemented) connectionist model to accommodate their effects. It is easy to verify—by derivation or simulation—that the model must predict repetition benefits, no matter how its parameters are set. The conspicuous absence of repetition priming is most easily explained by reference to an inefficient bias.

(Manuscript received September 17, 1996;
revision accepted for publication April 13, 1998.)